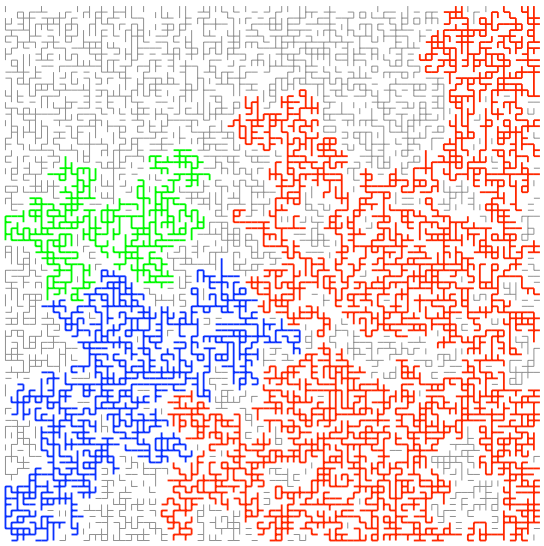
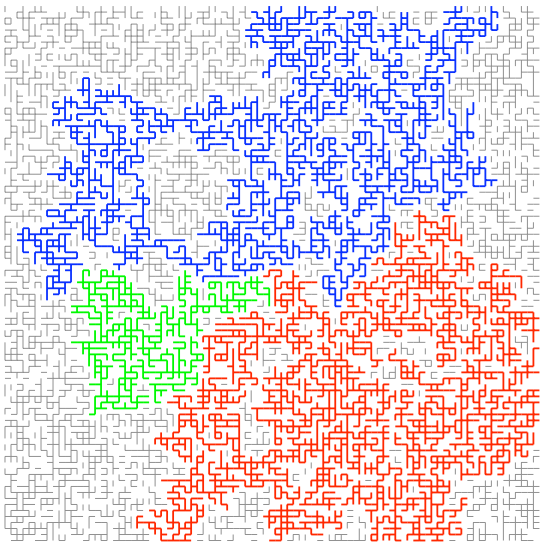
Lectures on noise sensitivity and percolation: preliminary version SECOND PART OF NOTES

Christophe Garban[1] Jeffrey E. Steif[2]





[1]ENS Lyon, CNRS
[2]Chalmers

Contents

Overview

The goal of this set of lectures is to combine two seemingly unrelated topics:

- The study of **Boolean functions**, a field particularly active in computer science
- Some models in statistical physics, mostly **percolation**

The link between these two fields can be loosely explained as follows: a percolation configuration is built out of a collection of i.i.d "bits" which determines whether the corresponding edges, sites, or blocks are present or absent. In that respect, any event concerning percolation can be seen as a Boolean function whose input are precisely these "bits".

Over the last 20 years, mainly thanks to the computer science community, a very rich structure has emerged concerning the properties of Boolean functions. The first part of this course will be devoted to a description of some of the main achievements in this field.

In some sense one can say, although this is an exaggeration, that computer scientists are mostly interested in the *stability* or *robustness* of Boolean functions. As we will see later in this course, the Boolean functions which 'encode' large scale properties of **critical** percolation will turn out to be very *sensitive* to small perturbations. This phenomenon corresponds to what we will call **noise sensitivity**. Hence, the Boolean functions one wishes to describe here are in some sense *orthogonal* to the Boolean functions one encounters, ideally, in computer science. Remarkably, it turns out that the tools developed by the computer science community to capture the properties and stability of Boolean functions are also suitable for the study of noise sensitive functions. This is why it is worth us first spending some time on the general properties of Boolean functions.

One of the main tools needed to understand properties of Boolean functions is Fourier analysis on the hypercube. Noise sensitivity will correspond to our Boolean function being of 'High Frequency' while stability will correspond to our Boolean function being of 'Low Frequency' . We will apply these ideas to some other models from statistical mechanics as well; namely, first passage percolation and dynamical percolation.

Some of the different topics here can be found (in a more condensed form) in [Gar10]. Also, updated versions of these lecture notes will be posted on `http://www.umpa.ens-lyon.fr/~cgarban`

Chapter I

Boolean functions and key concepts

1 Boolean functions

Definition I.1. A **Boolean function** is a function from the hypercube $\Omega_n := \{-1,1\}^n$ into either $\{-1,1\}$ or $\{0,1\}$.

Ω_n will be endowed with the uniform measure $\mathbb{P} = \mathbb{P}^n = (\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$ and $\mathbb{E}$ will denote the corresponding expectation. At various times, Ω_n will be endowed with the general product measure $\mathbb{P}_p = \mathbb{P}_p^n = ((1-p)\delta_{-1} + p\delta_1)^{\otimes n}$ but in such cases the p will be explicit. $\mathbb{E}_p$ will then denote the corresponding expectations.

An element of Ω_n will be denoted by either ω or ω_n and its n bits by $x_1, \ldots, x_n$ so that $\omega = (x_1, \ldots, x_n)$.

Depending on the context, concerning the range, it might be more pleasant to work with one of $\{-1,1\}$ or $\{0,1\}$ rather than the other and at some specific places in these lectures, we will even relax the Boolean constraint (i.e. taking only two possible values). In these cases (which will be clearly mentioned), we will consider instead **real-valued** functions $f : \Omega_n \to \mathbb{R}$.

A Boolean function f is canonically identified with a subset A_f of Ω_n via $A_f := \{\omega : f(\omega) = 1\}$.

Remark I.1. Often, Boolean functions are defined on $\{0,1\}^n$ rather than $\Omega_n = \{-1,1\}^n$. This does not make any fundamental difference at all but as we will see later, the choice of $\{-1,1\}^n$ turns out to be more convenient when one wishes to apply Fourier analysis on the hypercube.

2 Some Examples

We begin with a few examples of Boolean functions. Others will appear throughout this chapter.

Example 1 (Dictatorship)**.**

$$\mathbf{DICT}_n(x_1, \ldots, x_n) := x_1$$

The first bit determines what the outcome is.

Example 2 (Parity)**.**

$$\mathbf{PAR}_n(x_1, \ldots, x_n) := \prod_{i=1}^{n} x_i$$

This boolean function tells whether the number of -1's is even or odd.

These two examples are in some sense trivial, but they are good to keep in mind since in many case they turn out to be the "extreme cases" for properties concerning Boolean functions.

The next rather simple Boolean function is of interest in social choice theory.

Example 3 (Majority function)**.** Let n be odd and define

$$\mathbf{MAJ}_n(x_1, \ldots, x_n) := \mathrm{sign}(\sum_{i=1}^{n} x_i)\,.$$

Following are two further examples which will also arise in our discussions.

Example 4 (Iterated 3-Majority function)**.** Let $n = 3^k$ for some integer k. The bits are indexed by the leaves of a rooted 3-ary tree (so the root has degree 3, the leaves have degree 1 and all others have degree 4) with depth k. One iteratively applies the previous example (with $n = 3$) to obtain values at the vertices at level $k-1$, then level $k-2$, etc. until the root is assigned a value. The root's value is then the output of f. For example when $k = 2$, $f(-1, 1, 1; 1, -1, -1; -1, 1, -1) = -1$. The recursive structure of this Boolean function will enable explicit computations for various properties of interest.

Example 5 (Clique containnment)**.** If $r = \binom{n}{2}$ for some integer n, then Ω_r can be identified with the set of labelled graphs on n vertices. (x_i is 1 iff the ith edge is present.) Recall that a **clique** of size k of a graph $G = (V, E)$ is a complete graph on k vertices embedded in G.

Now for any $1 \leq k \leq \binom{n}{2} = r$, let $\mathbf{CLIQ}_n^k$ be the indicator function of the event that the random graph G_ω defined by $\omega \in \Omega_r$ contains a clique of size k. For most values of k the Boolean function $\mathbf{CLIQ}_n^k$ is very degenerate (i.e. has very small variance). Choosing $k = k_n$ so that this Boolean function is non-degenerate turns out to be a rather delicate issue. The interesting regime is near $k_n \approx 2\log_2(n)$. See the exercises for this "tuning" of $k = k_n$.

3 Pivotality and Influence

This section contains our first fundamental concepts.

Definition I.2. Given a Boolean function f from Ω_n into either $\{-1,1\}$ or $\{0,1\}$ and a variable $i \in [n]$, we say that i **is pivotal for** f **for** ω if $\{f(\omega) \neq f(\omega^i)\}$ where ω^i is ω but flipped in the ith coordinate. Note that this event is measurable with respect to $\{x_j\}_{j \neq i}$.

Definition I.3. The **pivotal set**, $\mathcal{P}$, of f is the random set of $[n]$ given by

$$\mathcal{P}(\omega) = \mathcal{P}_f(\omega) := \{i \in [n] : \ i \text{ is pivotal for } f \text{ for } \omega\}.$$

In words, it is the (random) set of bits with the property that if you flip the bit, then the function changes.

Definition I.4. The **influence** of the ith bit, $\mathbf{I}_i(f)$, is defined by

$$\mathbf{I}_i(f) := \mathbb{P}(i \in \mathcal{P}) = \mathbb{P}(\ i \text{ is pivotal for } f\ ).$$

Let also the **influence vector**, $\mathbf{Inf}(f)$, be the collection of all the influences: i.e. $\{\mathbf{I}_i(f)\}_{i \in [n]}$.

In words, the influence of the ith bit, $\mathbf{I}_i(f)$, is the probability that on flipping this bit, the function value changes.

Definition I.5. The **total influence**, $\mathbf{I}(f)$, is defined by

$$\mathbf{I}(f) := \sum_i \mathbf{I}_i(f) \ = \|\mathbf{Inf}(f)\|_1 \ (= \mathbb{E}(|\mathcal{P}|)).$$

It would now be instructive to go and compute these quantities for examples 1–3. See the exercises.

Later, we will need the last two concepts in the context when our probability measure is $\mathbb{P}_p$ instead. We give the corresponding definitions.

Definition I.6. The **influence vector at level** p, $\{\mathbf{I}_i^p(f)\}_{i \in [n]}$, is defined by

$$\mathbf{I}_i^p(f) := \mathbb{P}_p(i \in \mathcal{P}) = \mathbb{P}_p(\ i \text{ is pivotal for } f\ ).$$

Definition I.7. The **total influence at level** p, $\mathbf{I}^p(f)$, is defined by

$$\mathbf{I}^p(f) := \sum_i \mathbf{I}_i^p(f) \ (= \mathbb{E}_p(|\mathcal{P}|)).$$

It turns out that the total influence has a geometric-combinatorial interpretation as the size of the so-called Edge-Boundary. See the exercises.

Remark I.2. Aside from its natural definition as well as its geometric interpretation as measuring the edge-boundary of the corresponding subset of the hypercube (see the exercises), the notion of *Total influence* arises very naturally when one studies **sharp thresholds** for *monotone functions* (to be defined in Chapter III). Roughly speaking, as we will see in detail in Chapter III, for a monotone event A, one has that $d\mathbb{P}_p\big[A\big]/dp$ is the Total influence at level p (this is the Margulis-Russo formula). This tells us that the speed at which one changes from the event A "almost surely" **not** occurring to the case where it "almost surely" **does** occur is very sudden if the Boolean function happens to have a large total influence.

4 Kahn, Kalai, Linial Theorem

This section addresses the following question. Does there always exist some variable i with (reasonably) large influence? In other words, for large n, what is the smallest value (as we vary over Boolean functions) that the largest influence (over the different variables) can take on?

Since for the constant function, all influences are 0 and the function which is 1 only if all the bits are 1 has all influences $1/2^{n-1}$, clearly one wants to deal with functions which are reasonably balanced (or alternatively, obtain lower bounds on the maximal influence in terms of the variance of the Boolean function).

The first result in this direction is the following result. A sketch of the proof is given in the exercises.

Theorem I.1 (Discrete Poincaré)**.** *If f is a Boolean function mapping Ω_n into $\{-1,1\}$, then*

$$\mathrm{Var}(f) \leq \sum_i \mathbf{I}_i(f).$$

It follows that there exists some i such that

$$\mathbf{I}_i(f) \geq \mathrm{Var}(f)/n.$$

This gives a first answer to our question. For balanced functions (i.e., those whose variances are not so close to 0), there is some variable whose influence is at least of order $1/n$. **Can we find a better "universal" lower bound on the maximal influence ?** Note that for Example 3, all the influences are of order $1/\sqrt{n}$ (and the variance is 1). In terms of our question, this universal lower bound one is looking for should lie somewhere between $1/n$ and $1/\sqrt{n}$. The following celebrated result improves by a logarithmic factor on the above $\Omega(1/n)$ bound.

Theorem I.2 ([KKL88])**.** *There exists a universal $c > 0$ such that if f is a Boolean function mapping Ω_n into $\{0,1\}$, then there exists some i such that*

$$\mathbf{I}_i(f) \geq c\mathrm{Var}(f)\log n/n.$$

What is remarkable about this theorem is that this "logarithmic" lower bound on the maximal influence turns out to be **sharp**! This is shown by the following example:

Example 6 (Tribes)**.** Partition $[n]$ into subsequent blocks of length $\log_2(n) - \log_2(\log_2(n))$ with perhaps some leftover debris. Define $f = f_n$ to be 1 if there exists at least one block which contains all 1's, and 0 otherwise.

It turns out that one can check that the sequence of variances stays bounded away from 0 and 1 and that all the influences (including those belonging to the debris which are equal to 0) are smaller than $c \log n/n$ for some $c < \infty$. See the exercises for this. Hence the above theorem is indeed sharp.

Our next result tells us that if all the influences are 'small', then the total influence is large.

Theorem I.3 ([KKL88])**.** *There exists a $c > 0$ such that if f is a Boolean function mapping Ω_n into $\{0,1\}$ and $\delta := \max_i \mathbf{I}_i(f)$ then*

$$\mathbf{I}(f) \geq c \operatorname{Var}(f) \log(1/\delta).$$

Or equivalently,

$$\|\mathbf{Inf}(f)\|_1 \geq c \operatorname{Var}(f) \log \frac{1}{\|\mathbf{Inf}(f)\|_\infty}.$$

One can talk about the influence of a set rather than the influence of a variable.

Definition I.8. Given $S \subseteq [n]$, the **influence of** S, $\mathbf{I}_S(f)$, is defined by

$$\mathbf{I}_S(f) := \mathbb{P}(\ f \text{ is not determined by the bits in } S^c)$$

It is easy to see that when S is a single bit, this corresponds to our previous definition. The following is also proved in [KKL88]. We will not indicate the proof of this result in these lecture notes.

Theorem I.4 ([KKL88])**.** *Given a sequence f_n of Boolean functions mapping Ω_n into $\{0,1\}$ such that $0 < \inf_n \mathbb{E}_n(f) \leq \sup_n \mathbb{E}_n(f) < 1$ and any sequence a_n going to ∞ arbitrarily slowly, then there exist a sequence of sets $S_n \subseteq [n]$ such that $|S_n| \leq a_n n/\log(n)$ and $\mathbf{I}_{S_n}(f_n) \to 1$ as $n \to \infty$.*

Theorems I.2 and I.3 will be proved in Chapter V.

5 Noise Sensitivity and Noise Stability

This subsection introduces our second set of fundamental concepts.

Let ω be uniformly chosen from Ω_n and let ω_ϵ be ω but with each bit independently "rerandomized" with probability ϵ. "Rerandomized" means that each bit (independently of everything else) rechooses whether it is 1 or -1, each with probability 1/2. (Note that ω_ϵ then has the same distribution as ω.)

The following definition is *central* for these lecture notes. Let m_n be an increasing sequence of integers and let $f_n : \Omega_{m_n} \to \{\pm 1\}$ or $\{0, 1\}$.

Definition I.9. The sequence $\{f_n\}$ is **noise sensitive** if for every $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{E}[f_n(\omega) f_n(\omega_\epsilon)] - \mathbb{E}[f_n(\omega)]^2 = 0. \tag{I.1}$$

Since f_n just takes 2 values, this says that the random variables $f_n(\omega)$ and $f_n(\omega_\epsilon)$ are asymptotically independent for $\epsilon > 0$ fixed and n large. We will see later that (I.1) holds for one value of $\epsilon \in (0, 1)$ if and only if it holds for all such ϵ. The following notion captures the opposite situation where the two events above are close to being the same event if ϵ is small, uniformly in n.

Definition I.10. The sequence $\{f_n\}$ is **noise stable** if

$$\lim_{\epsilon\to 0} \sup_n \mathbb{P}(f_n(\omega) \neq f_n(\omega_\epsilon)) = 0.$$

It is an easy exercise to check that a sequence $\{f_n\}$ is both noise sensitive and noise stable if and only it is degenerate in the sense that the sequence of variances $\{\mathrm{Var}(f_n)\}$ goes to 0.

It is also an easy exercise to check that Example 1 (dictator) is noise stable and Example 2 (parity) is noise sensitive. We will see later (when Fourier analysis is brought into the picture) that these examples are the two opposite extreme cases.

For the other examples, it turns out that Example 3 (Majority) is noise stable, while Examples 4–6 are all noise sensitive. See the exercises.

Finally, note that a sequence of Boolean functions could be neither noise sensitive nor noise stable (see the exercises).

On figure 5, next page, we give a slightly *impressionistic* view of what "noise sensitivity" is.

6 Benjamini, Kalai and Schramm Noise Sensitivity Theorem

The following is the main theorem concerning noise sensitivity.

Theorem I.5 ([BKS99])**.** *If*

$$\lim_n \sum_k \mathbf{I}_k(f_n)^2 = 0,$$

then $\{f_n\}$ *is noise sensitive.*

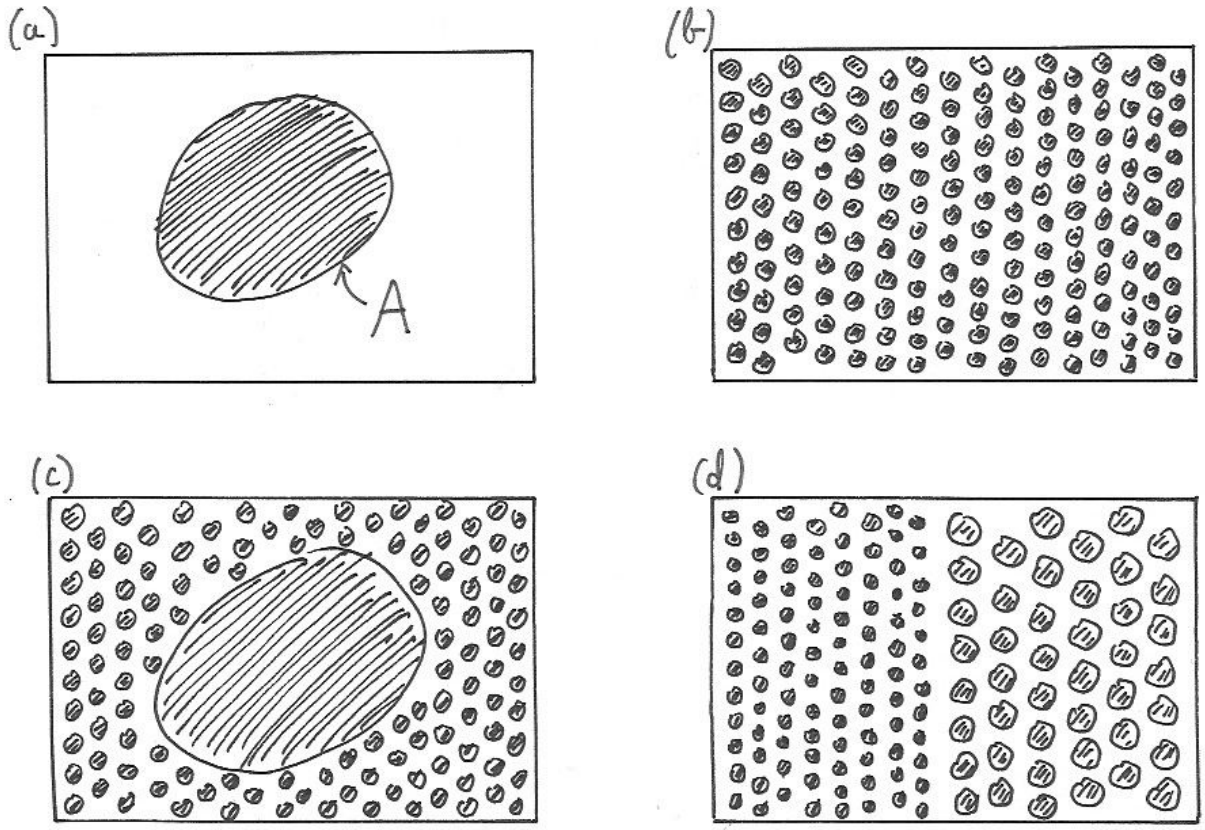


Figure I.1: Let us consider the following **"experiment"**: take a bounded domain in the plane, say a rectangle, and consider a measurable subset A of this domain. What would be an analog of the above definitions of being *noise sensitive* or *noise stable* in this case ? Start by sampling a point x uniformly in the domain according to Lebesgue measure. Then let us apply some noise to this position x so that we end-up with a new position x_ϵ. One can think of many natural "noising" procedures here. For example, let x_ϵ be a uniform point in the ball of radius ϵ around x, conditioned to remain in the domain. (This is not quite perfect yet since this procedure does not exactly preserve Lebesgue measure, but let's not worry here.) The natural analog of the above definitions is to ask whether $1_A(x)$ and $1_A(x_\epsilon)$ are decorrelated or not.
Question: According to this analogy, discuss the stability v.s. sensitivity of the sets A sketched in pictures (a) to (d) ? Note that in order to match with definitions I.9 and I.10, one should have considered sequences of subsets $\{A_n\}$ instead, since noise sensitivity is more of an asymptotic notion.

Remark I.3. The converse is clearly false as shown by Example 2. However, it turns out that the converse is true for so-called **monotone functions** (see the next chapter for the definition of this) as we will see in Chapter IV.

This theorem will allow us to conclude noise sensitivity of many of the examples we have introduced in this first chapter. See the exercises.

This theorem will also be proved in Chapter V.

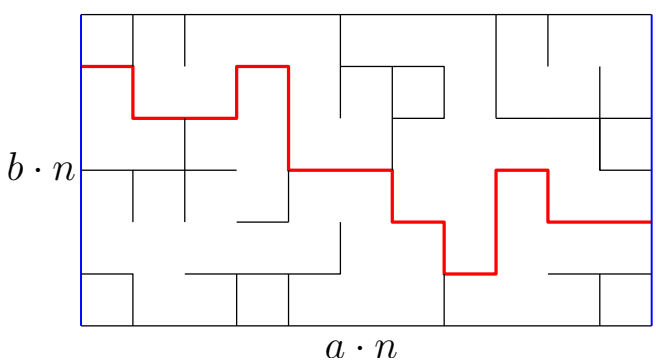
7 Percolation crossing: our final and most important example

We have saved our most important example to the end. This set of notes would not be being written if it were not for this example and for the results that have been proved for it.

Let us consider percolation on $\mathbb{Z}^2$ at the critical point $p_c(\mathbb{Z}^2) = 1/2$. (See Chapter II for a fast review on the model). At this critical point, there is no infinite cluster, but somehow clusters are 'large' (there are clusters at all scales). This can be seen using duality or with the RSW Theorem II.1. In order to understand the geometry of the critical picture, the following large-scale *observables* turn out to be very useful: Let Ω be a piece-wise smooth domain with two disjoint open arcs ∂_1 and ∂_2 on its boundary $\partial\Omega$. For each $n \geq 1$, we consider the scaled domain $n\Omega$. Let A_n be the event that there is an open path in ω from $n\partial_1$ to $n\partial_2$ which stays inside $n\Omega$. Such events are called **crossing events**. They are naturally associated with Boolean functions whose entries are the set of edges inside $n\Omega$ (there are $O(n^2)$ such variables).

For simplicity, let us consider the particular case of rectangle crossings:

Example 7 (Percolation crossings)**.**



Let $a, b > 0$ and let us consider the rectangle $[0, a \cdot n] \times [0, b \cdot n]$. The left to right crossing event corresponds to the Boolean function $f_n : \{-1, 1\}^{O(1)n^2} \to \{0, 1\}$ defined as follows

$$f_n(\omega) := \begin{cases} 1 & \text{if there is a left-right crossing} \\ 0 & \text{else} \end{cases}$$

We will later prove that this sequence of Boolean functions $\{f_n\}$ is noise sensitive. This means that if a percolation configuration $\omega \sim \mathbb{P}_{p_c=1/2}$ is given to us, one cannot predict anything about the large scale clusters of the slightly perturbed percolation configuration ω_ϵ (where only an ϵ-fraction of the edges have been resampled).

Remark I.4. The same statement holds for the above more general crossing events (i.e. in $(n\Omega, n\partial_1, n\partial_2)$).

Exercise sheet of Chapter I

Exercise 1. Determine the pivotal set, the influence vector and the total influence for Examples 1–3.

Exercise 2. Determine the influence vector for Example 4 (iterated 3-majority) and Example 6 (tribes).

Exercise 3. Show that in Example 6 (tribes), the variances stay bounded away from 0. If the blocks are taken to be of size $\log_2(n)$ instead, show the influences would all be of order $1/n$. Why does this not contradict the KKL theorem?

Exercise 4. Ω_n has a graph structure where two elements are neighbors if they differ in exactly one location. The **edge boundary** of a subset $A \subseteq \Omega_n$, denoted by $\partial_E(A)$, is the set of edges where exactly one of the endpoints are in A.

Show that for any Boolean function, $\mathbf{I}(f) = |\partial_E(A_f)|/2^{n-1}$.

Exercise 5. Prove Theorem I.1. This is a type of Poincaré inequality. **Hint:** use the fact that $\mathrm{Var}(f)$ can be written $\frac{1}{2}\mathbb{E}\big[(f(\omega) - f(\widetilde{\omega}))^2\big]$, where $\omega, \widetilde{\omega}$ are independent and try to "interpolate" from ω to $\widetilde{\omega}$.

Exercise 6. Show that Example 3 (Majority) is noise stable.

Exercise 7. Prove that Example 4 (iterated 3-majority) is noise sensitive directly without relying on Theorem I.5. **Hint:** use the recursive the structure of this example in order to show that the criterion of noise sensitivity is satisfied.

Exercise 8. Prove that Example 6 (tribes) is noise sensitive directly without using Theorem I.5. (Here there is no recursive structure, so a more "probabilistic" argument is needed).

Problem 9. Recall Example 5 (clique containment).

(a) Prove that when $k_n = o(n^{1/2})$, $\mathbf{CLIQ}_n^{k_n}$ is asymptotically noise sensitive. **Hint:** start by obtaining an upper bound on the influences (which are identical for each edge). Conclude by using Theorem I.5.

(b) *Open exercise:* Find a more direct proof of this fact (in the spirit of exercise 8) which would avoid using Theorem I.5.

As pointed out after Example 5, for most values of $k = k_n$, the Boolean function $\mathbf{CLIQ}_n^{k_n}$ *happens to be very degenerate. The purpose of the rest of this problem is to determine what is the interesting regime where* $\mathbf{CLIQ}_n^{k_n}$ *has a chance of being non-degenerate (i.e. variance bounded away from 0). The rest of this exercise is somewhat tangential to the course.*

(c) If $1 \leq k \leq \binom{n}{2} = r$, what is the expected number of cliques in G_ω, $\omega \in \Omega_r$?

(d) Explain why there should be at most one choice of $k = k_n$ such that the variance of $\mathbf{CLIQ}_n^{k_n}$ remains bounded away from 0 ? (no rigorous proof required.) Describe this choice of k_n. Check that it is indeed in the regime $2\log_2(n)$.

(e) Note retrospectively that in fact for any choice of $k = k_n$, $\mathbf{CLIQ}_n^{k_n}$ is noise sensitive.

Exercise 10. Deduce from Theorem I.5 that both Example 4 (iterated 3-majority) and Example 6 (tribes) are noise sensitive.

Exercise 11. Give a sequence of Boolean functions which is neither noise sensitive nor noise stable.

Exercise 12. In the sense of Definition I.8, show that for the majority function and for fixed ϵ, any set of size $n^{1/2+\epsilon}$ has influence approaching 1 while any set of size $n^{1/2-\epsilon}$ has influence approaching 0.

Problem 13. Do you think a *"generic"* Boolean function would be stable or sensitive? Justify your intuition.

Chapter II

Percolation in a nutshell

In this chapter, we review some facts about percolation. Most of these topics shall be explained by Vincent Beffara in his course. Nevertheless, in order to make these lecture notes as self-contained as possible, we include here a short summary of the main useful results.

Besides the lecture notes for the present summer school written by Vincent Beffara, for a complete account on percolation, see [Gri99] and more specifically in our context the lecture notes [Wer07].

1 The model

Let us briefly start by the model itself.

We will be concerned mainly in two-dimensional percolation and we will focus on two lattices: $\mathbb{Z}^2$ and the triangular lattice $\mathbb{T}$.

Let us describe the model on $\mathbb{Z}^2$. Let $\mathbb{E}^2$ denote the set of edges of the graph $\mathbb{Z}^2$. For any $p \in [0,1]$ we define a random subgraph of $\mathbb{Z}^2$ as follows: independently for each edge $e \in \mathbb{E}^2$, we keep this edge with probability p and remove it with probability $1-p$. Equivalently, this corresponds to defining a random configuration $\omega \in \{-1,1\}^{\mathbb{E}^2}$ where, independently for each edge $e \in \mathbb{E}^2$, we declare the edge to be open ($\omega(e) = 1$) with probability p or closed ($\omega(e) = -1$) with probability $1-p$. The law of the so-defined random subgraph (or configuration) is denoted by $\mathbb{P}_p$.

Percolation is defined similarly on the triangular grid $\mathbb{T}$, except that on this lattice we will rather consider **site** percolation (i.e. here we keep each **site** with probability p). The sites are the points $\mathbb{Z} + e^{i\pi/3}\mathbb{Z}$ so that neighboring sites have distance one from each other in the complex plane.

2 Russo-Seymour-Welsh

We will often rely on the following celebrated result known as the **RSW Theorem**.

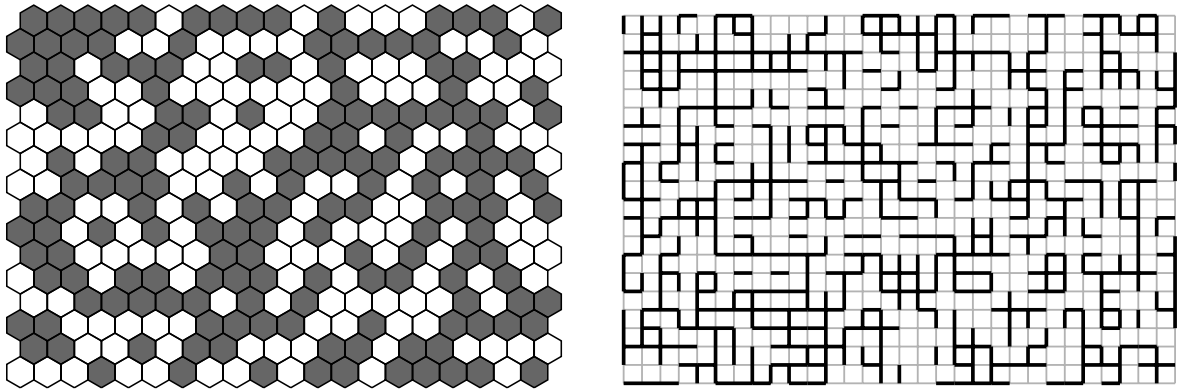


Figure II.1: Pictures (by Oded Schramm) representing two percolation configurations respectively on $\mathbb{T}$ and on $\mathbb{Z}^2$ (both at $p = 1/2$). The sites of the triangular grid are represented by hexagons.

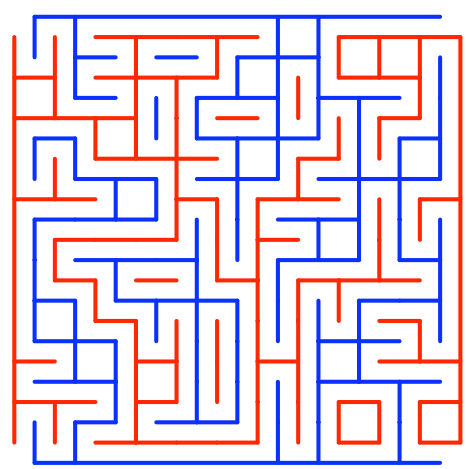
Theorem II.1 (RSW)**.** *For percolation on $\mathbb{Z}^2$ at $p = 1/2$, one has the following property concerning the* **crossing events**. *Let $a, b > 0$. There exists a constant $c = c(a, b) > 0$, such that for any $n \geq 1$, if A_n denotes the event that there is a left to right crossing in the rectangle $([0, a \cdot n] \times [0, b \cdot n]) \cap \mathbb{Z}^2$, then*

$$c < \mathbb{P}_{1/2}\big[A_n\big] < 1 - c\,.$$

In other words, this says that the Boolean functions f_n defined in Example 7 of Chapter I are non-degenerate.

The same result holds also in the case of site-percolation on $\mathbb{T}$ (also at $p = 1/2$).

The parameter $p = 1/2$ plays a very special role for the two models under consideration. Indeed, there is a natural way to associate to each percolation configuration $\omega_p \sim \mathbb{P}_p$ a dual configuration ω_{p^*} on the so-called dual graph. In the case of $\mathbb{Z}^2$, its dual graph can be realized as $\mathbb{Z}^2 + (\frac{1}{2}, \frac{1}{2})$. In the case of the triangular lattice, $\mathbb{T}^* = \mathbb{T}$. The figure on the right illustrates this **duality** for percolation on $\mathbb{Z}^2$. It is easy to see that in both cases $p^* = 1 - p$. Hence at $p = 1/2$, our two models happen to be **self-dual**.

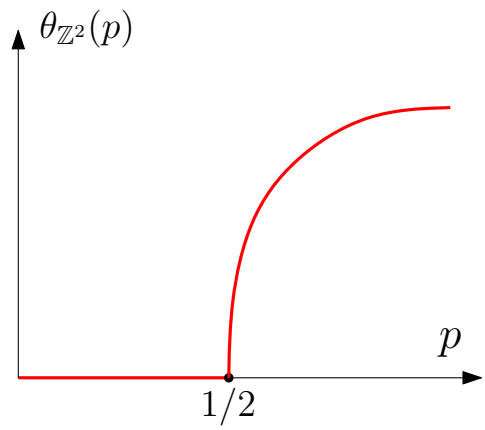


This duality has the following very important consequence. For a domain in $\mathbb{T}$ with two specified boundary arcs, there is a 'left-right' crossing of white hexagons if and only if there is no 'top-bottom' crossing black hexagons.

3 Phase transition

In percolation theory, one is interested in large scale connectivity properties of the random configuration $\omega = \omega_p$. In particular as one raises the level p, above a certain critical parameter $p_c(\mathbb{Z}^2)$, an infinite cluster (almost surely) emerges. This corresponds to the well-known *phase transition* of percolation. By a famous theorem of Kesten this transition takes place at $p_c(\mathbb{Z}^2) = \frac{1}{2}$. On the triangular grid, one has also $p_c(\mathbb{T}) = 1/2$.

This phase transition can be measured with the *density function* $\theta_{\mathbb{Z}^2}(p) := \mathbb{P}_p(0 \overset{\omega}{\longleftrightarrow} \infty)$ which encodes important properties of the large scale connectivities of the random configuration ω: it corresponds to the density averaged over the space $\mathbb{Z}^2$ of the (almost surely unique) infinite cluster. The shape of the function $\theta_{\mathbb{Z}^2}$ is pictured on the right (notice the infinite derivative at p_c).



4 Conformal invariance at criticality and SLE processes

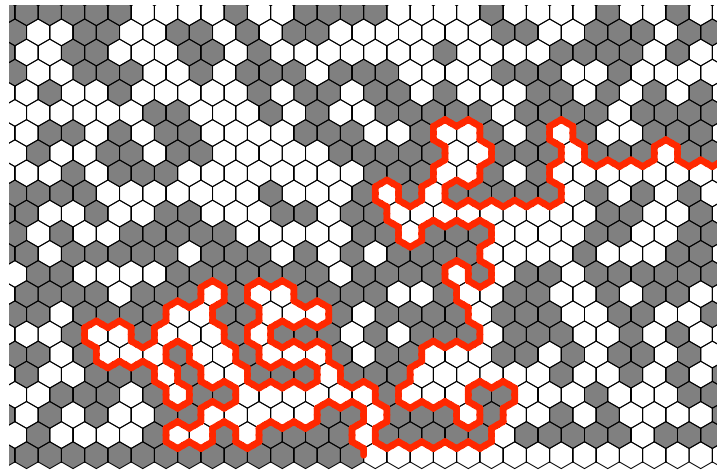
It has been conjectured for a long time that percolation should be *asymptotically* conformally invariant at the critical point. This should be understood in the same way as the fact that a Brownian motion (ignoring its time-parametrization) is a conformally invariant probabilistic object. One way to picture this conformal invariance is as follows: consider the 'largest' cluster C_δ surrounding 0 in $\delta\mathbb{Z}^2 \cap \mathbb{D}$ and such that $C_\delta \cap \partial\mathbb{D} = \emptyset$. Now consider some other simply connected domain Ω containing 0. Let $\hat{C}_\delta$ be the largest cluster surrounding 0 in a critical configuration in $\delta\mathbb{Z}^2 \cap \Omega$ and such that $\hat{C}_\delta \cap \partial\Omega = \emptyset$. Now let ϕ be the conformal map from $\mathbb{D}$ to Ω such that $\phi(0) = 0$ and $\phi'(0) > 0$. Even though the random sets $\phi(C_\delta)$ and $\hat{C}_\delta$ do not lie on the same lattice, the conformal invariance principle claims that when $\delta = o(1)$, these two random clusters are very close **in law**.

Over the last decade, two major breakthroughs enabled a much better understanding of the critical regime of percolation:

- The invention of the SLE processes by Oded Schramm([Sch00]).
- The proof of conformal invariance on $\mathbb{T}$ by Stanislav Smirnov ([Smi01]).

We will not define the SLE processes in these notes. See the lecture notes by Vincent Beffara and references therein. The illustration below explains how SLE curves arise naturally in the percolation picture.

This celebrated picture (by Oded Schramm) represents an **exploration path** on the triangular lattice. This exploration path, which turns right when encountering black hexagons and left when encountering white ones, asymptotically converges towards SLE_6 (as the mesh size goes to 0).



Note that, on $\mathbb{Z}^2$ at $p_c = 1/2$, proving the conformal invariance is still a challenging open problem.

5 Critical exponents

The proof of conformal invariance combined with the detailed information given by the SLE_6 process enables one to obtain very precise information on the critical and *near-critical* behavior of $\mathbb{T}$-percolation. For instance it is known that on the triangular lattice, the density function $\theta_{\mathbb{T}}(p)$ has the following behavior near $p_c = 1/2$

$$\theta(p) = (p - 1/2)^{5/36+o(1)} ,$$

when $p \to 1/2+$ (see [Wer07]).

In the rest of these lectures, we will often rely on three types of percolation events: namely the *one-arm*, *two-arm* and *four-arm* events. They are defined as follows: for any radius $R > 1$, let A_R^1 be the event that the site 0 is connected to distance R by some open path. Next, let A_R^2 be the event that there are two "arms" of different colors from the site 0 (which itself can be of either color) to distance R away. Finally, let A_R^4 be the event that there are four "arms" of alternating color from the site 0 (which itself can be of either color) to distance R away (i.e. there are four connected paths, two open, two closed from 0 to radius R and the closed paths lie between the open paths). See figure II.2 for a realization of two of these events.

It was proved in [LSW02] that the probability of the one-arm event decays like

$$\mathbb{P}\big[A_R^1\big] := \alpha_1(R) = R^{-\frac{5}{48}+o(1)} .$$

For the two-arms and four-arms events, it was proved by Smirnov and Werner in [SW01] that these probabilities decay like

$$\mathbb{P}\big[A_R^2\big] := \alpha_2(R) = R^{-\frac{1}{4}+o(1)}$$

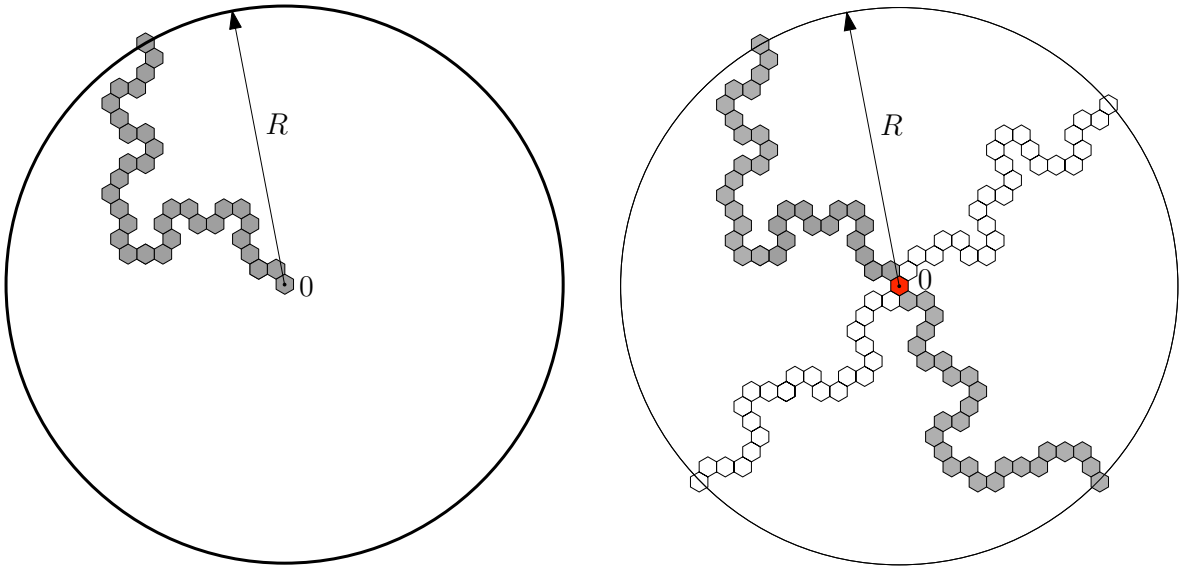


Figure II.2: A realization of the *one-arm* event is pictured on the left; the *four-arm* event is pictured on the right.

and

$$\mathbb{P}\big[A_R^4\big] := \alpha_4(R) = R^{-\frac{5}{4}+o(1)}\,.$$

Remark II.1. Note the $o(1)$ in the above statements (which means of course goes to zero as $R \to \infty$). Its presence reveals that the above critical exponents are known so far only up to 'logarithmic' corrections. It is conjectured that there are no such 'logarithmic' corrections, but at the moment one has to deal with their possible existence.

The three exponents we encountered concerning $\theta_{\mathbb{T}}$, α_1, α_2 and α_4 (i.e. $\frac{5}{36}$, $\frac{5}{48}$, $\frac{1}{4}$ and $\frac{5}{4}$) are known as *critical exponents*.

The *four-arm* event is clearly of particular relevance to us in these lectures. Indeed if a point x is in the 'bulk' of a domain $(n\Omega, n\partial_1, n\partial_2)$, the probability that this point is pivotal for the Left-Right crossing A_n (in other words its influence $\mathbf{I}_x(A_n)$) is intimately related to the four-arm event. See Chapter VI for more details.

6 Quasi-multiplicativity

Finally, let us end this overview by a type of scale invariance property of these arm-events. More precisely, it is often convenient to "divide" these arm-events into different scales. For this purpose, we introduce $\alpha_4(r, R)$ (with $r \leq R$) to be the probability that the four-arm event is realized from radius r to radius R ($\alpha_1(r, R)$ and $\alpha_2(r, R)$ are defined similarly). By independence on disjoint sets, it is clear that if $r_1 \leq r_2 \leq r_3$ then one has $\alpha_4(r_1, r_3) \leq \alpha_4(r_1, r_2)\,\alpha_4(r_2, r_3)$. A very useful property known as **quasi-multiplicativity** claims that up to constants, these two expressions are the same (this

makes the division into several scales practical). This property can be stated as follows.

Proposition II.2 (**quasi-multiplicativity**, [Kes87])**.** *For any $r_1 \leq r_2 \leq r_3$, one has (both for $\mathbb{Z}^2$ and $\mathbb{T}$ percolations)*

$$\alpha_4(r_1, r_3) \asymp \alpha_4(r_1, r_2)\, \alpha_4(r_2, r_3)\,,$$

where the constant involved in $\asymp$ are uniform constants.

See [Wer07, Nol09, SS10b] for more details. Note also that the same property holds for the one-arm event. However, this is much easier to prove: it is an easy consequence of the RSW Theorem II.1 and the so-called FKG inequality which says that increasing events are positively correlated.

Exercise sheet of Chapter II

Exercise 1. Prove using the RSW theorem that we have quasi-multiplicativity for the 1 arm event.

Chapter III

Sharp thresholds and the critical point for 2-d percolation

1 Monotone functions and the Margulus-Russo formula

The class of so-called monotone functions plays a very central role in this subject.

Definition III.1. A function f is **monotone** if $x \leq y$ (meaning $x_i \leq y_i$ for each i) implies that $f(x) \leq f(y)$. (An event is monotone if its indicator function is monotone.)

Recall that when the underlying variables are independent with 1 having probability p, we let $\mathbb{P}_p$ and $\mathbb{E}_p$ denote probabilities and expectations.

It is somewhat 'obvious' that for f monotone, $\mathbb{E}_p(f)$ should be increasing in p. The Margulus-Russo formula gives us an explicit formula for this (nonnegative) derivative.

Theorem III.1. *Let A be an increasing event in Ω_n. Then*

$$d(\mathbb{P}_p(A))/dp = \sum_i \mathbf{I}_i^p(A).$$

Proof.

Let us allow each variable x_i to have its own parameter p_i. Let $\mathbb{P}_{p_1,\ldots,p_n}$ and $\mathbb{E}_{p_1,\ldots,p_n}$ be the corresponding probability measure and expectation. It suffices to show that

$$\partial(\mathbb{P}_{(p_1,\ldots,p_n)}(A))/\partial p_i = \mathbf{I}_i^{(p_1,\ldots,p_n)}(A)$$

where the definition of this latter term is clear. WLOG, take $i = 1$. Now

$$\mathbb{P}_{(p_1,\ldots,p_n)}(A) = \mathbb{P}_{(p_1,\ldots,p_n)}(A \backslash \{1 \in \mathcal{P}_A\}) + \mathbb{P}_{(p_1,\ldots,p_n)}(A \cap \{1 \in \mathcal{P}_A\}).$$

The event in the first term is measurable with respect to the other variables and hence the first term does not depend on p_1 while the second term is

$$p_1 \mathbb{P}_{(p_2,\ldots,p_n)}(\{1 \in \mathcal{P}_A\})$$

since $A \cap \{1 \in \mathcal{P}_A\}$ is the event $\{x_1 = 1\} \cap \{1 \in \mathcal{P}_A\}$. □

2 KKL away from the uniform measure case

Recall now Theorem I.2. For sharp threshold results, one needs lower bounds on the total influence not just at the special parameter 1/2 but at all p.

The following are the two main results concerning the KKL result for general p that we will want to have at our disposal.

Theorem III.2 ([BKK+92]). *There exists a universal $c > 0$ such that for any Boolean function f mapping Ω_n into $\{0,1\}$ and for any p, there exists some i such that*

$$\mathbf{I}_i^p(f) \geq c\mathrm{Var}_p(f) \log n/n$$

Theorem III.3 ([BKK+92]). *There exists a universal $c > 0$ such that for any Boolean function f mapping Ω_n into $\{0,1\}$ and for any p,*

$$\mathbf{I}^p(f) \geq c\mathrm{Var}_p(f) \log(1/\delta_p)$$

where $\delta_p := \max_i \mathbf{I}_i^p(f)$.

3 Sharp Thresholds in General : Friedgut-Kalai Theorem

Theorem III.4 ([FK96]). *There exists a $c_1 < \infty$ such that for any monotone event A on n variables where all the influences are the same, if $\mathbb{P}_{p_1}(A) > \epsilon$, then*

$$\mathbb{P}_{p_1 + \frac{c_1 \log(1/(2\epsilon))}{\log(n)}}(A) > 1 - \epsilon.$$

Remark III.1. This says that for fixed ϵ, the probability of A moves from below ϵ to above $1-\epsilon$ in an interval of p of length of order at most $1/\log(n)$. The assumption of equal influences holds for example if the event is invariant under some transitive action which is often the case. The latter covers the case of the random graphs $G(n,p)$.

Proof. Theorem III.2 and all influences the same tells us that

$$\mathbf{I}(A) \geq c \min\{\mathbb{P}_p(A), 1 - \mathbb{P}_p(A)\} \log n$$

for some $c > 0$. Hence

$$d(\log(\mathbb{P}_p(A)))/dp \geq c \log n$$

if $\mathbb{P}_p(A) \leq 1/2$. Letting $p* := p_1 + \frac{\log(1/2\epsilon)}{c\log(n)}$, an easy computation (using the fundamental theorem of calculus) yields

$$\log(P_{p^*}(A)) \geq \log(1/2).$$

Next, if $\mathbb{P}_p(A) \geq 1/2$, then

$$d(\log(1 - \mathbb{P}_p(A)))/dp \leq -c \log n$$

from which another application of the fundamental theorem yields

$$\log(1 - P_{p^{**}}(A)) \leq -\log(1/\epsilon)$$

where $p^{**} := p^* + \frac{\log(1/2\epsilon)}{c\log(n)}$. Letting $c_1 = 2/c$ gives the result. □

4 The critical point for percolation for $\mathbb{Z}^2$ and $\mathbb{T}$ is $\frac{1}{2}$

Theorem III.5 ([Kes80])**.**

$$p_c(\mathbb{Z}^2) = p_c(\mathbb{T}) = \frac{1}{2}.$$

Proof.

We first show that $\theta(1/2) = 0$. Let $\mathrm{Ann}(\ell) := [-3\ell, 3\ell]\backslash[-\ell, \ell]$ and C_k be the event that there is a circuit in $Ann(4^k) + 1/2$ in the dual lattice around the origin consisting of closed edges. The C_k's are independent and RSW and FKG show that for some $c > 0$, $\mathbb{P}_{1/2}(C_k) \geq c$ for all k. This gives that $\mathbb{P}_{1/2}(C_k$ infinitely often$) = 1$ and hence $\theta(1/2) = 0$.

The next key step is a *finite criterion* which implies percolation and which is interesting in itself. We outline its proof afterwards.

Proposition III.6. *(Finite size criterion) J_n be the event that there is a crossing of a $2n \times (n-2)$ box. For any p, if there exists an n such that*

$$\mathbb{P}_p(J_n) \geq .98,$$

then $\mathbb{P}_p(|C(0)| = \infty) > 0$.

Assume now that $p_c = 1/2 + \delta$ with $\delta > 0$. Let $I = [1/2, 1/2 + \delta/2]$. Since $\theta(1/2 + \delta/2) = 0$, it is easy to see that the maximum influence over all variables and over all $p \in I$ goes to 0 with n since being pivotal implies the existence of an open path from a neighbor of the given edge to distance $n/2$ away. Next, by RSW, $\inf_n \mathbb{P}_{1/2}(J_n) > 0$. If for all n, $\mathbb{P}_{1/2+\delta/2}(J_n) < .98$, Theorems and III.1 and III.3 would allow us to conclude that the derivative of $\mathbb{P}_p(J_n)$ goes to ∞ uniformly on I as $n \to \infty$, giving a contradiction. Hence $\mathbb{P}_{1/2+\delta/2}(J_n) \geq .98$ for some n implying, by Proposition III.6, that $\theta(1/2 + \delta/2) > 0$, a contradiction. □

Outline of Proof of Proposition III.6.

The first step is to show that for any p and for any $\epsilon \leq .02$, if $\mathbb{P}_p(J_n) \geq 1 - \epsilon$, then $\mathbb{P}_p(J_{2n}) \geq 1 - \epsilon/2$. The idea is that by FKG and 'glueing' one can show that one can cross a $4n \times (n-2)$ box with probability at least $1 - 5\epsilon$ and hence one obtains that $\mathbb{P}_p(J_{2n}) \geq 1 - \epsilon/2$ since for this event to fail, it must fail in both the top and bottom halves of the box. It follows that if we place down (possibly rotated and translated) boxes of size $2n \times (n-2)$ anywhere, with probability 1, all but finitely are crossed. Finally, one can place these boxes down in an intelligent way such that crossing all but finitely many of them necessarily entails the existence of an infinite cluster. □

5 Further discussion

The Margulis-Russo formula is due independently to Margulis [Mar74] and Russo [Rus81].

The idea to use the results from KKL to show that $p_c = 1/2$ is due to Bollobas and Riordan (see [BR06]). It was understood much earlier that obtaining a sharp threshold was the key step. Kesten (see [Kes80]) showed the necessary sharp threshold by obtaining a lower bound on the expected number of pivotals in a hands on fashion. Russo (see [Rus82]) had developed an earlier weaker more qualitative version of KKL and showed how it also sufficed to show that $p_c = 1/2$.

Exercise sheet of Chapter III

Exercise 1. Develop an alternative proof of the Margulis-Russo formula using classical couplings.

Exercise 2. Study what are the **"threshold windows"** (i.e. where and how long does it take to go from a probability of order ϵ to a probability of order $1-\epsilon$) in the following examples:

(a) for $\mathbf{DICT}_n$

(b) for $\mathbf{MAJ}_n$

(c) for the **tribes** example

(d) for the **iterated majority** example.

(don't rely on [KKL88] type of results, but instead do hands-on computations specific to each case).

Exercise 3. Write out the details of the proof of Proposition III.6.

Problem 4 (*What is the "sharpest" monotone event ?)***.** Show that among all **monotone** Boolean functions on Ω_n, $\mathbf{MAJ}_n$ is the one with largest Total influence.
Hint: Use the Margulis-Russo formula.

Exercise 5. Find a monotone function $f : \Omega_n \to \{0,1\}$ such that $\partial_p \mathbb{E}_p(f)$ is very large at $p = 1/2$, but nevertheless there is no Sharp threshold for f (this means that a large Total influence at some value of p is not in general a sufficient condition for sharp threshold).

Chapter IV

Fourier analysis of Boolean functions (first facts)

1 Discrete Fourier analysis and the energy spectrum

It turns out that in order to understand and analyze the concepts previously introduced, which are in some sense purely probabilistic, a critical tool is Fourier analysis on the hypercube.

Recall that we consider our Boolean functions as functions from the hypercube $\Omega_n := \{-1,1\}^n$ into $\{-1,1\}$ or $\{0,1\}$ where Ω_n is endowed with the uniform measure $\mathbb{P} = \mathbb{P}^n = (\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)^{\otimes n}$.

In order to apply Fourier analysis, the natural setup is to enlarge our discrete space of Boolean functions and to consider instead the larger space $L^2(\{-1,1\}^n)$ of real-valued functions on Ω_n endowed with the inner product:

$$\begin{aligned} \langle f, g \rangle &:= \sum_{x_1,\ldots,x_n} 2^{-n} f(x_1,\ldots,x_n) g(x_1,\ldots,x_n) \\ &= \mathbb{E}\big[fg\big] \quad \text{for all } f,g \in L^2(\Omega_n)\,, \end{aligned}$$

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

For any subset $S \subseteq \{1,2\ldots,n\}$, let χ_S be the function on $\{-1,1\}^n$ defined for any $x = (x_1,\ldots,x_n)$ by

$$\chi_S(x) := \prod_{i \in S} x_i\,. \tag{IV.1}$$

(So $\chi_\emptyset \equiv 1$.) It is straightforward (check this!) to see that this family of 2^n functions forms an orthonormal basis of $L^2(\{-1,1\}^n)$. Thus, any function f on Ω_n (and a fortiori any Boolean function f) can be decomposed as

$$f = \sum_{S \subseteq \{1,\ldots,n\}} \hat{f}(S)\,\chi_S,$$

where $\{\hat{f}(S)\}_{S\subseteq[n]}$ are the Fourier coefficients of f. They are sometimes called the **Fourier-Walsh** coefficients of f and they satisfy

$$\hat{f}(S) := \langle f, \chi_S \rangle = \mathbb{E}\big[f\chi_S\big].$$

Note that $\hat{f}(\emptyset)$ corresponds to the average $\mathbb{E}\big[f\big]$. As in classical Fourier analysis, if f is some Boolean function, its Fourier(-Walsh) coefficients provide information on the "regularity" of f. We will sometime use the term *spectrum* when referring to the set of Fourier coefficients.

Of course one may find many other orthonormal bases for $L^2(\{-1,1\}^n)$, but there are many situations for which this particular set of functions $(\chi_S)_{S\subseteq\{1,\dots,n\}}$ arises naturally. First of all there is a well-known theory of Fourier analysis on groups, a theory which is particularly simple and elegant on Abelian groups (thus including our special case of $\{-1,1\}^n$, but also $\mathbb{R}/\mathbb{Z}$, $\mathbb{R}$ and so on). For Abelian groups, what turns out to be relevant for doing harmonic analysis is the set $\hat{G}$ of **characters** of G (i.e. the group homomorphisms from G to $\mathbb{C}^*$). In our case of $G = \{-1,1\}^n$, the characters are precisely our functions χ_S indexed by $S \subseteq \{1,\dots,n\}$ since they satisfy $\chi_S(x\cdot y) = \chi_S(x)\chi_S(y)$.

These functions also arise naturally if one performs simple random walk on the hypercube (equipped with the Hamming graph structure), since they are the eigenfunctions of the corresponding Markov chain (heat kernel) on $\{-1,1\}^n$. Last but not least, we will see later in the chapter that the basis (χ_S) turns out to be particularly adapted to our study of **noise sensitivity**.

We introduce one more concept here without motivation; it will be very well motivated later on in the chapter.

Definition IV.1. For any real-valued function $f : \Omega_n \to \mathbb{R}$, the **Energy Spectrum** E_f is defined by

$$E_f(m) := \sum_{|S|=m} \hat{f}(S)^2,\ \forall\, m \in \{1,\dots,n\}\,.$$

2 Examples

First note from the Fourier point of view, Dictator and Parity have simple representations since they are χ_1 and $\chi_{[n]}$ respectively. Each of the two corresponding energy spectrums are trivially concentrated on 1 point.

For Example 3, the Majority function, Bernasconi explicitly computed the Fourier coefficients and when n goes to infinity, one ends up with the following asymptotic formula:

$$E_{\mathbf{MAJ}_n}(m) = \sum_{|S|=m} \widehat{\mathbf{MAJ}}_n(S)^2 = \begin{cases} \frac{4}{\pi\, m 2^m}\binom{m-1}{\frac{m-1}{2}} + O(m/n) & \text{if } m \text{ is odd}\,, \\ 0 & \text{if } m \text{ is even}\,. \end{cases}$$

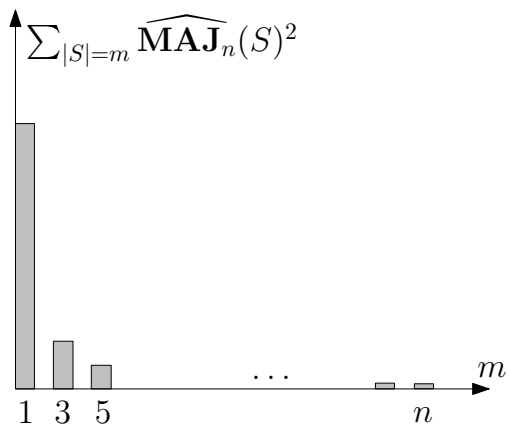


Figure IV.1: Shape of the energy spectrum for the Majority function

See [O'D03] for a nice overview and references therein concerning the spectral behavior of the majority function.

Picture IV.1 represents the shape of the Energy Spectrum of $\mathbf{MAJ}_n$: its Spectrum is concentrated on low frequencies which is typical of stable functions.

3 Noise sensitivity and stability in terms of the energy spectrum

In this section, we describe the concepts of noise sensitivity and noise stability in terms of the energy spectrum.

The first step is to note, given any real-valued function $f : \Omega_n \to \mathbb{R}$, the correlation between $f(\omega)$ and $f(\omega_\epsilon)$ is nicely expressed in terms of the Fourier coefficients of f as follows.

$$\begin{aligned}
\mathbb{E}\big[f(\omega)f(\omega^\epsilon)\big] &= \mathbb{E}\Big[\Big(\sum_{S_1} \hat{f}(S_1)\chi_{S_1}(\omega)\Big)\Big(\sum_{S_2} \hat{f}(S_2)\chi_{S_2}(\omega_\epsilon)\Big)\Big] \\
&= \sum_S \hat{f}(S)^2 \mathbb{E}\big[\chi_S(\omega)\,\chi_S(\omega_\epsilon)\big] \\
&= \sum_S \hat{f}(S)^2 (1-\epsilon)^{|S|}\,. \qquad \text{(IV.2)}
\end{aligned}$$

Moreover, we immediately obtain

$$\mathrm{Cov}(f(\omega), f(\omega_\epsilon)) = \sum_{m=1}^{n} E_f(m)(1-\epsilon)^m, \qquad \text{(IV.3)}$$

Note that either of the last two expressions tell us that $\mathrm{Cov}(f(\omega), f(\omega_\epsilon))$ is nonnegative and decreasing in ϵ. Also, we see that the "level of noise sensitivity" of a Boolean function is naturally encoded in its Energy Spectrum. It is now an an easy exercise to prove the following proposition.

Proposition IV.1. *A sequence of Boolean functions $f_n : \{-1,1\}^{m_n} \to \{0,1\}$ is* **noise sensitive** *if and only if, for any $k \geq 1$*

$$\sum_{m=1}^{k} \sum_{|S|=m} \hat{f}_n(S)^2 = \sum_{m=1}^{k} E_{f_n}(m) \underset{n\to\infty}{\longrightarrow} 0\,.$$

Moreover, (I.1) holding does not depend on the value of $\epsilon \in (0,1)$ chosen.

There is a similar spectral description of noise stability which, given (IV.2), is an easy exercise.

Proposition IV.2. *A sequence of Boolean functions $f_n : \{-1,1\}^{m_n} \to \{0,1\}$ is* **noise stable** *if and only if, for any $\epsilon > 0$, there exists k such that for all n*

$$\sum_{m=k}^{\infty} \sum_{|S|=m} \hat{f}_n(S)^2 = \sum_{m=k}^{\infty} E_{f_n}(m) < \epsilon.$$

So, as argued in the introduction, a function of "high frequency" will be sensitive to noise while a function of "low frequency" will be stable.

4 Link between the spectrum and influence

In this section, we relate the notion of influence with that of the spectrum.

Proposition IV.3. *If $f : \Omega_n \to \{0,1\}$, then for all k*

$$\mathbf{I}_k(f) = 4 \sum_{S:k\in S} \hat{f}(S)^2$$

and

$$\mathbf{I}(f) = 4 \sum_{S} |S| \hat{f}(S)^2.$$

Proof. If $f : \Omega_n \to \mathbb{R}$, we introduce the functions

$$\nabla_k f : \left\{ \begin{array}{ccc} \Omega_n & \to & \mathbb{R} \\ \omega & \mapsto & f(\omega) - f(\sigma_k(\omega)) \end{array} \right. \quad \text{for all } k \in [n]\,,$$

where σ_k acts on Ω_n by flipping the kth bit (thus $\nabla_k f$ corresponds to a discrete derivative along the kth bit).

Observe that

$$\nabla_k f(\omega) = \sum_{S \subseteq \{1,\ldots,n\}} \hat{f}(S)\,[\chi_S(\omega) - \chi_S(\sigma_k(\omega))] = \sum_{S \subseteq \{1,\ldots,n\}, k \in S} \hat{f}(S)\,\chi_S(\omega),$$

from which it follows that for any $S \subseteq [n]$,

$$\widehat{\nabla_k f}(S) = \begin{cases} 2\hat{f}(S) & \text{if } k \in S \\ 0 & \text{otherwise} \end{cases} \tag{IV.4}$$

Clearly if f maps into $\{0,1\}$, then $\mathbf{I}_k(f) := \|\nabla_k f\|_1$ and since $\nabla_k f$ takes values in $\{-1,0,1\}$ in this case, we have $\|\nabla_k f\|_1 = \|\nabla_k f\|_2^2$. Applying Parseval to $\nabla_k f$ and using (IV.4), one obtains the first statement of the proposition. The second is obtained by summing over k and exchanging the order of summation. □

Remark IV.1. If f maps into $\{-1,1\}$ instead, then one can easily check that $\mathbf{I}_k(f) = \sum_{S:k\in S} \hat{f}(S)^2$ and $\mathbf{I}(f) = \sum_S |S|\hat{f}(S)^2$.

5 Monotone functions and their spectrum

It turns out that for monotone functions, there is an alternative useful spectral description of the influences.

Proposition IV.4. *If $f : \Omega_n \to \{0,1\}$ is monotone, then for all k*

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

If f maps into $\{-1,1\}$ instead, then one has that $\mathbf{I}_k(f) = \hat{f}(\{k\})$.

Proof. We prove only the first statement; the second is proved in the same way.

$$\hat{f}(\{k\}) := \mathbb{E}\big[f\chi_{\{k\}}\big] = \mathbb{E}\big[f\chi_{\{k\}}I_{\{k\notin\mathcal{P}\}}\big] + \mathbb{E}\big[f\chi_{\{k\}}I_{\{k\in\mathcal{P}\}}\big]$$

It is easily seen that the first term is 0 (independent of whether f is monotone or not) and the second term is $\frac{\mathbf{I}_k(f)}{2}$ due to monotonicity. □

Remark IV.2. This tells us that for monotone functions mapping into $\{-1,1\}$, the sum in Theorem I.5 is exactly the total weight of the level 1 Fourier coefficients, that is the energy spectrum at 1, $E_f(1)$. (If we map into $\{0,1\}$ instead, there is simply an extra irrelevant factor of 4.) So Theorem I.5 and Proposition IV.1 say that for monotone functions, if the energy spectrum at 1 goes to 0, then this is true for any fixed level. Theorem I.5 and Proposition IV.1 also now easily imply that for monotone functions, the converse of Theorem I.5 holds.

Another application of Proposition IV.4 gives a general upper bound for the total influence for monotone functions.

Proposition IV.5. *If $f : \Omega_n \to \{-1, 1\}$ or $\{0, 1\}$ is monotone, then*

$$\mathbf{I}(f) \leq \sqrt{n}.$$

Proof. If the image is $\{-1, 1\}$, then by PropositionIV.4, we have

$$\mathbf{I}(f) = \sum_{k=1}^{n} \mathbf{I}_k(f) = \sum_{k=1}^{n} \hat{f}(\{k\}).$$

By Cauchy-Schwartz, this is at most $(\sum_{k=1}^{n} \hat{f}^2(\{k\}))^{1/2}\sqrt{n}$. By Parseval's theorem, the first term is at most 1 and we are done. If the image is $\{0, 1\}$, the above proof can easily be modified or one can deduce it from the first case since the total influence of the corresponding ± 1 valued function is the same. $\square$

Remark IV.3. The above result with some universal c on the right hand side follows (for odd n) from the exercise showing that majority has the largest influence together with the known influences for majority. However, the above result yields a more direct proof of the $\sqrt{n}$ bound.

Exercise sheet of chapter IV

Exercise 1. Prove the discrete Poincaré inequality, Theorem I.1 using the spectrum.

Exercise 2. Compute the Fourier coefficients for the indicator function that there are all 1's.

Exercise 3. Show that all even size Fourier coefficients for the majority function are 0. Can you extend this result to a broader class of Boolean functions?

Exercise 4. For the Majority function $\mathbf{MAJ}_n$, find the limit (as the number of voters n goes to infinity) of the following quantity (total weight of the level-3 Fourier coefficients)

$$E_{\mathbf{MAJ}_n}(3) := \sum_{|S|=3} \widehat{\mathbf{MAJ}}_n(S)^2 \,.$$

Exercise 5. Let f_n be a sequence of Boolean functions which is noise sensitive and g_n be a sequence of Boolean functions which is noise stable. Show that f_n and g_n are asymptotically uncorrelated.

Exercise 6 (Another equivalent definition of noise sensitivity)**.** Assume that $\{A_n\}$ is a noise sensitive sequence. (This of course means that the indicator functions of these events is a noise sensitive sequence.)

(a) Show for each $\epsilon > 0$, we have that $\mathbb{P}\big[\omega_\epsilon \in A_n \,\big|\, \omega\big] - \mathbb{P}\big[A_n\big]$ approaches 0 in probability. (Hint: use the Fourier representation.)

(b) Can you show this implication without the Fourier representation?

(c) Discuss if this implication is surprising.

(d) Show that the condition in part (a) implies that the sequence is noise sensitive directly without the Fourier representation.

Exercise 7. How does the spectrum of a generic Boolean function look ? Use this to answer rigorously the question asked in problem 13 of Chapter I.

Exercise 8. *(Open exercise).* For Boolean functions, can one have ANY shape of the Energy spectrum (besides $\sum_{k\geq 1} E_f(k) \leq 1$), or are there restrictions ?

For the next exercises, we introduce the following functional which measures the stability of Boolean functions. For any Boolean function $f : \Omega_n \to \{-1, 1\}$, let

$$\mathbb{S}_f \,:\, \epsilon \mapsto \mathbb{P}\big[f(\omega) \neq f(\omega^\epsilon)\big]\,.$$

Obviously, the smaller $\mathbb{S}_f$ is, the more stable f is.

Exercise 9. Express the functional $\mathbb{S}_f$ in terms of the Fourier expansion of f.

Exercise 10. Among **balanced** Boolean functions, does there exist some function f^* which is "stablest" in the sense that for any balanced Boolean function f and any $\epsilon > 0$,

$$\mathbb{S}_{f^*}(\epsilon) \leq \mathbb{S}_f(\epsilon)\,?$$

If yes, describe the set of these extremal functions (and prove that these are the only ones).

Problem 11. In this problem, we wish to understand the <u>asymptotic shape</u> of the Energy Spectrum for $\mathbf{MAJ}_n$.

(a) Compute for all $\epsilon > 0$, the limit as $n \to \infty$ of $\mathbb{S}_{\mathbf{MAJ}_n}(\epsilon)$.

(b) For any $n \geq 1$, define the following holomorphic function on $\mathbb{D}$ defined by

$$\begin{aligned}\phi_n(z) \;&:=\; \sum_{k\geq 1}\Big[\sum_{|S|=k}\widehat{\mathbf{MAJ}}_n(S)^2\Big]z^k \\ &=\; \sum_k E_{\mathbf{MAJ}_n}(k)z^k\end{aligned}$$

Show that $(\phi_n)_n$ is a normal family of holomorphic functions on $\mathbb{D}$.

(c) Deduce from (a) and (b), an asymptotic formula for $E_{\mathbf{MAJ}_n}(k) = \sum_{|S|=k}\widehat{\mathbf{MAJ}}_n(S)^2$ for all $k \geq 1$. Check that the answer is consistent with $k = 1$ and $k = 3$ (Exercise 4).

Chapter V

Hypercontractivity and its applications

In this lecture, we will prove the main theorems about influences stated in Chapter I (in particular the famous Theorem I.2 from [KKL88] that there always exists a $\Omega(\frac{\log n}{n})$ influential variable). As we will see, these proofs rely on techniques brought from harmonic analysis, in particular *hypercontractivity*. As we will see later in this chapter and in Chapter VII, these type of proofs extend to other contexts which will be of interest to us: noise sensitivity and sub-Gaussian fluctuations.

1 Heuristics of proofs

All the subsequent proofs which will be based on *hypercontractivity* will have more or less the same flavor. Let us now explain in the particular case of Theorem I.2 what is the overall scheme of the proof.

Recall that we want to prove that there exists a universal constant $c > 0$ such that for any function $f : \Omega_n \to \{0, 1\}$, one of its variables has influence at least $c\frac{\log n \operatorname{Var}(f)}{n}$.

Let f be a Boolean function. Suppose all its influences $\mathbf{I}_k(f)$ are 'small' (this would need to be made quantitative). From section 4 of Chapter IV, this means that all the functions $\nabla_k f$ have small L^1 norm. Now, since f is Boolean (into $\{0, 1\}$), then as noticed above, $\nabla_k f$ is almost Boolean (its values are in $\{-1, 0, 1\}$). In particular $\nabla_k f$ must have a small support and its L^2 norm is small. Using the intuition coming from the Weyl-Heisenberg uncertainty, $\widehat{\nabla_k f}$ should then be quite spread, in particular, most of its spectral mass should be concentrated on high frequencies.

This intuition (which is still vague at this point) somehow says that having small influences pushes the spectrum of $\nabla_k f$ towards high frequencies. Now, summing up as we did in section 4 of Chapter IV, but only restricting ourselves to frequencies S of size smaller than some large (well-chosen) $1 \ll M \ll n$, one easily obtains

$$\begin{aligned}
\sum_{0<|S|<M} \hat{f}(S)^2 &\leq 4 \sum_{0<|S|<M} |S| \hat{f}(S)^2 \\
&= \sum_k \sum_{0<|S|<M} \widehat{\nabla_k f}(S)^2 \\
&\text{"}\ll\text{"} \sum_k \|\widehat{\nabla_k f}\|_2^2 \\
&= \mathbf{I}(f)\,, \qquad \text{(V.1)}
\end{aligned}$$

where in the third line, we used the informal statement that $\widehat{\nabla_k f}$ should be supported on high frequencies if f has small influences. Now recall (or observe) that

$$\sum_{|S|>0} \hat{f}(S)^2 = \mathrm{Var}(f)\,.$$

Therefore, in the above equation (V.1), if we are in the case where a positive fraction of the Fourier mass of f is concentrated below M, then (V.1) says that $\mathbf{I}(f)$ is much larger than $\mathrm{Var}(f)$. In particular, at least one of the influences has to be 'large'. If, on the other hand, we are in the case where most of the Spectral mass of f is supported on frequencies of size higher than M, then we also obtain that $\mathbf{I}(f)$ is large by using the formula:

$$\mathbf{I}(f) = 4 \sum_S |S| \hat{f}(S)^2\,.$$

Remark V.1. Note that these heuristics suggest that there is a subtle balance between $\sum_k \mathbf{I}_k(f) = \mathbf{I}(f)$ and $\sup_k \mathbf{I}_k(f)$. Namely, if influences are 'collectively small' (i.e. $\|\cdot\|_\infty$ small), then their sum on the other hand has to be 'large'. The right balance is exactly quantified by Theorem I.3.

Of course it now remains to convert the above sketch into a proof. The main difficulty in the above program is to obtain quantitative spectral information on functions with values in $\{-1, 0, 1\}$ knowing that they have small support. This is done ([KKL88]) using techniques brought from harmonic analysis, namely *Hypercontractivity.*

2 About hypercontractivity

First, let us state what it corresponds to. Let $(K_t)_{t\geq 0}$ be the heat kernel on $\mathbb{R}^n$. Hypercontractivity is a statement which quantifies how functions are regularized under the heat flow. The statement, which goes back to a number of authors can be simply stated as follows:

Theorem V.1 (Hypercontractivity)**.** *If $1 < q < 2$, there is some $t = t(q) > 0$ (which does not depend on the dimension n) such that for any $f \in L^q(\mathbb{R}^n)$,*

$$\|K_t * f\|_2 \leq \|f\|_q\,.$$

The dependence $t = t(q)$ is explicit but will not concern us in the Gaussian case. Hypercontractivity is thus a regularization statement: if one starts with some initial "rough" L^q function f outside of L^2 and waits long enough $(t(q))$ under the heat flow, we end up being in L^2 with a good control on its L^2 norm.

This concept has an interesting history as is nicely explained in O'Donnell lecture notes (see [O'D]). It was originally invented by Nelson in [Nel66] where he needed regularization estimates on Free Fields (which are the building blocks of quantum field theory) in order to apply these in "constructive field theory". It was then generalized by Gross in his elaboration of Logarithmic Sobolev Inequalities ([Gro75]), which are an important tool in analysis. Hypercontractivity is intimately related to these Log-Sobolev Inequalities (they are somewhat equivalent concepts) and thus has many applications in the theory of Semi-Groups, mixing of Markov chains and so on.

We now state the result in the case which concerns us, the hypercube. For any $\rho \in [0,1]$, let T_ρ be the following **Noise Operator** on the set of functions on the hypercube: recall from Chapter I that if $\omega \in \Omega_n$, we denote by ω_ϵ an ϵ-noised configuration of ω. For any $f : \Omega_n \to \mathbb{R}$, we define $T_\rho f : \omega \mapsto \mathbb{E}\big[f(\omega^{1-\rho}) \bigm| \omega\big]$. This Noise Operator acts in a very simple way on the Fourier coefficients:

$$T_\rho : f = \sum_S \hat{f}(S)\, \chi_S \mapsto \sum_S \rho^{|S|} \hat{f}(S)\, \chi_S \,.$$

We have the following analog of Theorem V.1

Theorem V.2 (Bonami-Gross-Beckner)**.** *For any $f : \Omega_n \to \mathbb{R}$ and any $\rho \in [0,1]$,*

$$\|T_\rho f\|_2 \leq \|f\|_{1+\rho^2} \,.$$

The analogy with the classical result V.1 is clear: the Heat flow is replaced here by the random Walk on the hypercube. You can find the proof of this result in the appendix attached to the present chapter.

Remark V.2. The term *hypercontractive* refers here to the fact that one has an operator which maps L^q into L^2 $(q < 2)$, which is a contraction.

Before going into the detailed proof, let us see why such an estimate provides us with the type of spectral information we need. In the above sketch, we assumed that all influences were small. This can be written as

$$\mathbf{I}_k(f) = \|\nabla_k f\|_1 = \|\nabla_k f\|_2^2 \ll 1 \,,$$

for any $k \in [n]$. Now if one applies the hypercontractive estimate to these functions $\nabla_k f$ for some fixed $0 < \rho < 1$, we obtain that

$$\|T_\rho \nabla_k f\|_2 \leq \|\nabla_k f\|_{1+\rho^2} = \|\nabla_k f\|_2^{2/(1+\rho^2)} \ll \|\nabla_k f\|_2 \,,$$

where, for the equality, we used once again that $\nabla_k f \in \{-1, 0, 1\}$. After squaring, this gives on the Fourier side:

$$\sum_S \rho^{2|S|} \widehat{\nabla_k f}(S)^2 \ll \sum_S \widehat{\nabla_k f}(S)^2 \,.$$

This shows (under the assumption that $\mathbf{I}_k(f)$ is small) that the spectrum of $\nabla_k f$ is indeed mostly concentrated on high frequencies.

3 Proof of the KKL theorems on the influences of Boolean functions

We will start by proving Theorem I.2, and then Theorem I.3. In fact, it turns out that one can recover Theorem I.2 directly from Theorem I.3, see the exercises. Nevertheless, since the proof of theorem I.2 is slightly simpler, we start by this one.

3.1 Proof of Theorem I.2

Let $f : \Omega_n \to \{0, 1\}$. Recall that we want to show that there is some $k \in [n]$ such that

$$\mathbf{I}_k(f) \geq c \mathrm{Var}(f) \frac{\log n}{n} \,, \tag{V.2}$$

for some universal constant $c > 0$.

We divide the analysis into the following two cases.

Case 1:

Suppose that there is some $k \in [n]$ such that $\mathbf{I}_k(f) \geq n^{-3/4} \mathrm{Var}(f)$. Then the bound V.2 is clearly satisfied for a small enough $c > 0$.

Case 2:

Now, if f does not belong to the first case, this means that for all $k \in [n]$,

$$\mathbf{I}_k(f) = \|\nabla_k f\|_2^2 \leq \mathrm{Var}(f) n^{-3/4} \,. \tag{V.3}$$

Following the above heuristics, we will show that under this assumption, most of the Fourier spectrum of f is supported on high frequencies. Let $M \geq 1$, whose value will be chosen later. We wish to bound from above the bottom part (up to M) of the Fourier spectrum of f.

$$\begin{aligned}
\sum_{1 \leq |S| \leq M} \hat{f}(S)^2 &\leq \sum_{1 \leq |S| \leq M} |S| \hat{f}(S)^2 \\
&\leq 2^{2M} \sum_{|S| \geq 1} (1/2)^{2|S|} |S| \hat{f}(S)^2 \\
&= \frac{1}{4} 2^{2M} \sum_k \|T_{1/2} \nabla_k f\|_2^2 \,,
\end{aligned}$$

(see section 4 of Chapter IV). Now by applying hypercontractivity (Theorem V.2) with $\rho = 1/2$ to the above sum, we obtain

$$\begin{aligned}
\sum_{1\le|S|\le M} \hat{f}(S)^2 &\le \frac{1}{4}\, 2^{2M} \sum_k \|\nabla_k f\|_{5/4}^2 \\
&\le 2^{2M} \sum_k \mathbf{I}_k(f)^{8/5} \\
&\le 2^{2M}\, n \operatorname{Var}(f)^{8/5} n^{\frac{-3}{4}\cdot\frac{8}{5}} \\
&\le 2^{2M}\, n^{-1/5} \operatorname{Var}(f)\,,
\end{aligned}$$

where we used the assumption V.3 and the obvious fact that $\operatorname{Var}(f)^{8/5} \le \operatorname{Var}(f)$ (recall $\operatorname{Var}(f) \le 1$ since f is Boolean). Now with $M := \lfloor \frac{1}{20}\log_2 n \rfloor$, this gives

$$\sum_{1\le|S|\le\frac{1}{20}\log_2 n} \hat{f}(S)^2 \le n^{1/10-1/5} \operatorname{Var}(f) = n^{-1/10} \operatorname{Var}(f)\,.$$

This shows that under our above assumption, most of the Fourier spectrum is concentrated above $\Omega(\log n)$. We are now ready to conclude:

$$\begin{aligned}
\sup_k \mathbf{I}_k(f) \ge \frac{\sum_k \mathbf{I}_k(f)}{n} &= \frac{4\sum_{|S|\ge 1} |S|\hat{f}(S)^2}{n} \\
&\ge \frac{1}{n}\Big[\sum_{|S|>M} |S|\hat{f}(S)^2\Big] \\
&\ge \frac{M}{n}\Big[\sum_{|S|>M} \hat{f}(S)^2\Big] \\
&= \frac{M}{n}\Big[\operatorname{Var}(f) - \sum_{1\le|S|\le M} \hat{f}(S)^2\Big] \\
&\ge \frac{M}{n}\operatorname{Var}(f)\big[1 - n^{-1/10}\big] \\
&\ge c_1 \operatorname{Var}(f)\frac{\log n}{n}\,,
\end{aligned}$$

with $c_1 = \frac{1}{20\log 2}(1 - 2^{-1/10})$. By combining with the constant given in case 1, this completes the proof. □

Remark V.3. We did not try here to optimize the proof in order to find the best possible universal constant $c > 0$. Note though, that even without optimizing at all, the constant we obtain is not that bad.

3.2 Proof of Theorem I.3

We now proceed to the proof of the stronger result, Theorem I.3, which states that there is a universal constant $c > 0$ such that for any $f : \Omega_n \to \{0,1\}$:

$$\|\mathbf{I}(f)\| = \|\mathbf{Inf}(f)\|_1 \geq c\,\mathrm{Var}(f) \log \frac{1}{\|\mathbf{Inf}(f)\|_\infty}\,.$$

The strategy is very similar. Let $f : \Omega_n \to \{0,1\}$ and let $\delta := \|\mathbf{Inf}(f)\|_\infty = \sup_k \mathbf{I}_k(f)$. Assume for the moment that $\delta \leq 1/1000$. As in the above proof, we start by bounding the bottom part of the spectrum up to some integer M (whose value will be fixed later). Exactly in the same way as above, one has

$$\begin{aligned}\sum_{1\leq|S|\leq M} \hat{f}(S)^2 &\leq 2^{2M} \sum_k \mathbf{I}_k(f)^{8/5} \\ &\leq 2^{2M}\delta^{3/5} \sum_k \mathbf{I}_k(f) = 2^{2M}\delta^{3/5}\,\mathbf{I}(f)\,.\end{aligned}$$

Now,

$$\begin{aligned}\mathrm{Var}(f) = \sum_{|S|\geq 1} \hat{f}(S)^2 &\leq \sum_{1\leq|S|\leq M} \hat{f}(S)^2 + \frac{1}{M}\sum_{|S|>M} |S|\hat{f}(S)^2 \\ &\leq \big[2^{2M}\delta^{3/5} + \frac{1}{M}\big]\,\mathbf{I}(f)\,.\end{aligned}$$

Choose $M := \frac{3}{10}\log_2(\frac{1}{\delta}) - \frac{1}{2}\log_2\log_2\frac{1}{\delta}$. Since $\delta < 1/1000$, it is easy to check that $M \geq \frac{1}{10}\log_2(1/\delta)$ which leads us to

$$\mathrm{Var}(f) \leq \left[\frac{1}{\log_2(1/\delta)} + \frac{10}{\log_2(1/\delta)}\right]\mathbf{I}(f) \tag{V.4}$$

which gives

$$\mathbf{I}(f) = \|\mathbf{Inf}(f)\|_1 \geq \frac{1}{11\log 2}\,\mathrm{Var}(f)\log\frac{1}{\|\mathbf{Inf}(f)\|_\infty}\,.$$

This gives us the result for $\delta \leq 1/1000$.

Next Poincarés inequality, which says that $\mathbf{I}(f) \geq \mathrm{Var}(f)$, tells us that the claim is true for $\delta \geq 1/1000$ if we take c to be $1/\log 1000$. Since this is larger than $\frac{1}{11\log 2}$, we obtain the result with the constant $c = \frac{1}{11\log 2}$. □

4 KKL away from the uniform measure

In Chapter III (on sharp thresholds), we needed an extension of the above KKL theorems to the p-biased measures $\mathbb{P}_p = (p\delta_1 + (1-p)\delta_{-1})^{\otimes n}$. These extensions are respectively theorems III.2 and III.3.

A first natural idea in order to extend the above proofs would be to extend the hypercontractive estimate (Theorem V.2) to these p-biased measures $\mathbb{P}_p$. This extension of Bonami-Gross-Beckner is possible, but it turns out that the control it gives gets worse and worse near the edges (p close to 0 or 1). This is problematic since both in theorems III.2 and III.3, we need bounds which are **uniform** in $p \in [0,1]$.

Hence, one needs a different approach to extend the KKL theorems. A nice approach was provided in [BKK+92], where they prove the following general Theorem.

Theorem V.3 ([BKK+92])**.** *There exists a universal $c > 0$ such that for any measurable function $f : [0,1]^n \to \{0,1\}$, there exists a variable k such that*

$$\mathbf{I}_k(f) \geq c\,\mathrm{Var}(f)\frac{\log n}{n}\,.$$

Here the 'continuous' hypercube is endowed with the uniform (Lebesgue) measure and for any $k \in [n]$, $\mathbf{I}_k(f)$ denotes the probability that f is not almost-surely constant on the fiber given by $(x_i)_{i\neq k}$.

In other words,

$$\mathbf{I}_k(f) = \mathbb{P}\big[\mathrm{Var}\big(f(x_1,\ldots,x_n) \;\big|\; x_i, i \neq k\big) > 0\big]\,.$$

It is clear how to obtain theorem III.2 from the above theorem. If $p \in [0,1]$ and $f : \Omega_n \to \{0,1\}$, consider $\bar{f}_p : [0,1]^n \to \{0,1\}$ defined by

$$\bar{f}_p(x_1,\ldots,x_n) = f((1_{x_i<p} - 1_{x_i\geq p})_{i\in[n]})\,.$$

Friedgut noticed in [Fri04] that one can recover Theorem V.3 from Theorem I.3. The first idea is to use a symmetrization argument in such a way that the problem reduces to the case of monotone functions. Then, the main idea is the approximate the uniform measure on $[0,1]$ by the dyadic random variable

$$X_M : (x_1,\ldots,x_M) \in \{-1,1\}^M \mapsto \sum_{m=1}^{M} \frac{x_m+1}{2} 2^{-m}\,.$$

On can then approximate $f : [0,1]^n \to \{0,1\}$ by the Boolean function $\hat{f}_M$ defined on $\{-1,1\}^{M\times n}$ by

$$\hat{f}_M(x_1^1,\ldots,x_M^1,\ldots,x_1^n,\ldots,x_M^n) := f(X_M^1,\ldots,X_M^n)\,.$$

Still (as mentioned in the above heuristics) this proof requires two 'technical' steps: a monotonization procedure and an 'approximation' one (going from f to $\hat{f}_M$). Since in

our applications to sharp thresholds we used theorems III.2 and III.3 only in the case of monotone functions, for the sake of simplicity we will not present the monotonization procedure in these notes.

Furthermore, it turns out that for our specific needs (the applications in Chapter III), we do not need to deal with the approximation part either. The reason is that for any Boolean function f, the function $p \mapsto \mathbf{I}_k^p(f)$ is continuous. Hence it is enough to obtain uniform bounds on $\mathbf{I}_k^p(f)$ for dyadic values of p (i.e. $p \in \{m2^{-M}\} \cap [0,1]$).

See the exercises for the proof of theorems III.2 and III.3 when f is assumed to be monotone (problem 4).

Remark V.4. We mentioned above that generalizing hypercontractivity would not allow us to obtain uniform bounds (with p taking any value in $[0,1]$) on the influences. It should be noted though that Talagrand obtained ([Tal94]) results similar to Theorems III.2 and III.3 by somehow generalizing hypercontractivity, but along a less naive line. Finally let us point out that both Talagrand ([Tal94]) and Friedgut and Kalai ([FK96]) obtain sharper versions of Theorems III.2 and III.3 where the constant $c = c_p$ in fact improves (i.e. blows up) near the edges.

5 The noise sensitivity theorem

In this section, we prove the milestone Theorem I.5 from [BKS99]. Before recalling what the statement is, let us define the following functional on Boolean functions. For any $f : \Omega_n \to \{0,1\}$, let

$$\mathbf{H}(f) := \sum_k \mathbf{I}_k(f)^2 = \|\mathbf{Inf}(f)\|_2^2 .$$

Recall the Benjamini-Kalai-Schramm theorem.

Theorem V.4 ([BKS99])**.** *Consider a sequence of Boolean functions $f_n : \Omega_{m_n} \to \{0,1\}$. If*

$$\mathbf{H}(f_n) = \sum_{k=1}^{m_n} \mathbf{I}_k(f)^2 \to 0$$

as $n \to \infty$, then $\{f_n\}_n$ is noise sensitive.

We will in fact prove this Theorem under a stronger condition, namely that $\mathbf{H}(f_n) \leq (m_n)^{-\delta}$ for some exponent $\delta > 0$. Without this assumption of "polynomial decay" on $\mathbf{H}(f_n)$, the proof is more technical and relies on estimates obtained by Talagrand. See the remark at the end of this proof. For our application to the *noise sensitivity of percolation* (see Chapter VI), this stronger assumption will be satisfied and hence we stick to this simpler case in these notes.

The assumption of polynomial decay in fact enables us to prove the following more quantitative result.

Proposition V.5. *For any $\delta > 0$, there exists a constant $M = M(\delta) > 0$ such that if $f_n : \Omega_{m_n} \to \{0,1\}$ is any sequence of Boolean functions satisfying*

$$\mathbf{H}(f_n) \leq (m_n)^{-\delta}\,,$$

then

$$\sum_{1 \leq |S| \leq M \log m_n} \widehat{f_n}(S)^2 \to 0\,.$$

Using PropositionIV.1, this proposition obviously implies Theorem I.5 (when $\mathbf{H}(f_n)$ decays fast enough). Furthermore, this gives a quantitative "logarithmic" control on the noise sensitivity of such functions.

Let us now proceed to the proof. The strategy will be very similar to the one used in the KKL theorems (even though the goal is very different). The main difference here is that the regularization term ρ used in the hypercontractive estimate must be chosen in a more delicate way than in the proofs of KKL results (where we simply fixed $\rho = 1/2$).

Let $M > 0$ be a constant whose value will be chosen later.

$$\begin{aligned}
\sum_{1 \leq |S| \leq M \log(m_n)} \widehat{f_n}(S)^2 &\leq 4 \sum_{1 \leq |S| \leq M \log(m_n)} |S| \widehat{f_n}(S)^2 = \sum_k \sum_{1 \leq |S| \leq M \log(m_n)} \widehat{\nabla_k f_n}(S)^2 \\
&\leq \sum_k (\frac{1}{\rho^2})^{M \log(m_n)} \|T_\rho \nabla_k f_n\|_2^2 \\
&\leq \sum_k (\frac{1}{\rho^2})^{M \log(m_n)} \|\nabla_k f_n\|_{1+\rho^2}^2 \text{ by hypercontractivity.}
\end{aligned}$$

Now, since f_n is Boolean, one has $\|\nabla_k f_n\|_{1+\rho^2} = \|\nabla_k f_n\|_2^{2/(1+\rho^2)}$, hence

$$\begin{aligned}
\sum_{0 < |S| < M \log(m_n)} \hat{f}_n(S)^2 &\leq \rho^{-2M \log(m_n)} \sum_k \|\nabla_k f_n\|_2^{4/(1+\rho^2)} = \rho^{-2M \log(m_n)} \sum_k \mathbf{I}_k(f_n)^{2/(1+\rho^2)} \\
&\leq \rho^{-2M \log(m_n)} (m_n)^{\rho^2/(1+\rho^2)} \Big(\sum_k \mathbf{I}_k(f_n)^2 \Big)^{\frac{1}{1+\rho^2}} \text{ (by Hölder)} \\
&= \rho^{-2M \log(m_n)} (m_n)^{\rho^2/(1+\rho^2)}\, \mathbf{H}(f_n)^{\frac{1}{1+\rho^2}} \\
&\leq \rho^{-2M \log(m_n)} (m_n)^{\frac{\rho^2 - \delta}{1+\rho^2}}\,.
\end{aligned}$$

Now by choosing $\rho \in (0,1)$ close enough to 0, and then by choosing $M = M(\delta)$ small enough, we obtain the desired logarithmic noise Sensitivity. □

Some words on the proof of Theorem I.5 in the general case.

Recall that Theorem I.5 is true independently of the speed of convergence of $\mathbf{H}(f_n) = \sum_k \mathbf{I}_k(f_n)^2$. The proof of this general result is a bit more involved than the one we outlined here. The main lemma is as follows:

Lemma V.6. *There exist absolute constants C_k such that for any monotone Boolean function f and for any $k \geq 1$, one has*

$$\sum_{|S|=k} \hat{f}(S)^2 \leq C_k \mathbf{H}(f) \left(-\log \mathbf{H}(f)\right)^{k-1}.$$

This lemma "mimics" a result from Talagrand's [Tal96]. Indeed proposition 2.2 in [Tal96] can be translated as follows: for any monotone Boolean function f, its level-2 Fourier weight (i.e. $\sum_{|S|=2} \hat{f}(S)^2$) is bounded by $O(1)\mathbf{H}(f)\log(1/\mathbf{H}(f))$. It obviously implies Theorem I.5 in the monotone case, the general case being deduced from it by a monotonization procedure. It is worth pointing out that hypercontractivity is used in the proof of this lemma.

Appendix: sketch of proof for hypercontractivity

The purpose of this appendix is to show that we are not using a giant "hammer" but rather that this needed inequality arising from Fourier analysis is understandable from first principles. In fact, historically, the proof by Gross of the Gaussian case first looked at the case of the hypercube and so we have the tools to obtain the Gaussian case should we want to. Before starting the proof, observe that for $\rho = 0$ (where 0^0 is defined to be 1), this simply reduces to $|\int f| \leq \int |f|$.

Proof.

1 Tensorization

In this first section, we show that it is sufficient, via a tensorization procedure, that the result holds for $n = 1$ in order for us to be able to conclude by induction the result for all n.

The key step of the argument is the following lemma.

Lemma V.7. *Let $q \geq p \geq 1$, $(\Omega_1, \mu_1), (\Omega_2, \mu_2)$ be two finite probability spaces, $K_i : \Omega_i \times \Omega_i \to \mathbb{R}$ and assume that for $i = 1, 2$*

$$\|T_i(f)\|_{L_q(\Omega_i,\mu_i)} \leq \|f\|_{L_p(\Omega_i,\mu_i)}$$

where $T_i(f)(x) := \int_{\Omega_i} f(y)K_i(x, y)d\mu_i(y)$. Then

$$\|T_1 \otimes T_2(f)\|_{L_q((\Omega_1,\mu_1)\times(\Omega_2,\mu_2))} \leq \|f\|_{L_p((\Omega_1,\mu_1)\times(\Omega_2,\mu_2))}$$

where $T_1 \otimes T_2(f)(x_1, x_2) := \int_{\Omega_1\times\Omega_2} f(y_1, y_2)K_1(x_1, y_1)K_2(x_2, y_2)d\mu_1(y_1) \times d\mu_2(y_2)$.

Proof. One first needs to recall Minkowski's inequality for integrals which states that for $g \geq 0$ and $r \in [1, \infty)$, we have that

$$\left(\int \left(\int g(x, y)d\nu(y)\right)^r d\mu(x)\right)^{1/r} \leq \int \left(\int g(x, y)^r d\mu(x)\right)^{1/r} d\nu(y).$$

One can think of T_1 acting on functions of both variables by leaving the second variable untouched and analogously for T_2. It is then easy to check that $T_1 \otimes T_2 = T_1 \circ T_2$. By thinking of x_2 as fixed, our assumption on T_1 yields

$$\|T_1 \otimes T_2(f)\|^q_{L_q((\Omega_1,\mu_1)\times(\Omega_2,\mu_2))} \leq \int_{\Omega_2} \left(\int_{\Omega_1} |T_2(f)|^p d\mu_1(x_1) \right)^{q/p} d\mu_2(x_2).$$

(It might be helpful here to think of $T_2(f)(x_1, x_2)$ as a function $g^{x_2}(x_1)$ where x_2 is fixed).

Applying Minkowski's integral inequality to $|T_2(f)|^p$ with $r = q/p$, this in turn is at most

$$\left[\int_{\Omega_1} \left(\int_{\Omega_2} |T_2(f)|^q d\mu_2(x_2) \right)^{p/q} d\mu_1(x_1) \right]^{q/p}.$$

Fixing now the x_1 variable and applying our assumption on T_2 gives that this is at most $\|f\|^q_{L_p((\Omega_1,\mu_1)\times(\Omega_2,\mu_2))}$, as desired. □

The next key observation, easily obtained by expanding and interchanging of summation, is that our operator T_ρ acting on functions on Ω_n corresponds to an operator of the type dealt with in the previous lemma with $K(x, y)$ being

$$\sum_{S \subseteq \{1,\ldots,n\}} \rho^{|S|} \chi_S(x) \chi_S(y).$$

In addition, it is easily checked that the function K for the Ω_n is simply an n-fold product of the function for the $n = 1$ case.

Assuming the result for the case $n = 1$, Lemma V.7 and the above observations allows us to conclude by induction the result for all n.

2 The one-dimensional case ($n = 1$)

We now establish the case $n = 1$. We abbreviate T_ρ by T.

Since $f(x) = (f(-1) + f(1))/2 + (f(1) - f(-1))/2 \ x$, we have $Tf(x) = (f(-1) + f(1))/2 + \rho(f(1) - f(-1))/2 \ x$. Denoting $(f(-1) + f(1))/2$ by a and $(f(1) - f(-1))/2$ by b, it suffices to show that for all a and b, we have

$$(a^2 + \rho^2 b^2)^{(1+\rho^2)/2} \leq \frac{|a + b|^{1+\rho^2} + |a - b|^{1+\rho^2}}{2}.$$

Using $\rho \in [0, 1]$, the case $a = 0$ is immediate. For the case, $a \neq 0$, it is clear we can assume $a > 0$. Dividing both sides by $a^{1+\rho^2}$, we need to show that

$$(1 + \rho^2 y^2)^{(1+\rho^2)/2} \leq \frac{|1 + y|^{1+\rho^2} + |1 - y|^{1+\rho^2}}{2} \tag{V.5}$$

for all y and clearly it suffices to assume $y \geq 0$.

We first do the case that $y \in [0, 1)$. By the generalized Binomial formula, the right hand side of (V.5) is

$$\frac{1}{2}\left[\sum_{k=0}^{\infty}\binom{1+\rho^2}{k}y^k + \sum_{k=0}^{\infty}\binom{1+\rho^2}{k}(-y)^k\right] = \sum_{k=0}^{\infty}\binom{1+\rho^2}{2k}y^{2k}.$$

For the left hand side of (V.5), we first note the following. For $0 < \lambda < 1$, a simple calculation shows that the function $g(x) = (1+x)^\lambda - 1 - \lambda x$ has a negative derivative on $[0, \infty)$ and hence $g(x) \leq 0$ on $[0, \infty)$.

This yields that the left hand side of (V.5) is at most

$$1 + \left(\frac{1+\rho^2}{2}\right)\rho^2 y^2$$

which is precisely the first two terms of the right hand side of (V.5). On the other hand, the binomial coefficients appearing in the other terms are nonnegative since in the numerator, there are an even number of terms with the first two terms being positive and all the other terms being negative. This verifies the desired inequality for $y \in [0, 1)$.

The case $y = 1$ for (V.5) follows by continuity.

For $y > 1$, we let $z = 1/y$ and note, by multiplying both sides of (V.5) by $z^{1+\rho^2}$, we need to show

$$(z^2 + \rho^2)^{(1+\rho^2)/2} \leq \frac{|1+z|^{1+\rho^2} + |1-z|^{1+\rho^2}}{2}. \tag{V.6}$$

Now, expanding $(1 - z^2)(1 - \rho^2)$, one sees that $z^2 + \rho^2 \leq 1 + z^2\rho^2$ and hence the desired inequality follows precisely from (V.5) for the case $y \in (0, 1)$ already proved. This completes the $n = 1$ case and thereby the proof. $\square$

Exercise sheet of Chapter V

Exercise 1. Find a direct proof that Theorem I.3 implies Theorem I.2.

Exercise 2. Is it true that the smaller the influences are, the more noise sensitive the function is ?

Exercise 3. Prove that BKKKL indeed implies Theorem III.3 (use the natural projection).

Problem 4. In this problem, we prove Theorems III.2 and III.3 for the monotone case.

1. Show that theorem III.3 implies III.2 and hence one needs to prove only Theorem III.3 (This is the basically the same as Exercise 1).
2. Show that it suffices to prove the result when $p = k/2^\ell$ for integers k and ℓ.
3. Let $\Pi : \{0,1\}^\ell \to \{0, 1/2^\ell, \ldots, (2^\ell - 1)/2^\ell\}$ by $\Pi(x_1, \ldots, x_\ell) = \sum_{i=1}^{\ell} x_i/2^i$. Observe that if x is uniform, then $\Pi(x)$ is uniform on its range and that $\mathbb{P}(\Pi(x) \geq i/2^\ell) = (2^\ell - i)/2^\ell$.
4. Define $g : \{0,1\}^\ell \to \{0,1\}$ by $g(x_1, \ldots, x_\ell) := I_{\{\Pi(x) \geq 1-p\}}$. Note that $\mathbb{P}(g(x) = 1) = p$.
5. Define $\tilde{f} : \{0,1\}^{n\ell} \to \{0,1\}$ by

   $$\tilde{f}(x_1^1, \ldots, x_\ell^1, x_1^2, \ldots, x_\ell^2, \ldots, x_1^n, \ldots, x_\ell^n) = f(g(x_1^1, \ldots, x_\ell^1), g(x_1^2, \ldots, x_\ell^2), \ldots, g(x_1^n, \ldots, x_\ell^n)).$$

   Observe that $\tilde{f}$ (defined on $(\{0,1\}^{n\ell}, \pi_{1/2})$) and f (defined on $(\{0,1\}^n, \pi_p)$) have the same distribution and hence the same variance.
6. Show (or observe) that $\mathbf{I}_{(r,j)}(\tilde{f}) \leq \mathbf{I}_r^p(f)$ for each $r = 1, \ldots, n$ and $j = 1, \ldots, \ell$. Deduce from Theorem I.3 that

   $$\sum_{r,j} \mathbf{I}_{(r,j)}(\tilde{f}) \geq c\mathrm{Var}(f) \log(1/\delta_p)$$

   where $\delta_p := \max_i \mathbf{I}_i^p(f)$ where c comes from Theorem I.3.

7. (Key step). Show that for each $r = 1, \ldots, n$ and $j = 1, \ldots, \ell$,

$$\mathbf{I}_{(r,j)}(\tilde{f}) \leq \mathbf{I}_r^p(f)/2^{j-1}.$$

8. Combine questions 6 and 7 to complete the proof.

Chapter VI

First evidence of noise sensitivity of percolation

In this lecture, our goal is to collect some of the facts and theorems we have seen so far in order to conclude that percolation crossings are indeed **noise sensitive**. Recall from the "BKS" Theorem (Theorem I.5), that it is enough for this purpose to prove that influences are "small" in the sense that $\sum_k \mathbf{I}_k(f_n)^2$ goes to zero.

In the first section, we will deal with a careful study of influences in the case of percolation crossings on the triangular lattice. Then, we will treat the case of $\mathbb{Z}^2$, where conformal invariance is not known. Finally, we will speculate to what "extent" percolation is noise sensitive.

This whole chapter should be considered somewhat of a "pause" in our program, where we take the time to summarize what we have achieved so far in our understanding of the noise sensitivity of percolation, and what remains to be done if one wishes to prove things such as the existence of exceptional times in dynamical percolation.

1 Bounds on influences for crossing events in critical percolation on the triangular lattice

1.1 Setup

Let us consider some rectangle $[0, a \cdot n] \times [0, b \cdot n]$, and let R_n be the set of of hexagons in $\mathbb{T}$ which intersect $[0, a \cdot n] \times [0, b \cdot n]$. Let f_n be the event that there is a left to right crossing event in R_n. (This is the same event as in Example 7 in chapter I, but with $\mathbb{Z}^2$ replaced by $\mathbb{T}$). By the RSW Theorem, we know that $\{f_n\}$ is non-degenerate. Conformal invariance and SLEs tell us that $\mathbb{E}\big[f_n\big] = \mathbb{P}\big[f_n = 1\big]$ converges as $n \to \infty$. The limit is given by the so called **Cardy's formula**.

In order to prove that this sequence of Boolean functions $\{f_n\}$ is noise sensitive,

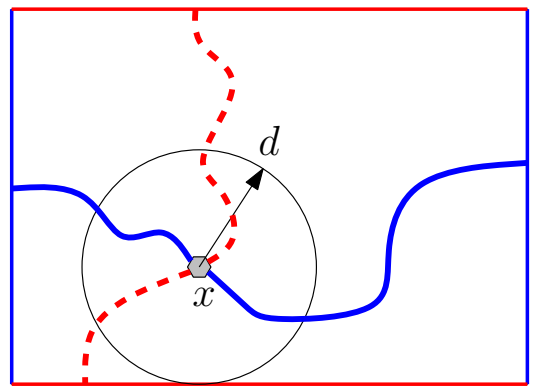
we wish to study its influence vector $\mathbf{Inf}(f_n)$ and we would like to prove that $\mathbf{H}(f_n) = \|\mathbf{Inf}(f_n)\|_2^2 = \sum \mathbf{I}_k(f_n)^2$ decays **polynomially fast** towards 0. (Recall that in these notes, we gave a complete proof of Theorem I.5 only in the case where $\mathbf{H}(f_n)$ decreases as a polynomial of the number of variables.)

1.2 Study of the set of influences

Let x be a site (i.e. an hexagon) in the rectangle R_n. One needs to understand

$$\mathbf{I}_x(f_n) := \mathbb{P}\big[x \text{ is pivotal for } f_n\big]$$

It is easy but crucial to note that if x is at distance d from the boundary of R_n, in order x to be pivotal, the *four-arm* event described in Chapter II (see figure II.2) has to be satisfied in the ball $B(x,d)$ of radius d around the hexagon x. See the figure on the right.



In particular, this implies (still under the assumption that $\mathrm{dist}(x, \partial R_n) = d$) that

$$\mathbf{I}_x(f_n) \le \alpha_4(d) = d^{-\frac{5}{4}+o(1)}\,,$$

where $\alpha_4(d)$ denotes the probability of the four-arm event up to distance d. See Chapter II. There is a $o(1)$ next to the exponent since the probabilities of arm events are known only up to logarithmic error terms. The statement

$$\alpha_4(R) = R^{-5/4+o(1)}$$

implies that for any $\epsilon > 0$, there exists a constant $C = C_\epsilon$, such that for all $R \ge 1$:

$$\alpha_4(R) \le C\, R^{-5/4+\epsilon}\,.$$

The above bound gives us a very good control on the influences of the points in the *bulk* of the domain (i.e. the points far from the boundary). Indeed for any fixed $\delta > 0$, let Δ_n^δ be the set of hexagons which are at distance at least δn from ∂R_n. Most of the points in R_n (except a proportion $O(\delta)$ of these) lie in Δ_n^δ, and for any such point $x \in \Delta_n^\delta$, one has by the above argument

$$\mathbf{I}_x(f_n) \le \alpha_4(\delta n) \le C\,(\delta n)^{-5/4+\epsilon} \le C\delta^{-5/4} n^{-5/4+\epsilon}\,. \tag{VI.1}$$

Therefore, the contribution of these points to $\mathbf{H}(f_n) = \sum_k \mathbf{I}_k(f_n)^2$ is bounded by $O(n^2)(C\delta^{-5/4}n^{-5/4+\epsilon})^2 = O(\delta^{-5/2}n^{-1/2+2\epsilon})$. As $n \to \infty$, this goes to zero **polynomially fast**. Since this estimate concerns "almost" all points in R_n, it seems we are close to proving the BKS criterion.

1.3 Influence of the boundary

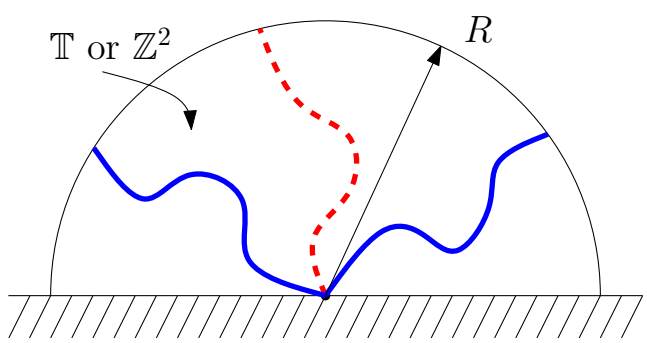
Still, in order to complete the above analysis, one has to estimate what the influence of the points near the boundary is. The main difficulty here is that if x is close to the boundary, the probability for x to be pivotal is not related any more to the above *four-arm* event. Think of the above figure when d gets very small compared to n. One has to distinguish two cases:

- x is close to a **corner**. This will correspond to a *two-arm* event in a quarter plane.
- x is close to an **edge**. This involves the *three-arm* event in the half-plane $\mathbb{H}$.

Before detailing how to estimate the influence of points near the boundary, let us start by giving the necessary background on the involved critical exponents.

The three-arm event in $\mathbb{H}$. For this particular event, it turns out that the critical exponent is known to be *universal*: it is one of the few critical exponents which are known also on the square lattice $\mathbb{Z}^2$. The derivations of these type of exponents do not rely on SLE technology but are "elementary". Therefore in this discussion, we will consider both lattices $\mathbb{T}$ and $\mathbb{Z}^2$.

The *three arm* event in $\mathbb{H}$ corresponds to the event that there are three arms (two open arms and one 'closed' arm in the dual) going from 0 to distance R and such that they remain in the upper half plane. See the figure for a self-explanatory definition.



Let $\alpha_3^+(R)$ denote the probability of this event. As in chapter II, let $\alpha_3^+(r,R)$ be the natural extension to the annulus case (i.e. the probability that this event is satisfied in the annulus between radii r and R).

We will rely on the following result which goes back as far as we know to Aizenman. See [Wer07] for a proof of this result.

Proposition VI.1. *Both on the triangular lattice $\mathbb{T}$ and on $\mathbb{Z}^2$, one has that*

$$\alpha_3^+(R) \asymp R^{-2}\,.$$

Furthermore, one has that

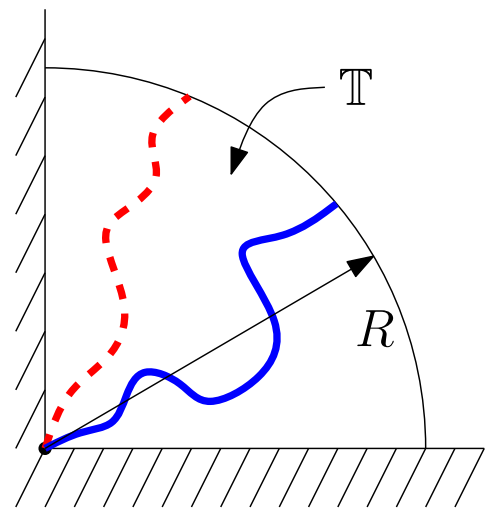
$$\alpha_3^+(r,R) \asymp (r/R)^2\,.$$

Note that in this special case, there are no $o(1)$ correction terms. The exponent is known in this case **up to constants**.

We don't use it here but we mention that the exponent for $\alpha_2^+(r,R)$, the 2 arm half-plane event, is also universal and can be obtained "elementarily" without using SLE technology. One has $\alpha_2^+(r,R) \asymp (r/R)$ and hence the exponent is 1.

The two-arm event in the quarter plane. In this case, the corresponding exponent is unfortunately not known on $\mathbb{Z}^2$, so we will need to do some work here in the next section, where we will prove noise sensitivity of percolation crossings on $\mathbb{Z}^2$.

The *two arm* event in a corner corresponds to the event illustrated on the following picture. We will use the following proposition:



Proposition VI.2 ([SW01]). *If $\alpha_2^{++}(R)$ denotes the probability of this event, then*

$$\alpha_2^{++}(R) = R^{-2+o(1)}\,,$$

and with the obvious notations

$$\alpha_2^{++}(r,R) = (r/R)^{2+o(1)}\,.$$

Now, back to our study of influences, we are in good shape (at least for the triangular lattice) since the two critical exponents arising from the boundary effects are larger than the *bulk* exponent 5/4. This means that it is less likely for a point near the boundary to be pivotal than for a point in the bulk. Therefore in some sense the boundary helps us here.

More formally, summarizing the above facts, for any $\epsilon > 0$, there is a constant $C = C(\epsilon)$ such that for any $1 \le r \le R$

$$\max\{\alpha_4(r,R), \alpha_3^+(r,R), \alpha_2^{++}(r,R)\} \le C(r/R)^{\frac{5}{4}-\epsilon}\,. \tag{VI.2}$$

Now, if x is some site in R_n, let n_0 be the distance to the closest edge of ∂R_n and let x_0 be the point on ∂R_n such that $\mathrm{dist}(x,x_0) = n_0$. Next, let $n_1 \ge n_0$ be the distance from x_0 to the closest corner and let x_1 be this closest corner. It is easy to see that for x to be pivotal for f_n, the following events all have to be satisfied:

- The four-arm event in the ball of radius n_0 around x.
- The $\mathbb{H}$-three-arm event in the annulus centered at x_0 of radii $2n_0$ and n_1.
- The corner-two-arm event in the annulus centered at x_1 of radii $2n_1$ and n.

By independence on disjoint sets, one thus conclude that

$$\begin{aligned} \mathbf{I}_x(f_n) &\le \alpha_4(n_0)\,\alpha_3^+(2n_0,n_1)\,\alpha_2^{++}(2n_1,n) \\ &\le O(1)n^{-5/4+\epsilon}\,. \end{aligned}$$

1.4 Noise sensitivity of crossing events

This uniform bound on the influences over the whole domain R_n enables us to conclude that the BKS criterion is indeed verified. Indeed (for any choice of $\epsilon < 1/4$)

$$\mathbf{H}(f_n) = \sum_{x \in R_n} \mathbf{I}_x(f_n)^2 \leq Cn^2(n^{-5/4+\epsilon})^2 = Cn^{-1/2+2\epsilon}\,, \tag{VI.3}$$

where $C = C(a, b, \epsilon)$ is a universal constant.

This gives us the desired polynomial decay on $\mathbf{H}(f_n)$, which by the restricted proof we gave of Theorem I.5 implies

Theorem VI.3. *The sequence of percolation crossing events $\{f_n\}$ on $\mathbb{T}$ is noise sensitive.*

We will give some other consequences (to sharp thresholds etc..) of the above analysis on the influences of the crossing events in a later section.

2 The case of $\mathbb{Z}^2$ percolation

Let R_n denote similarly the $\mathbb{Z}^2$ rectangle closest to $[0, a \cdot n] \times [0, b \cdot n]$ and let f_n be the left-right crossing event (so here this corresponds exactly to example 7). Here one has to face two main difficulties:

- The main one is that due to the missing ingredient of *conformal invariance*, one does not have at our disposal the value of the *four-arm* critical exponent (which is of course believed to be 5/4). In fact, even the *existence* of a critical exponent is an open problem.
- The second difficulty (also due to the lack of conformal invariance) is that it is now slightly harder to deal with boundary effects. Indeed, one can still use the above bounds on α_3^+ which are *universal*, but the exponent 2 for α_2^{++} is not known for $\mathbb{Z}^2$. So this requires some more analysis.

Let us start by taking care of the boundary effects.

2.1 Handling the boundary effect

All what we need to do if we want the above analysis to carry through for $\mathbb{Z}^2$ is to obtain a reasonable estimate on α_2^{++}. The idea is to try to compare α_2^{++} with α_3^+ (which is known on $\mathbb{Z}^2$) using some geometrical argument. By drawing two quarter-planes next to each other so that they from an half-plane, it is not hard to deduce that

$$\alpha_2^{++}(r, R)^2 \leq \alpha_4^+(r, R) \leq \alpha_3^+(r, R) \leq c\,\Big(\frac{r}{R}\Big)^2\,,$$

by the previous section, for some universal constant $c > 0$. Basically $\alpha_2^{++}(r,R)^2$ corresponds to obtaining independently in each (disjoint) quarter-planes two arms from radius r to R. When this happens, this in particular induces 4 arms in $\mathbb{H}$ from r to R. Some care is needed though to check that the two quarter-plane can be correctly glued together, but this works fine.

This geometric argument thus implies that there exists $C > 0$ such that for any $1 \le r \le R$:

$$\alpha_2^{++}(r,R) \le C\frac{r}{R}\,. \tag{VI.4}$$

Now let e be an edge in R_n. We wish to bound from above $\mathbf{I}_e(f_n)$. We will use the same notations as in the case of the triangular lattice: recall the definitions of n_0, x_0, n_1, x_1 there.

We obtain in the same way

$$\mathbf{I}_e(f_n) \le \alpha_4(n_0)\,\alpha_3^{+}(2n_0,n_1)\,\alpha_2^{++}(2n_1,n)\,. \tag{VI.5}$$

At this point, we need another *universal* exponent which goes back also to Aizenman:

Theorem VI.4 (Aizenman, see [Wer07])**.** *Let $\alpha_5(r,R)$ denote the probability that there are 5 arms (with four of them being of 'alternate colors'). Then there are some universal constants $c, C > 0$ such that both for $\mathbb{T}$ and $\mathbb{Z}^2$, one has for all $1 \le r \le R$*

$$c\big(\frac{r}{R}\big)^2 \le \alpha_5(r,R) \le C\big(\frac{r}{R}\big)^2\,.$$

This result allows us to get a lower bound on $\alpha_4(r,R)$. Indeed, it is clear that

$$\alpha_4(r,R) \ge \alpha_5(r,R) \ge \Omega(1)\alpha_3^{+}(r,R)\,. \tag{VI.6}$$

In fact, using this comparison with the 5-arm event, one can obtain a more precise estimate on $\alpha_4(r,R)$. Since we will need this stronger estimate later, let us now state what it is.

Lemma VI.5. *There exists some $\epsilon > 0$ and some constant $c > 0$ such that for any $1 \le r \le R$*

$$\alpha_4(r,R) \ge c(r/R)^{2-\epsilon}\,.$$

There are several ways to see why this holds (none of them being neither very hard, nor very easy). One of them is to use **Reimer's inequality** which in this case would state that

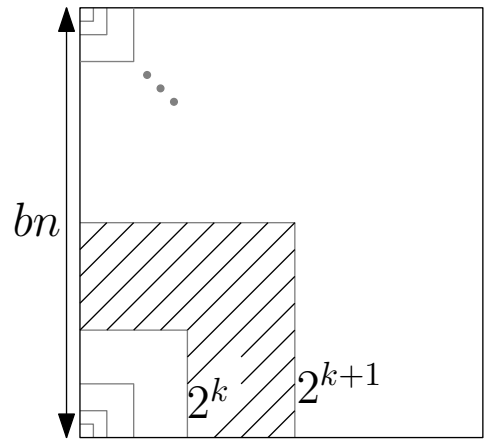
$$\alpha_5(r,R) \le \alpha_1(r,R)\alpha_4(r,R)\,.$$

See [[GPS08], section 2.2 as well as the appendix] for more on these bounds.
Combining (VI.5) with (VI.6), one obtains

$$\begin{aligned}\mathbf{I}_e(f_n) &\le \alpha_4(n_0)\alpha_4(2n_0,n_1)\alpha_2^{++}(2n_1,n)\\ &\le O(1)\alpha_4(n_1)\,\frac{n_1}{n}\,,\end{aligned}$$

where we used quasi-multiplicativity (proposition II.2) as well as the bound given by (VI.4).

Recall we want an upper bound on $\mathbf{H}(f_n) = \sum \mathbf{I}_e(f_n)^2$. In this sum over edges $e \in R_n$, let us divide the set of edges into dyadic annuli centered around the 4 corners as in the next picture.



Notice that there are $O(1)2^{2k}$ edges in an annulus of radius 2^k. This enables us to bound $\mathbf{H}(f_n)$ as follows

$$\begin{aligned} \sum_{e\in R_n} \mathbf{I}_e(f_n)^2 &\leq O(1) \sum_{k=1}^{O(1)\log_2 n} 2^{2k}\Big(\alpha_4(2^k)\frac{2^k}{n}\Big)^2 \\ &\leq O(1)\,\frac{1}{n^2} \sum_{k\leq \log_2 n + O(1)} 2^{4k}\,\alpha_4(2^k)^2\,. \end{aligned} \tag{VI.7}$$

It now remains to obtain a good upper bound on $\alpha_4(R)$, for all $R \geq 1$.

2.2 An upper bound on the four-arm event in $\mathbb{Z}^2$

This turns out to be a rather non-trivial problem. Recall we obtained an easy lower bound on α_4 using α_5. (and lemma VI.5 strengthens this lower bound). For an upper bound, completely different ideas are required. On $\mathbb{Z}^2$, the following estimate is available for the 4-arm event.

Proposition VI.6. *For critical percolation on $\mathbb{Z}^2$, there exists constants $\epsilon, C > 0$ such that for any $1 \leq r \leq R$, one has*

$$\alpha_4(r,R) \leq C\Big(\frac{r}{R}\Big)^{1+\epsilon}.$$

In fact, one can take $\epsilon = \frac{2}{11}$.

Before discussing where such an estimate comes from (we will not give a proof of this proposition in the present chapter, but we will discuss various approaches which lead to this estimate), let us see that it indeed implies a polynomial decay for $\mathbf{H}(f_n)$.

Recall equation (VI.7). Plugging in the above estimate, this gives us

$$\begin{aligned}\sum_{e\in R_n} \mathbf{I}_e(f_n)^2 &\leq O(1)\,\frac{1}{n^2}\sum_{k\leq \log_2 n+O(1)} 2^{4k}\,(2^k)^{-2-2\epsilon}\\ &\leq O(1)\frac{1}{n^2}n^{2-2\epsilon} = O(1)n^{-2\epsilon}\,,\end{aligned}$$

which implies the desired polynomial decay and thus the fact that $\{f_n\}$ is **noise sensitive**.

Let us now discuss different approaches which enable one to prove Proposition VI.6.

(a) Kesten proved implicitly this estimate in the case $r = 1$ in his celebrated paper [Kes87]. His main motivation for such an estimate was to obtain bounds on the corresponding critical exponent which governs the so called *critical length*.

(b) In [BKS99], in order to prove noise sensitivity of percolation using their criterion (on $\mathbf{H}(f_n)$), the author referred to [Kes87], but they also gave a completely different approach which also yields the estimate with $r = 1$ in Proposition VI.6. Furthermore their approach gives an explicit bound on $\epsilon > 0$, namely that (both on $\mathbb{Z}^2$ and $\mathbb{T}$)

$$\alpha_4(R) \leq \left(\frac{1}{R}\right)^{13/12+o(1)}.$$

Their alternative approach is very nice: finding an upper bound for $\alpha_4(R)$ is related to finding an upper bound for the influences of f_R. For this, they noticed the nice following phenomenon: if a monotone function f happens to be very little correlated with majority, then its influences have to be small. The proof of this phenomenon uses for the first time in this context the concept of "randomized algorithms". For more on this approach, see the later Chapter VIII, which is devoted to these type of ideas.

(c) In [SS10b], the concept of randomized algorithms is used in a more powerful way. See again Chapter VIII. One can extract easily from the results in [SS10b] a similar bound on $\alpha_4(R)$.

(d) Note that all these approaches imply Proposition VI.6 only in the case $r = 1$. The extension to the multi-scale case $1 \leq r \leq R$ is non-trivial (quasi-multiplicativity is not enough for this purpose.) Yet, one can prove this by extending an approach of O'Donnell and Servedio (see Theorem VIII.5 later) which yields upper bounds on influences. See also a related sketch in [SS10a]. We will not develop this more here but will rely on this estimate in a later chapter.

3 Some other consequences of our study of influences

In the previous sections, we handled the boundary effects in order to check that $\mathbf{H}(f_n)$ indeed decays polynomially fast. Let us list some related results implied by this analysis.

3.1 Energy spectrum of f_n

We start by a straightforward observation: since f_n are monotone, we have seen in Chapter IV that

$$\widehat{f_n}(\{x\}) = \frac{1}{2}\mathbf{I}_x(f_n)\,,$$

for any site x (or edge e) in R_n. Therefore, the bounds we obtained on $\mathbf{H}(f_n)$ imply the following control on the first layer of the Energy Spectrum of the crossing events $\{f_n\}$.

Corollary VI.7. *Let $\{f_n\}$ be the crossing events of the rectangles R_n.*

- *If we are on the triangular lattice $\mathbb{T}$, then we have the bound*

$$E_{f_n}(1) = \sum_{|S|=1} \widehat{f_n}(S)^2 \leq n^{-1/2+o(1)}\,.$$

- *On the square lattice $\mathbb{Z}^2$, we end up with the weaker estimate*

$$E_{f_n}(1) \leq C\, n^{-\epsilon}\,,$$

  for some $\epsilon, C > 0$.

3.2 Sharp threshold of percolation

The above analysis gave an upper bound on $\sum_k \mathbf{I}_k(f_n)^2$. As we have seen in the first chapters, the total influence $\mathbf{I}(f_n) = \sum_k \mathbf{I}_k(f_n)$ is also a very interesting quantity. Recall that by Russo's formula, this is the quantity which shows "how sharp" the threshold is for $p \mapsto \mathbb{P}_p[f_n = 1]$.

The above analysis allows us to get following upper bound

Proposition VI.8. *There exists a constant $c > 0$ such that both on $\mathbb{T}$ and $\mathbb{Z}^2$, one has*

$$\mathbf{I}(f_n) \leq c\, n^2 \alpha_4(n)\,.$$

In particular, this shows on $\mathbb{T}$ that

$$\mathbf{I}(f_n) \leq n^{3/4+o(1)}\,.$$

Remark VI.1. Since f_n is defined on $\{-1,1\}^{O(n^2)}$, note that the Majority function defined on the same hypercube has a **much sharper** threshold than the percolation crossings f_n.

Sketch of proof:
In the same vein (i.e. using dyadic annuli, quasi-multiplicativity) and with the same notations one has

$$\begin{aligned}\mathbf{I}(f_n) = \sum_e \mathbf{I}_e(f_n) &\leq \sum_e O(1)\alpha_4(n_1)\frac{n_1}{n} \\ &\leq \frac{1}{n}\sum_{k\leq \log_2 n + O(1)} 2^{3k}\alpha_4(2^k)\,.\end{aligned}$$

Now, and this is the main step here, using quasi-multiplicativity one has $\alpha_4(2^k) \leq O(1)\frac{\alpha_4(n)}{\alpha_4(2^k,n)}$, which gives us

$$\begin{aligned}\mathbf{I}(f_n) &\leq O(1)\frac{\alpha_4(n)}{n}\sum_{k\leq \log_2 n + O(1)} 2^{3k}\frac{1}{\alpha_4(2^k,n)} \\ &\leq O(1)\frac{\alpha_4(n)}{n}\sum_{k\leq \log_2 n + O(1)} 2^{3k}\frac{n^2}{2^{2k}} \quad \text{since } \alpha_4(l) \geq \alpha_5(l) \asymp l^{-2} \\ &\leq O(1)n\,\alpha_4(n)\sum_{k\leq \log_2 n + O(1)} 2^k \\ &\leq O(1)n^2\alpha_4(n) \quad \text{as desired.}\end{aligned}$$

□

3.3 A lower bound on the total influence

In fact, one can show that the lower bound on $\mathbf{I}(f_n)$ matches, namely one can prove

Proposition VI.9. *There are universal constants $c, C > 0$ such that both on $\mathbb{Z}^2$ and $\mathbb{T}$:*

$$c\,n^2\alpha_4(n) \leq \mathbf{I}(f_n) = \sum_k \mathbf{I}_k(f_n) \leq Cn^2\alpha_4(n)\,.$$

One obtains a lower bound by just summing over the influences of points whose distance to the boundary is at least $n/4$. It would suffice if we knew that for such edges or hexagons, the influence is at least a constant times $\alpha_4(n)$. This is in fact known to be true. It is not very involved (part of the *folklore* results in percolation), but still would bring us too far from our topic. The needed technique is known under the name of **separation of arms** and is very related to the statement of quasi-multiplicativity. See [Wer07] for more details.

4 Quantitative noise sensitivity

Combining the BKS Theorem I.5 with the present Chapter, we proved that the sequence of crossing events $\{f_n\}$ is noise sensitive. This can be roughly translated as follows: for any fixed level of noise $\epsilon > 0$, as $n \to \infty$, the large scale clusters of ω in the window $[0, n]^2$ are asymptotically independent of the large clusters of ω_ϵ.

Remark VI.2. Note that this intuition is correct, but in order to make it rigorous, this would require some work, since so far we only worked with left-right crossing events. The non-trivial step here is to prove that in some sense, in the scaling limit $n \to \infty$, any macroscopic property concerning percolation is measurable with respect to the σ-algebra generated by the crossing events. This is a rather subtle problem since we need to precise what kind of information we keep in what we call the "scaling limit" of percolation (or subsequential scaling limits in the case of $\mathbb{Z}^2$). For example, having more open sites than closed one is a non-trivial information on the discrete level, but this information is not present at the scaling limit, since by noise sensitivity we know that it is asymptotically decorrelated with crossing events. We will not need to discuss these notions of scaling limits more in these lecture notes, since the focus is mainly on the discrete model itself (including the model of dynamical percolation at the end).

At this stage, a natural question is to wonder to what extend the percolation picture is sensitive to noise. In other words, can we let the noise $\epsilon = \epsilon_n$ go to zero with the "size of the system" n, and yet keep this independence of large scale structures between ω and ω_{ϵ_n} ? If yes, can we give **quantitative estimates** on how fast may the noise $\epsilon = \epsilon_n$ go to zero ? One can state this question more precisely as follows

Question VI.1. *If $\{f_n\}$ denotes our left-right crossing events. For which sequences of noise-levels $\{\epsilon_n\}$ do we have*

$$\lim_{n\to\infty} \mathrm{Cov}[f_n(\omega), f_n(\omega_{\epsilon_n})] = 0 \ ?$$

The purpose of this section is to briefly discuss this question based on the results we have obtained so far.

4.1 Link with the Energy Spectrum of $\{f_n\}$

Check, using section 3 from Chapter IV, that the above question is equivalent to the following one

Question VI.2. *For which sequences $\{k_n\}$ going to infinity do we have*

$$\sum_{m=1}^{k_n} E_{f_n}(m) = \sum_{1\le|S|\le k_n} \widehat{f_n}(S)^2 \underset{n\to\infty}{\longrightarrow} 0 \ ?$$

Recall that we already obtained some relevant information on this question. Indeed we proved in this Chapter that $\mathbf{H}(f_n) = \sum_x \mathbf{I}_x(f_n)^2$ decays polynomially fast towards 0 (both on $\mathbb{Z}^2$ and $\mathbb{T}$). Therefore, using Proposition V.5, this tells us that for some constant $c > 0$, one has

$$\sum_{1 \leq |S| \leq c \log n} \widehat{f_n}(S)^2 \to 0\,. \tag{VI.8}$$

Therefore, back to our original question VI.1, this gives us the following quantitative statement: if for each n, the noise ϵ_n satisfies $\epsilon_n \gg \frac{1}{\log n}$, then $f_n(\omega)$ and $f_n(\omega_{\epsilon_n})$ are asymptotically independent.

4.2 Noise stability regime

Of course, one cannot be too demanding on the rate of decay of $\{\epsilon_n\}$. For example if $\epsilon_n \ll \frac{1}{n^2}$, then in the window $[0,n]^2$, with high probability, the configurations ω and ω_{ϵ_n} are identical. This brings us to the next natural question on the **noise stability regime** of crossing events.

Question VI.3. *Let $\{f_n\}$ be our sequence of crossing events. For which sequences $\{\epsilon_n\}$, do we have*

$$\mathbb{P}\big[f_n(\omega) \neq f_n(\omega_{\epsilon_n})\big] \underset{n\to\infty}{\longrightarrow} 0\ ?$$

Or equivalently, for which sequences $\{k_n\}$, do we have

$$\sum_{|S|>k_n} \widehat{f_n}(S)^2 \to 0\ ?$$

Using the estimates of the present Chapter, one can give the following non-trivial bound on the "noise stability regime" of $\{f_n\}$.

Proposition VI.10. *Both on $\mathbb{Z}^2$ and $\mathbb{T}$, if*

$$\epsilon_n = o\Big(\frac{1}{n^2\alpha_4(n)}\Big)\,,$$

then the crossing events (f_n) from ω to ω_{ϵ_n} remain unchanged with high probability.

On the triangular grid, using the critical exponent, this gives us a bound of $n^{-3/4+o(1)}$ on the **" noise stability regime of percolation"**.

Proof. Let us sketch how this works. Let $\{\epsilon_n\}$ be a sequence satisfying the above assumption. Recall f_n is the crossing event of the rectangle R_n. There are $O(n^2)$ bits concerned. In order to work with simpler notations, assume WLOG that there are exactly n^2 bits. Let us order these in some arbitrary way: $\{x_1, \ldots, x_{n^2}\}$ (or on $\mathbb{Z}^2$, $\{e_1, \ldots, e_{n^2}\}$).

Let $\omega = \omega_0 = (x_1, \ldots, x_{n^2})$ be sampled according to the uniform measure. Recall that the noised configuration ω_{ϵ_n} is produced as follows: for each $i \in [n^2]$, resample the bit x_i with probability ϵ_n, independently of everything else, obtaining the bit y_i. (In particular $y_i \neq x_i$ with probability $\epsilon_n/2$).

Now for each $i \in [n^2]$ define the intermediate configuration

$$\omega_i := (y_1, \ldots, y_i, x_{i+1}, \ldots, x_{n^2})$$

Notice that for each $i \in [n^2]$, ω_i is also sampled according to the uniform measure. Therefore, one has for each $i \in \{1, \ldots, n^2\}$ that

$$\mathbb{P}\big[f_n(\omega_{i-1}) \neq f_n(\omega_i)\big] \leq \epsilon_n/2\, \mathbf{I}_{x_i}(f_n)\,.$$

summing over all i, one obtains

$$\begin{aligned}
\mathbb{P}\big[f_n(\omega) \neq f_n(\omega_{\epsilon_n})\big] &= \mathbb{P}\big[f_n(\omega_0) \neq f_n(\omega_{n^2})\big] \\
&\leq \epsilon_n \sum_{i=1}^{n^2} \mathbf{I}_{x_i}(f_n) \\
&= \epsilon_n\, \mathbf{I}(f_n) \\
&\leq \epsilon_n O(1) n^2 \alpha_4(n) \ \text{ by Proposition VI.8,}
\end{aligned}$$

which concludes the proof.

□

4.3 Where does the spectral mass lies ?

The above proposition shows that, necessarily the Fourier coefficients of f_n satisfy

$$\sum_{|S| \gg n^2 \alpha_4(n)} \widehat{f_n}(S)^2 \underset{n\to\infty}{\longrightarrow} 0\,. \tag{VI.9}$$

From lemma VI.5, we know that even on $\mathbb{Z}^2$, $n^2\alpha_4(n)$ is greater than some polynomial n^ϵ for some exponent $\epsilon > 0$. Combining the estimates on the spectrum we achieved so far (equations (VI.8) and (VI.9)), we see that in order to **localize the Spectral mass** of $\{f_n\}$, there is still a missing gap. See figure 4.3.

For our later applications to the model of dynamical percolation (last Chapter of these lecture notes), a better understanding on the noise sensitivity of percolation than the "logarithmic" control we achieved so far will be needed.

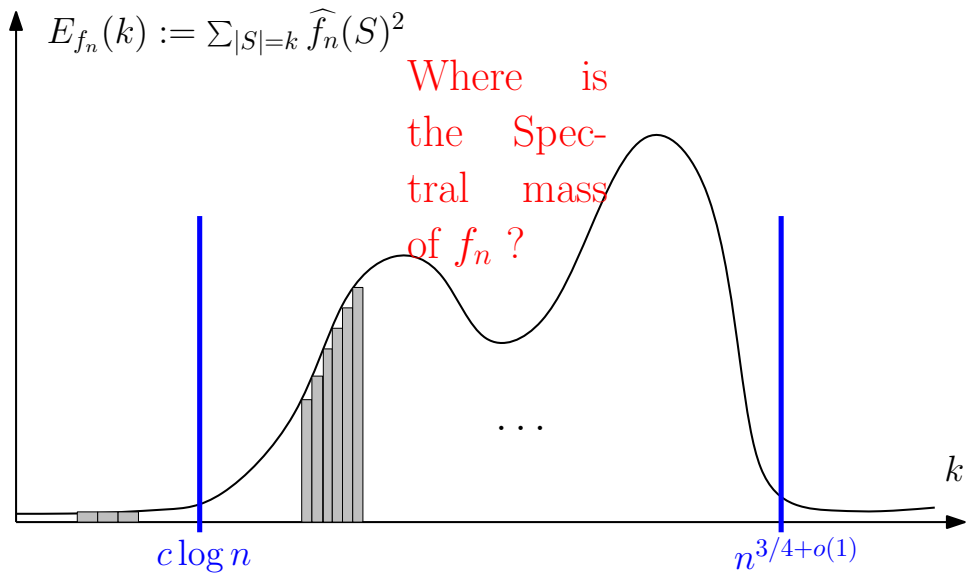


Figure VI.1: The picture on the right-hand side summarizes our present knowledge on the Energy Spectrum of $\{f_n\}$ on the triangular lattice $\mathbb{T}$. Much remains to be understood to know where, in the range $[\Omega(\log n), n^{3/4+o(1)}]$, does the Spectral mass lie. This question will be analyzed in the next Chapters.

Exercise sheet on chapter VI

Instead of being the usual exercise sheet, this page will be devoted to a single Problem whose goal will be to do "hands-on" computations of the first layers of the Energy Spectrum of the percolation crossing events f_n. Recall from Proposition IV.1 that a sequence of Boolean functions $\{f_n\}$ is **noise sensitive** if and only if for any fixed $k \geq 1$

$$\sum_{m=1}^{k} \sum_{|S|=m} \hat{f}_n(S)^2 = \sum_{m=1}^{k} E_{f_n}(m) \underset{n\to\infty}{\longrightarrow} 0\,.$$

In the present chapter, we obtained that it is indeed the case for $k = 1$. The purpose here is to check by simple combinatorial arguments (without relying on hypercontractivity) that it is still the case for $k = 2$ and to convince ourselves that it works for all layers $k \geq 3$.

To start with, we will simplify our task by working on the torus $\mathbb{Z}^2/n\mathbb{Z}^2$. This has the very nice advantage that there are no boundary issues here.

Energy Spectrum of crossing events in the torus (study of the first layers)

Let T_n be either the square grid torus $\mathbb{Z}^2/n\mathbb{Z}^2$ or the triangular grid torus $\mathbb{T}/n\mathbb{T}$. Let f_n be the indicator of the event that there is an open circuit along the first coordinate of T_n.

1. Using RSW, prove that there is a constant $c > 0$ such that for all $n \geq 1$,

$$c \leq \mathbb{P}\big[f_n = 1\big] \leq 1 - c\,.$$

   (In other words, $\{f_n\}$ is non-degenerate.)

2. Show that for all edge e (or site x) in T_n

$$\mathbf{I}_e(f) \leq \alpha_4(\frac{n}{2})\,.$$

3. Check that the BKS criterion (about $\mathbf{H}(f_n)$) is satisfied. Therefore $\{f_n\}$ is noise-sensitive

From now on, one would like to forget about the BKS theorem and try to do some hands-on computations in order to get a feeling why most frequencies should be large.

4. Show that if x, y are two sites of T_n (or similarly if e, e' are two edges of T_n), then
$$|\hat{f}(\{x,y\})| \leq 2\mathbb{P}\big[\ x \text{ and } y \text{ are pivotal points}\big]\,,$$

5. Show that if $d := |x-y|$, then
$$\mathbb{P}\big[\ x \text{ and } y \text{ are pivotal points }\big] \leq O(1)\frac{\alpha_4(n/2)^2}{\alpha_4(\frac{d}{2},\frac{n}{2})}\,.$$
(Hint: use quasi-multiplicativity II.2.)

6. On the square lattice $\mathbb{Z}^2$, by carefully summing over all edges $e, e' \in T_n \times T_n$, show that
$$E_{f_n}(2) = \sum_{|S|=2} \widehat{f_n}(S)^2 \leq\ O(1)n^{-\epsilon}\,,$$
for some exponent $\epsilon > 0$.

   Hint: you might decompose the sum in a dyadic way (as we did many times in the present section) depending on the mutual distance $d(e,e')$.

7. On the triangular grid, what exponent does it give for the decay of $\mathbb{E}_{f_n}(2)$? Compare with the decay we found in Corollary VI.7 about the decay of the first layer $E_{f_n}(1)$ (i.e. $k=1$). Discuss this.

8. What do you expect for higher (fixed) values of k ? (I.e. for $E_{f_n}(k)$, $k \geq 3$) ?

9. *(Quite hard)* Prove rigorously, using a combinatorial argument similar as the one above in the particular case $k=2$, that for any fixed layer $k \geq 1$,
$$E_{f_n}(k) \underset{n\to\infty}{\longrightarrow} 0\,,$$
which gives us an **alternative proof of noise sensitivity of percolation** (at least in the case of the Torus T_n).

Back to our crossing events in rectangles

In this short problem, we will only consider the case of the triangular grid $\mathbb{T}$, since it is much easier to deal with boundary issues in this case, thanks to the estimate (VI.2).

1. Prove rigorously that $E_{f_n}(k=2)$ goes to zero and decays with the same exponent as in the case of Torus.

2. Discuss the general case $k \geq 2$ (non-rigorously).

Chapter VII

Anomalous fluctuations

In this lecture, our goal is to extend the technology we used to prove the KKL theorems on influences and the BKS theorem on noise sensitivity to a slightly different context: the study of fluctuations in **first passage percolation**.

1 The model of first-passage percolation

Let us first explain what the model is. Let $0 < a < b$ be two positive *numbers*. We define a **random metric** on the graph $\mathbb{Z}^d$, $d \geq 2$ as follows: independently for each edge $e \in \mathbb{E}^d$, fix its length τ_e to be a with probability $1/2$ and b with probability $1/2$. This is represented by a uniform configuration $\omega \in \{-1, 1\}^{\mathbb{E}^d}$.

This procedure induces a well-defined (random) metric dist_ω on $\mathbb{Z}^d$ in the usual fashion. For any vertices $x, y \in \mathbb{Z}^d$, let

$$\mathrm{dist}_\omega(x, y) := \inf_{\substack{\text{paths } \gamma = \{e_1, \ldots, e_k\} \\ \text{connecting } x \to y}} \left\{ \sum \tau_{e_i}(\omega) \right\}.$$

Remark VII.1. In greater generality, the lengths of the edges are i.i.d. non-negative random variables, but here, following [BKS03], we will restrict ourselves to the above uniform distribution on $\{a, b\}$ to simplify the exposition (see [BR08] for an extension to more general laws).

One of the main goals in first-passage-percolation is to understand the large-scale properties of this random metric space. For example, for any $T \geq 1$, one may consider the (random) ball

$$B_\omega(x, T) := \{y \in \mathbb{Z}^d, \mathrm{dist}_\omega(x, y) \leq T\}.$$

To understand the name *first-passage-percolation*, one can think of this model as follows. Imagine that a lot of water is pumped in at vertex x, and that for each edge e, it takes $\tau_e(\omega)$ units of time for the water to travel across the edge e. Then, $B_\omega(x, T)$ represents the region of the space that has been wetted by time T.

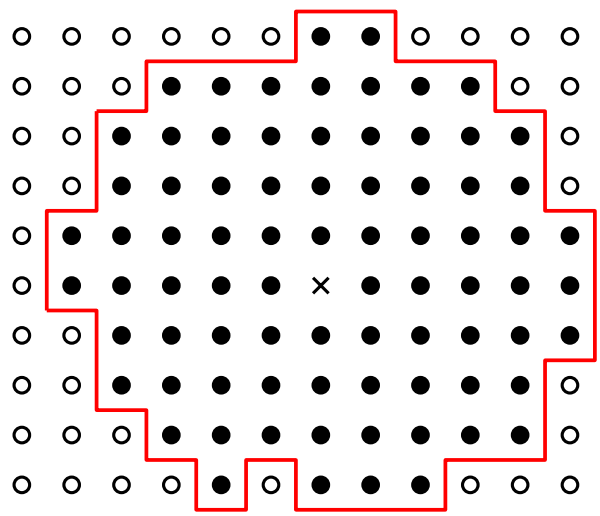


Figure VII.1: A sample of a wetted region at time T, i.e. $B_\omega(x,T)$, in first passage percolation.

An easy application of **sub-additivity** shows that the renormalized ball $\frac{1}{T}B_\omega(0,T)$ converges as $T \to \infty$ towards a deterministic shape (which can be in certain cases computed explicitly). This is a kind of "geometric law of large numbers". Whence the natural question:

Question VII.1. *Describe the* **fluctuations** *of $B_\omega(0,T)$ around its asymptotic deterministic shape.*

This question has received tremendous interest in the last 15 years or so. It is widely believed that these fluctuations should be in some sense **"universal"**. More precisely, the behavior of $B_\omega(0,T)$ around its limiting shape should not depend on the "microscopic" particularities of the model (for example the law on the edges) but only on the dimension d of the underlying graph. The shape itself depends on the other hand of course on the microscopic parameters, in the same way as the critical point depends on the graph in percolation.

In the two-dimensional case, using very beautiful combinatorial bijections with Random Matrices, certain cases of **directed** Last Passage Percolation (where the law on the edges is taken to be geometric or exponential) have been understood very deeply. For example, it is known ([Joh00]) that the fluctuations of the Ball of radius n (i.e. the points whose Last Passage Times are below n) around n times its asymptotic deterministic shape, are of order $n^{1/3}$, and the law of these fluctuations properly renormalized follow the Tracy-Widom distribution (as do the fluctuations of the largest eigenvalue of GUE ensembles).

2 State of the art

Returning to our initial model of (non-directed) First Passage Percolation, it is thus conjectured that, for dimension $d = 2$, fluctuations are of order $n^{1/3}$ following a Tracy-Widom Law. Still, the current state of understanding of this model is far from this conjecture.

Kesten first proved that the fluctuations of the ball of radius n are at most $\sqrt{n}$ (this did not exclude yet a possible Gaussian behavior). Benjamini, Kalai and Schramm then strengthened this result by showing that the fluctuations are sub-Gaussian. This is still far from the conjectured $n^{1/3}$-fluctuations, but their approach has the great advantage to be very general; in particular their result holds in any dimension $d \geq 2$.

Let us now state their main theorem concerning the fluctuations of the metric dist.

Theorem VII.1 ([BKS03])**.** *For all a, b, d, there exists an absolute constant $C = C(a, b, d)$ such that in $\mathbb{Z}^d$*

$$\mathrm{Var}(\mathrm{dist}_\omega(0, v)) \leq C \frac{|v|}{\log |v|}$$

for any $v \in \mathbb{Z}^d, |v| \geq 2$.

To keep things simple in these notes, we will only prove the analogous statement in the Torus (where one has more symmetries and invariance to play with).

3 The case of the Torus

Let $\mathbb{T}_m^d$ be the d-dimensional Torus $(\mathbb{Z}/m\mathbb{Z})^d$. As in the above (lattice) model, independently for each edge of $\mathbb{T}_m^d$, we choose its length to be either a or b equally likely. We are interested here in the smallest (random) length among all closed paths γ "winding" around the torus along the first coordinate $\mathbb{Z}/m\mathbb{Z}$ (i.e. these paths γ, when projected onto the first coordinate, have winding number one). In [BKS03], this is called the shortest *circumference*. For any configuration $\omega \in \{a, b\}^{E(\mathbb{T}_m^d)}$, this shortest circumference is denoted by $\mathrm{Circ}_m(\omega)$.

Theorem VII.2. *There is a constant $C = C(a, b)$ (which does not depend on the dimension d), such that*

$$\mathrm{var}(\mathrm{Circ}_m(\omega)) \leq C \frac{m}{\log m}$$

Remark VII.2. A similar analysis as the one carried out below works in greater generality: if $G = (V, E)$ is some finite connected graph endowed with a random metric d_ω with $\omega \in \{a, b\}^{\otimes E}$, then one can obtain bounds on the fluctuations of the random diameter $D = D_\omega$ of (G, d_ω). See [BKS03, Theorem 2] for a precise statement in this more general context.

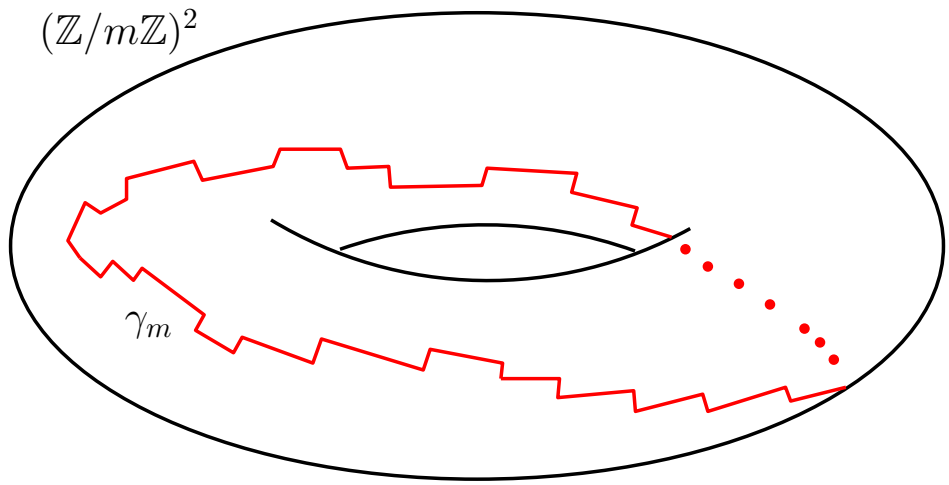


Figure VII.2: The shortest geodesic along the first coordinate for the random metric dist_ω on $(\mathbb{Z}/m\mathbb{Z})^2$.

Sketch of proof of Theorem VII.2.

For any edge e, let us consider the gradient along the edge e: $\nabla_e \mathrm{Circ}_m$. These gradient functions have values in $\{-(b-a), 0, b-a\}$, since changing the length of e can only have this effect on the circumference. By dividing our distances by the constant factor $b-a$, we can even assume our gradient functions to have values in $\{-1, 0, 1\}$. Doing so, we end up being in a setup very similar to the one we had in Chapter V. The **influence** of an edge e corresponds here to $\mathbf{I}_e(\mathrm{Circ}_m) := \|\nabla_e \mathrm{Circ}_m\|_1 = \|\nabla_e \mathrm{Circ}_m\|_2^2$. We will prove later on that Circ_m has very small influences. In other words, we will show that the above gradient functions (which are 'almost' Boolean) have small support, and hypercontractivity will imply the desired bound.

We thus reduced the problem to the following general framework. Consider a real-valued function $f : \{-1,1\}^n \to \mathbb{R}$, such that for any variable k, $\nabla_k f \in \{-1, 0, 1\}$. We are interested in $\mathrm{Var}(f)$ and we want to show that if "influences are small" then $\mathrm{Var}(f)$ is small. It is easy to check that the variance can be written

$$\mathrm{Var}(f) = \frac{1}{4} \sum_k \sum_{\emptyset \neq S \subset [n]} \frac{1}{|S|} \widehat{\nabla_k f}(S)^2 \,.$$

We see on this expression, that if variables have very small influence, then as previously, the almost Boolean $\nabla_k f$ will be of High frequency. Heuristically, this should then imply that

$$\begin{aligned} \mathrm{Var}(f) &\ll \sum_k \sum_{S \neq \emptyset} \widehat{\nabla_k f}(S)^2 \\ &= \sum_k \mathbf{I}_k(f) \,. \end{aligned}$$

This intuition is quantified by the following lemma on the link between the fluctuations of a real-valued function f on Ω_n and its influence vector.

Lemma VII.3. *Let $f : \Omega_n \to \mathbb{R}$ be a (real-valued) function such that each of its discrete derivatives $\nabla_k f$, $k \in [n]$ have their values in $\{-1, 0, 1\}$. Assume that the influences of f are small in the sense that there exists some $\alpha > 0$ such that for any $k \in \{1, \dots, n\}$, $\mathbf{I}_k(f) \leq n^{-\alpha}$. Then there is some constant $C = C(\alpha)$, such that*

$$\mathrm{Var}(f) \leq \frac{C}{\log n} \sum_k \mathbf{I}_k(f) \,.$$

Remark VII.3. If f is Boolean, then this follows from Theorem I.3 with $C(\alpha) = c/\alpha$ with c universal.

The proof of this lemma is postponed to the next section. In the meantime, let us show that in our special case of First Passage Percolation on the torus, the assumption on small influences is indeed verified. Since the edges' length are in $\{a, b\}$, the smallest contour $\mathrm{Circ}_m(\omega)$ in $\mathbb{T}_m^d$ around the first coordinate lies somewhere in $[am, bm]$. Hence, if γ is a geodesic (a path in the Torus) satisfying $\mathrm{length}(\gamma) = \mathrm{Circ}_m(\omega)$, then γ uses at most $\frac{b}{a}m$ edges. There might be several different geodesics minimizing the circumference. Let us choose randomly one of these in an 'invariant' way and call it $\tilde{\gamma}$. For any edge $e \in E(\mathbb{T}_m^d)$, if by changing the length of e, the circumference increases, then e has to be contained in any geodesic γ, and in particular in $\tilde{\gamma}$. This implies that $\mathbb{P}\big[\nabla_e \mathrm{Circ}_m(\omega) > 0\big] \leq \mathbb{P}\big[e \in \tilde{\gamma}\big]$. By symmetry we obtain that

$$\mathbb{P}\big[\nabla_e \mathrm{Circ}_m(\omega) \neq 0\big] \leq 2\mathbb{P}\big[e \in \tilde{\gamma}\big] \,.$$

As we have seen above, up to a constant factor $b-a$, $\nabla_e \mathrm{Circ}_m \in \{-1, 0, 1\}$; therefore $\mathbf{I}_e(\mathrm{Circ}_m) \leq O(1)\mathbb{P}\big[e \in \tilde{\gamma}\big]$. Now using the symmetries both of the Torus $\mathbb{T}_m^d$ and of our observable Circ_m, if $\tilde{\gamma}$ is chosen in an appropriate invariant way (uniformly among all geodesics would work), then it is clear that all the vertical edges (the edges which, once projected on the first cycle, project on a single vertex) have the same probability to lie in $\tilde{\gamma}$; same for horizontal edges. In particular:

$$\sum_{\text{“vertical edges” } e} \mathbb{P}\big[e \in \tilde{\gamma}\big] \leq \mathbb{E}\big[|\tilde{\gamma}|\big] \leq \frac{b}{a} m \,.$$

Since there are $O(1)m^d$ vertical edges, the influence of these is bounded by $O(1)m^{1-d}$. Same thing for horizontal edges. All together this gives the desired assumption needed in lemma VII.3. Applying this lemma, we indeed obtain that

$$\mathrm{Var}(\mathrm{Circ}_m(\omega)) \leq O(1) \frac{m}{\log m} \,,$$

where the constant does not depend on the dimension d (the dimension in fact helps us here, since it makes the influences smaller). □

Remark VII.4. At this point, we know that for any edge e, $\mathbf{I}_e(\mathrm{Circ}_m) = O(\frac{m}{m^d})$. Hence, at least in the case of the torus, one easily deduce from Poincaré's inequality, the theorem by Kesten which says that $\mathrm{Var}(\mathrm{Circ}_m) = O(m)$.

4 Upper bounds on fluctuations in the spirit of KKL

In this section, we prove Lemma VII.3.

Proof:

Similarly as in the proofs of Chapter V, the proof relies on implementing Hypercontractivity in the right way.

$$\begin{aligned}
\mathrm{var}(f) &= \frac{1}{4}\sum_k \sum_{S\neq\emptyset} \frac{1}{|S|}\widehat{\nabla_k f}(S)^2 \\
&\leq \frac{1}{4}\sum_k \sum_{0<|S|<c\log n} \widehat{\nabla_k f}(S)^2 + \frac{O(1)}{\log n}\sum_k \mathbf{I}_k(f)
\end{aligned}$$

Hence it is enough to bound the contribution of small frequencies, $0 < |S| < c\log n$, for some constant c which will be chosen later. As previously we have for any $\rho \in (0,1)$ and using Hypercontractivity,

$$\begin{aligned}
\sum_k \sum_{0<|S|<c\log n} \widehat{\nabla_k f}(S)^2 &\leq \rho^{-2c\log n}\sum_k \|T_\rho \nabla_k f\|_2^2 \\
&\leq \rho^{-2c\log n}\sum_k \|\nabla_k f\|_{1+\rho^2}^2 \\
&= \rho^{-2c\log n}\sum_k \mathbf{I}_k(f)^{2/(1+\rho^2)} \\
&\leq \rho^{-2c\log n}\big(\sup_k \mathbf{I}_k(f)\big)^{\frac{1-\rho^2}{1+\rho^2}}\sum_k \mathbf{I}_k(f) \\
&\leq \rho^{-2c\log n} n^{-\alpha\frac{1-\rho^2}{1+\rho^2}} \sum_k \mathbf{I}_k(f) \ \text{ by our assumption}.
\end{aligned} \tag{VII.1}$$

Now fixing $\rho \in (0,1)$, and then choosing the constant c depending on ρ and α, the lemma follows (by optimizating on the choice of ρ, one could get better constants). □

5 Further discussion

Some words on the proof of Theorem VII.1

The main difficulty here is that the quantity of interest: $f(\omega) := \mathrm{dist}_\omega(0, v)$ is not anymore invariant under a large class of graph automorphisms. This lack of symmetry makes the study of influences more difficult. For example, edges near the endpoints 0 or v have very high influence (of order one). To gain some more symmetry, the authors in [BKS03] rely on a very nice "averaging" procedure. We refer to this paper for more details.

Known lower bounds on the fluctuations

We discussed mainly here, ways to obtain upper bounds on the fluctuations of the shapes in First Passage Percolation. It is worth pointing out that some non-trivial **lower** bounds on the fluctuations are known for $\mathbb{Z}^2$. See [PP94, NP95].

Exercise sheet of Chapter VII

Problem 1. Let $n \geq 1$ and $d \geq 2$. Consider the random metric on the torus $\mathbb{Z}^d/n\mathbb{Z}^d$ as described in this chapter. For any $k \geq 1$, let $\mathcal{A}_n^k$ be the event that the shortest "horizontal" circuit is $\leq k$. If $d \geq 3$, show that for any choice of $k_n = k(n)$, the family of events $\mathcal{A}_n^{k_n}$ is noise sensitive. (Note that the situation here is similar to the Problem 9 in Chapter I.) Finally, discuss the two-dimensional case, $d = 2$ (non-rigorously).

Chapter VIII

Randomized algorithms and noise sensitivity

In this chapter, we explain how the notion of **revealment** for so-called randomized algorithms can in some cases yield direct information concerning the energy spectrum which may allow not only noise sensitivity results but even **quantitative** noise sensitivity results.

1 BKS and Randomized algorithms

In the previous chapter, we explained how Theorem I.5 together with bounds on the pivotal exponent for percolation yields noise sensitivity for percolation crossings. However, in [BKS99], a different approach was in fact used for showing noise sensitivity, which while using Theorem I.5, it did not use these bounds on the critical exponent. In that approach, one sees the first appearance of randomized algorithms. In a nutshell, the authors showed that (1) if a **monotone** function is very uncorrelated with all majority functions, then it is noise sensitive (in a precise quantitative sense) and (2) percolation crossings are very uncorrelated with all majority functions by utilizing certain algorithms which, due to RSW, look at very few bits.

2 The Revealment Theorem

A **algorithm** for a Boolean function f is an algorithm A which queries (asks the values of) the bits one by one, where the decision of which bit to ask can be based on the values of the bits previously queried, and stops once f is determined (being determined means that f takes the same value no matter how the remaining bits are set).

A **randomized algorithm** for a Boolean function f is the same as above but auxiliary randomness may also be used to decide the next value queried (including the first bit). [In computer science, the term randomized decision tree would be used for

our notion of randomized algorithm but we will not use this terminology.]

The following definition of *revealment* will be crucial. Given a randomized algorithm A for a Boolean function f, we let J_A denote the random set of bits queried by A. (Note that this set depends both on the randomness corresponding to the choice of ω and the randomness corresponding to the choice of the algorithm, which are of course independent.)

Definition VIII.1. The **revealment of a randomized algorithm** A for a Boolean function f, denoted by δ_A, is defined by

$$\delta_A := \max_{i \in [n]} \mathbb{P}(i \in J_A).$$

The **revealment of a Boolean function** f, denoted by δ_f, is defined by

$$\delta_f := \inf_A \delta_A$$

where the infimum is over all randomized algorithms A for f.

This section presents a connection between noise sensitivity and randomized algorithms. It will be used later to yield an alternative proof of noise sensitivity for percolation crossings which is not based upon Theorem I.5 as was the approach outlined in the previous section. Two advantages of the algorithmic approach of the present section over that mentioned in the previous section is that it applies to nonmonotone functions and yields a more "quantitative" version of noise sensitivity.

We have only defined algorithms, randomized algorithms and revealment for Boolean functions but the definitions immediately extend to functions $f : \Omega_n \to \mathbb{R}$.

The main theorem of this section is the following.

Theorem VIII.1. *For any function $f : \Omega_n \to \mathbb{R}$ and for each $k = 1, 2, \ldots,$ we have that*

$$\sum_{S \subseteq [n],\, |S| = k} \hat{f}(S)^2 \leq \delta_f \, k \, \|f\|^2, \tag{VIII.1}$$

where $\|f\|$ denotes the L^2 norm of f with respect to the uniform probability measure on Ω and δ_f is the revealment of f.

Before giving the proof, we make some comments to help the reader see what is happening and suggest why a result like this might be true. Our original function is a sum of monomials with coefficients given by the Fourier coefficients. Each time a bit is revealed by the algorithm, we get a new Boolean function obtained by just substituting in the value of the bit we obtained into the corresponding variable. On the algebraic side, those monomials which contain this bit go down by 1 in degree while the other monomials are unchanged. There might however be cancellation in the process which is what we hope for since when the algorithm stops, all the monomials (except the

constant) must have been killed. The way cancellation occurs is illustrated as follows. The Boolean function at some stage might contain $1/3x_2x_4x_5 + 1/3x_2x_4$ and then the bit x_5 might be revealed and take the value -1. When we substitute this value into the variable, the terms cancel and disappear thereby bringing us 1 step closer to a constant (and hence determined) function.
As far as why the result might be true, the intuition, very roughly speaking, is as follows. Letting, $g^k(\omega) := \sum_{|S|=k} \hat{f}(S)\,\chi_S(\omega)$ (whose L^2 norm squared is the left hand side in the statement of the theorem), if the revealment δ is small, then for low level k, there are few "frequencies" in $g^{(k)}$ which will be "seen" by the algorithm. More precisely, for any fixed "frequency" S, if $|S| = k$ is small, then with high probability none of the bits in S will be visited by the algorithm. This means that $\mathbb{E}\big[g^{(k)} \mid J\big]$ should be of small L^2 norm compared to $g^{(k)}$. Now since $f = \mathbb{E}\big[f \mid J\big] = \sum_k \mathbb{E}\big[g^{(k)} \mid J\big]$, most of the Fourier transform should be supported on high frequencies. There is some difficulty in implementing this intuition, since the conditional expectations $\mathbb{E}\big[g^{(k)} \mid J\big]$ are not orthogonal.

Proof. In the following, we let $\tilde{\Omega}$ denote the probability space that includes the randomness in the input bits of f and the randomness used to run the algorithm (which we assume to be independent) and we let $\mathbb{E}$ denote the corresponding expectation. , need? Without loss of generality, elements of $\tilde{\Omega}$ can be represented as $\tilde{\omega} = (\omega, \tau)$ where ω are the random bits and τ represents the randomness necessary to run the algorithm.

Now, fix $k \geq 1$. Let

$$g(\omega) := \sum_{|S|=k} \hat{f}(S)\,\chi_S(\omega)\,, \qquad \omega \in \Omega.$$

The left hand side of (VIII.1) is equal to $\|g\|^2$.

Let $J \subseteq [n]$ be the random set of all bits examined by the algorithm. Let $\mathcal{A}$ denote the minimal σ-field for which J is measurable and every ω_i, $i \in J$, is measurable; this can be viewed as the relevant information gathered by the algorithm. For any function $h : \Omega \to \mathbb{R}$, let $h_J : \Omega \to \mathbb{R}$ denote the random function obtained by substituting the values of the bits in J. More precisely, if $\tilde{\omega} = (\omega, \tau)$ and $\omega' \in \Omega$, then $h_J(\tilde{\omega})(\omega')$ is $h(\omega'')$ where ω'' is ω on $J(\tilde{\omega})$ and is ω' on $[n]\backslash J(\tilde{\omega})$. In this way, h_J is a random variable on $\tilde{\Omega}$ taking values in the set of mappings from Ω to $\mathbb{R}$ and it is immediate that this random variable is $\mathcal{A}$-measurable. When the algorithm terminates, the unexamined bits in Ω are unbiased and hence $\mathbb{E}\big[h \mid \mathcal{A}\big] = \int h_J (= \hat{h_J}(\emptyset))$ where $\int$ is defined, as usual, to be integration with respect to uniform measure on Ω. It follows that $\mathbb{E}[h] = \mathbb{E}[\int h_J]$.

Similarly, for all h,

$$\|h\|^2 = \mathbb{E}\big[h^2\big] = \mathbb{E}\Big[\int h_J^2\Big] = \mathbb{E}\big[\|h_J\|^2\big]. \tag{VIII.2}$$

Since the algorithm determines f, it is $\mathcal{A}$ measurable, and we have

$$\|g\|^2 = \mathbb{E}[g\,f] = \mathbb{E}\Big[\mathbb{E}\big[g\,f \mid \mathcal{A}\big]\Big] = \mathbb{E}\Big[f\,\mathbb{E}\big[g \mid \mathcal{A}\big]\Big].$$

Since $\mathbb{E}\big[g \mid \mathcal{A}\big] = \hat{g_J}(\emptyset)$, Cauchy-Schwarz therefore gives

$$\|g\|^2 \leq \sqrt{\mathbb{E}[\hat{g}_J(\emptyset)^2]}\,\|f\|\,. \tag{VIII.3}$$

We now apply Parseval to the (random) function g_J: this gives (for any $\tilde{\omega} = (\omega, \tau) \in \tilde{\Omega}$),

$$\hat{g}_J(\emptyset)^2 = \|g_J\|_2^2 - \sum_{|S|>0} \hat{g}_J(S)^2.$$

Taking the expectation over $\tilde{\omega} \in \tilde{\Omega}$, this leads to

$$\begin{aligned}
\mathbb{E}\big[\hat{g}_J(\emptyset)^2\big] &= \mathbb{E}\big[\|g_J\|_2^2\big] - \sum_{|S|>0} \mathbb{E}\big[\hat{g}_J(S)^2\big] \\
&= \|g\|_2^2 - \sum_{|S|>0} \mathbb{E}\big[\hat{g}_J(S)^2\big] \quad \text{by (VIII.2)} \\
&= \sum_{|S|=k} \hat{g}(S)^2 - \sum_{|S|>0} \mathbb{E}\big[\hat{g}_J(S)^2\big] \left\{ \begin{array}{l} \text{since } g \text{ is supported} \\ \text{on level-}k \text{ coefficients} \end{array} \right. \\
&\leq \sum_{|S|=k} \mathbb{E}\big[\hat{g}(S)^2 - \hat{g}_J(S)^2\big] \left\{ \begin{array}{l} \text{by restricting to} \\ \text{level-}k \text{ coefficients} \end{array} \right.
\end{aligned}$$

Now, since g_J is built randomly from g by fixing the variables in $J = J(\tilde{\omega})$, and since g by definition does not have frequencies larger than k, it is clear that for any S with $|S| = k$ we have

$$\hat{g}_J(S) = \left\{ \begin{array}{ll} \hat{g}(S) = \hat{f}(S), & \text{if } S \cap J(\tilde{\omega}) = \emptyset \\ 0, & \text{otherwise.} \end{array} \right.$$

Therefore, we obtain

$$\|\mathbb{E}\big[g \mid J\big]\|_2^2 = \mathbb{E}\big[\hat{g}_J(\emptyset)^2\big] \leq \sum_{|S|=k} \hat{g}(S)^2\, \mathbb{P}\big[S \cap J \neq \emptyset\big] \leq \|g\|_2^2\, k\, \delta\,.$$

Combining with (VIII.3) completes the proof. □

Proposition IV.1 and Theorem VIII.1 immediately imply the following corollary.

Corollary VIII.2. *If the revealments satisfy*

$$\lim_{n\to\infty} \delta_{f_n} = 0,$$

then $\{f_n\}$ *is noise sensitive.*

In the exercises, one is asked to show that certain sequences of Boolean functions are noise sensitive by applying the above corollary.

3 An application to noise sensitivity of percolation

In this section, we apply Corollary VIII.2 to prove noise sensitivity of percolation crossings. The following result gives the necessary assumption that the revealments approach 0.

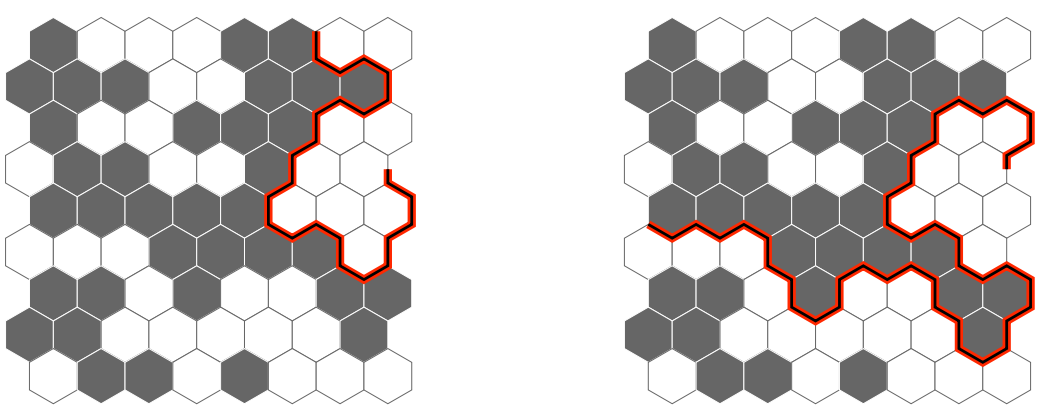
Theorem VIII.3. *Let $f = f_R$ be the indicator function for the event that critical site percolation on the standard triangular grid contains a left to right crossing of our $R \times R$ box. Then $\delta_{f_R} \leq R^{-1/4+o(1)}$ as $R \to \infty$.*

For critical bond percolation on the square grid, this holds with $1/4$ replaced by some positive constant $a > 0$.

Outline of Proof. We first give a first attempt at such an algorithm. We consider from Chapter II the interface from the bottom left of the square to the top right used to detect a left right crossing. This (deterministic) algorithm simply asks the bits that it needs to know in order to continue the interface. Observe that if a bit is queried, it is necessarily the case that there is both a black and white path from next to the hexagon to the boundary. It follows, from the exponent of $1/4$ for the 2 arm event in Chapter II, that for hexagons far from the boundary, the probability that they are revealed is at most $R^{-1/4+o(1)}$ as desired. However, one can not conclude that points near the boundary have small revealment and of course the left bottom point as revealment 1.

The way that we modify the above algorithm so that all points have small revealment is as follows. We first choose a point x at random from the middle third of the left side. We then run two algorithms, the first one which checks whether there is a left right path from the left side *above* x to the right side and the second one which checks whether there is a left right path from the left side *below* x to the right side. The first part is done by looking at an interface from x to the top right corner as above. The second part is done by looking at an interface from x to the bottom right corner as above (but where the colors on the two sides of the interface need to be swapped.)

It can then be shown with a little work (but no new conceptual ideas) that this modified algorithm has the desired revealment of at most $R^{-1/4+o(1)}$ as desired. One of the things that one needs to use in this analysis is the so-called half-plane exponent, which has a known value of $1/3$. $\square$



3.1 First quantitative noise sensitivity result

In this subsection, we give our first "polynomial bound" on the noise sensitivity of percolation. This is a big step in our understanding of **quantitative noise sensitivity** of percolation initiated in the previous Chapter.

Recall that in the definition of noise sensitivity, ϵ is held fixed. However, as we have seen in the previous Chapter, it is of interest to ask if the correlations can still go to 0 when $\epsilon = \epsilon_n$ goes to 0 with n but not so fast. The techniques of the present Chapter imply the following result.

Theorem VIII.4. *Let f_n be as in Theorem VIII.3. Then for all $\gamma < 1/8$,*

$$\lim_{n\to\infty} \mathbb{E}[f_n(\omega)f_n(\omega_{1/n^\gamma})] - \mathbb{E}[f_n(\omega)]^2 = 0. \tag{VIII.4}$$

On the square lattice, there exists some $\gamma > 0$ with the above property.

Proof. We prove only the first statement; the square lattice case is handled similarly. First, (IV.3) gives us that every n and γ

$$\mathbb{E}[f_n(\omega)f_n(\omega_{1/n^\gamma})] - \mathbb{E}[f_n(\omega)]^2 = \sum_{k=1} E_{f_n}(k)(1 - 1/n^\gamma)^k. \tag{VIII.5}$$

Note that there are order n^2 terms in the sum. Fix $\gamma < 1/8$. Choose $\epsilon > 0$ so that $\gamma + \epsilon < 1/8$. For large n, we have that $\delta_{f_n} \le n^{-1/4+\epsilon}$. Break up the right hand side in (VIII.5) into two parts, the first being the sum up to $n^{\gamma+\epsilon/2}$ and the second part being the other terms.

For the terms in the first part, replace $(1 - 1/n^\gamma)^k$ by 1 and use Theorem VIII.1 to bound the $E_{f_n}(k)$ terms by $Cn^{-1/4+\epsilon}k$. By the way ϵ was chosen, it is easy to see that the sum in the first parts approach 0 as $n \to \infty$. Next, using $1 - x \le e^{-x}$ and replacing the $E_{f_n}(k)$ terms by 1, it is also easy to see that the sum in the second parts also approach 0 as $n \to \infty$. □

4 Bounds on revealments

With the Revealment Theorem, we now show that one cannot hope to reach the 3/4-sensitivity exponent. Theorem VIII.4 told us that we obtain asymptotic decorrelation if the noise is $1/n^\gamma$ for $\gamma < 1/8$. This differs from the conjectured "critical exponent" of 3/4 by a factor of 6. In this section, we investigate the degree to which the 1/8 could potentially be improved and in the discussion, we will bring up an interesting open problem.

Given a randomized algorithm A for a Boolean function f, let $C(A)$ (the cost of A) be the expected number of queries that the algorithm A makes. Clearly $n\delta_A \ge C(A)$. Note also that this is at least the total influence $\mathbf{I}(f)$ since for any i, the event that i is pivotal necessarily implies that the bit i is queried. The following result due to O'Donnell and Servedio ([OS07])is an essential improvement on this remark.

Theorem VIII.5. *Let f be a monotone Boolean function mapping Ω_n into $\{-1,1\}$ and A a randomized algorithm for f. Then $C(A) \geq \mathbf{I}(f)^2$ and hence $\delta_A \geq \mathbf{I}(f)^2/n$.*

Proof. Fix a deterministic decision tree obtaining the minimum in the definition of $C(f)$. Let J be the random set of bits queried by this decision tree. We then have

$$\mathbf{I}(f) = \mathbb{E}[\sum_i f(\omega)\omega_i] = \mathbb{E}[f(\omega)\sum_i \omega_i I_{\{i\in J\}}] \leq \sqrt{\mathbb{E}[f(\omega)^2]}\sqrt{\mathbb{E}[(\sum_i \omega_i I_{\{i\in J\}})^2]}$$

where the first equality uses monotonicity. We now bound the first term by 1. For the second moment inside the second square root, the sum of the diagonal terms yields $\mathbb{E}[|J|]$ while the cross terms are all 0 since for $i \neq j$, $\mathbb{E}[\omega_i I_{\{i\in J\}}\omega_j I_{\{j\in J\}}] = 0$ as can be seen by breaking up the sum depending on whether i or j is queried first. This yields the result. □

Returning to our event f_n of percolation crossings, since the sum of the influences is $n^{3/4+o(1)}$, Theorem VIII.5 tells us that $\delta_{f_R} \geq R^{-1/2+o(1)}$. It follows from the method of proof in Theorem VIII.4 that the present version of the Revealment Theorem could never improve the result of Theorem VIII.4 past $\gamma = 1/4$ which is still a factor of 3 from the critical value 3/4. Of course, one could investigate the degree to which the Revealment Theorem itself could be improved.

Theorem VIII.3 tells us that there are algorithms A_R for f_R such that $C(A_R) \leq R^{7/4+o(1)}$. On the other hand, Theorem VIII.5 tell us that it is necessarily the case that $C(A) \geq R^{6/4+o(1)}$.

Open Question: Find the smallest σ such that there are algorithms A_R with $C(A_R) = R^{\sigma+o(1)}$. (We know $\sigma \in [6/4, 7/4]$ if it exists.)

We mention another inequality relating revealment with influences which is a consequence of the results in [OSSS05].

Theorem VIII.6. *Let f be a Boolean function mapping Ω_n into $\{-1,1\}$. Then $\delta_f \geq \mathrm{Var}(f)/n \max_i \mathbf{I}_i(f)$*

It is interesting to compare Theorems VIII.5 and VIII.6. Assuming $\mathrm{Var}(f)$ is of order 1, and all the influences are of order $1/n^\alpha$, then it is easy to check that Theorem VIII.5 gives a better bound when $\alpha < 2/3$ and Theorem VIII.6 gives a better bound when $\alpha > 2/3$. For crossings of percolation, where α should be around 5/8, it is better to use Theorem VIII.5 rather than VIII.6.

Finally, there are a number of interesting results concerning revealment obtained in the paper [BSW05]. Four results are as follows.
1. If f is reasonably balanced, the revealment is at least of order $1/n^{1/2}$.
2. There is a reasonably balanced function whose revealment is at most $O(1)\log n/n^{1/2}$.
3. If f is reasonably balanced and monotone, the revealment is at least of order $1/n^{1/3}$.
4. There is a reasonably balanced monotone function whose revealment is at most $O(1)\log n/n^{1/3}$.

5 Does noise sensitivity imply low revealment?

As far as these lectures are concerned, this subsection will not connect to anything that follows and hence can be viewed as tangential.

It is natural to ask if the converse of Corollary VIII.2 might be true. A quick thought reveals that example 2, Parity, provides a counterexample. However, it is more interesting perhaps that there is a monotone counterexample to the converse which is provided by example 5, Clique containnment.

Proposition VIII.7. *Clique containment provides an example showing that the converse of Corollary VIII.2 is false for monotone functions.*

Outline of Proof. We have already seen in one of the exercises that this example is noise sensitive. To show that the revealments do not go to 0, we prove something stronger. To do this, we must first give a few more definitions.

Definition VIII.2. A Boolean function f is given. A **witness** for ω is any subset W of the variables such that the elements of ω in W determine f in the sense that for every ω' which agrees with ω on W, we have that $f(\omega) = f(\omega')$. The **witness size** of ω, denoted $w(\omega)$, is the size of the smallest witness for ω. The **expected witness size**, denoted by $w(f)$, is $\mathbb{E}(w(\omega))$.

Observe that for any Boolean function f, the bits revealed by any algorithm T for f and for any ω is always a witness for ω. It easily follows that $n\delta_f \geq w(f)$. Therefore, in order to prove the proposition, it suffices to show that

$$w(f_n) = \Omega(n^2).$$

Remark VIII.1. (i). The above also implies that with a fixed uniform probability, $w(\omega)$ is $\Omega(n^2)$.
(ii). Of course when f_n is 1, there is always a (small) witness of size $\binom{k_n}{2} << n$ and so the large average witness size comes from when f_n is -1.
(iii). However, it is not deterministically true that when f_n is -1, $w(\omega)$ is necessarily of size $\Omega(n^2)$. For example, for $\omega \equiv -1$ (corresponding to the empty graph), the witness size is $o(n^2)$ as is easily checked. Clearly this ω has the smallest witness size among ω with $f_n = -1$.

Lemma VIII.8. *Let E_n be the event that all sets of vertices of size at least $.97n$ contains C_{k_n-3}. Then $\lim_{n\to\infty} \mathbb{P}(E_n) = 1$.*

Proof. This follows, after some work, from the Janson inequalities. □

Lemma VIII.9. *Let U be any collection of at most $n^2/1000$ edges in C_n. Then there exist distinct v_1, v_2, v_3 such that*
(1) no edge in U goes between any v_i and v_j and

$$(2)|\{e \in U : e \text{ is an edge between } \{v_1, v_2, v_3\} \text{ and } \{v_1, v_2, v_3\}^c\}| \leq n/50.$$

Proof. We use the probabilistic method where we choose $\{v_1, v_2, v_3\}$ to be a uniformly chosen 3-set. It is immediate that the probability that (1) fails is at most $3|U|/\binom{n}{2} \leq 1/100$. Letting Y be the number of edges in the set in (2) and Y' be the number of U edges touching v_1, it is easy to see that

$$E(Y) \leq 3E(Y') = 6|U|/n \leq n/100$$

where the equality follows from the fact that for any graph, the number of edges is half the total degree. By Markov's inequality, the probability of the event in (2) holds with probably at least 1/2. This shows the random 3-set $\{v_1, v_2, v_3\}$ satisfies (1) and (2) with positive probability and hence such a 3-set exists. □

By Lemma 1, we have $P(A_n^c \cap E_n) \geq c > 0$ for all large n. To prove the theorem, it therefore suffices to show that if $A_n^c \cap E_n$ occurs, there is no witness of size smaller than $n^2/1000$. Assume U be any set of edges of size smaller than $n^2/1000$. Choose $\{v_1, v_2, v_3\}$ from Lemma 2. By (2) of Lemma 2, there exists a set S of size at least $.97n$, which is disjoint from $\{v_1, v_2, v_3\}$ which has no U-edge to $\{v_1, v_2, v_3\}$. Since E_n occurs, S contains a C_{k_n-3}, whose vertices we denote by T. Since there are no U-edges between T and $\{v_1, v_2, v_3\}$ or within $\{v_1, v_2, v_3\}$ (by (1)) and T is the complete graph, U cannot be a witness since A_n^c occured. □

The result of this section is extracted from a paper by Friedgut, Kahn and Wigdersen.

Exercise sheet of Chapter VIII

Exercise 1. Compute the revealment for Majority function on 3 bits.

Exercise 2. Use Corollary VIII.2 to show that Examples 4 and 6, Iterated 3-Majority function and tribes, are noise sensitive.

Exercise 3. Show that for transitive monotone functions, Theorem VIII.5 yields the same result as Theorem VIII.1 does for the case $k = 1$.

Exercise 4. What can you say concerning quantitative noise sensitivity for the Iterated 3-Majority sequence? What can you say about the sequence of revealments for this example?

Exercise 5. You are given a sequence of Boolean functions and told that it is not noise sensitive using noise $\epsilon_n = 1/n^{1/5}$. What, if anything, can you conclude about the sequence of revealments δ_n?

Exercise 6. For transitive monotone functions, is there a relationship between revealment and the minimal cost over all algorithms?

Chapter IX

The Spectral sample

The following is the crucial definition in this chapter.

1 Definition

Definition IX.1. Given a Boolean function $f : \Omega_n \to \{\pm 1\}$ or $\{0, 1\}$, we let the **spectral measure** $\mathbb{Q} = \mathbb{Q}_f$ of f be the measure on subsets $\{1, \ldots, n\}$ given by

$$\mathbb{Q}_f(S) = \hat{f}(S)^2.$$

We let $\mathscr{S}_f = \mathscr{S}$ denote a subset of $\{1, \ldots, n\}$ chosen according to this measure and call this the **spectral sample**. We let $\mathbb{Q}$ also denote the corresponding expectation (even when it is not a probability measure).

Remark IX.1. 1. Note that if f maps into $\{\pm 1\}$, then by Parseval's formula, the distribution of $\mathscr{S}$ is a probability measure while if it maps into $\{0, 1\}$, it will be a subprobability measure.
2. There is no relationship between ω and $\mathscr{S}_f$ as they are defined on different probability spaces. The spectral sample will just be a convenient point of view in order to understand the questions we are studying.

Some of the formulas and results we have previously derived in these notes have very simple formulations in terms of the spectral sample. For example, it is immediate to check that (IV.2) and (IV.3) simply become

$$\mathbb{E}[f(\omega)f(\omega_\epsilon)] = \mathbb{Q}_f[(1-\epsilon)^{|\mathcal{S}|}] \tag{IX.1}$$

and

$$\mathbb{E}[f(\omega)f(\omega_\epsilon)] - (\mathbb{E}[f(\omega)])^2 = \mathbb{Q}_f[(1-\epsilon)^{|\mathcal{S}|} I_{\mathcal{S}\neq\emptyset}]. \tag{IX.2}$$

respectively.

Next, in terms of the spectral sample, Propositions IV.1 and IV.2 simply become the following proposition.

Proposition IX.1. *1. $\{f_n\}$ is noise sensitive if and only if $|\mathscr{S}_{f_n}| \to \infty$ in probability on the set $\{|\mathscr{S}_{f_n}| \neq 0\}$.*
2. $\{f_n\}$ is noise stable if and only if the random variables $\{|\mathscr{S}_{f_n}|\}$ are tight.

There is also a nice relationship between the pivotal set $\mathcal{P}$ and the spectral sample. The following result, which is simply Proposition IV.3 (see also the remark after this proposition), tells us that the two random sets $\mathcal{P}$ and $\mathscr{S}$ have the same 1-dimensional marginals.

Proposition IX.2. *If f is a Boolean function mapping into $\{\pm 1\}$, then for all $i \in [n]$, we have that*

$$\mathbb{P}(i \in \mathcal{P}) = \mathbb{Q}(i \in \mathscr{S})$$

and hence $\mathbb{E}(|\mathcal{P}|) = \mathbb{Q}(|\mathscr{S}|)$.

Even though $\mathscr{S}$ and $\mathscr{P}$ have the same '1-dimensional' marginals, it is not however true that these two random sets have the same distribution. For example, it is easily checked that for $\mathbf{MAJ}_3$, these two distributions are differ. Interestingly, as we will see in the next section, $\mathscr{S}$ and $\mathscr{P}$ also always have the same '2-dimensional' marginals. This will prove useful when applying second moment method arguments.

Before ending this section, let us give an alternative proof of Proposition VI.10 using this point of view of thinking of $\mathscr{S}$ as a random set.

Alternative proof of Proposition VI.10 The statement of the proposition when converted to the spectrum states that for any $a_n \to \infty$

$$\lim_{n\to\infty} \mathbb{Q}(|\mathscr{S}_n| \geq a_n n^2 \alpha_4(n)) = 0.$$

However this immediately follows from Markov's inequality using Propositions VI.9 and IX.2. □

2 A way to sample the Spectral sample in a subdomain

In this section, we describe a method of "sampling" the spectral measure restricted to a subset of the bits. As an application of this, we show that $\mathscr{S}$ and $\mathscr{P}$ in fact have the same 2-dimensional marginals, namely that for all i and j, $\mathbb{P}(i, j \in \mathcal{P}) = \mathbb{Q}(i, j \in \mathscr{S})$.

In order to first get a little intuitition about the spectral measure, we start with an easy proposition.

Proposition IX.3. *For a Boolean function f and $A \subseteq \{1, 2, \ldots, n\}$, we have*

$$\mathbb{Q}(\mathscr{S}_f \subseteq A) = \mathbb{E}[|\mathbb{E}(f|A)|^2]$$

where conditioning on A means conditioning on the bits in A.

Proof. Noting that $\mathbb{E}(\chi_S|A)$ is χ_S if $S \subseteq A$ and 0 otherwise, we obtain by expanding that

$$\mathbb{E}(f|A) = \sum_{S \subseteq A} \hat{f}(S)\, \chi_S.$$

Now apply Parseval's formula. □

If we have a subset $A \subseteq \{1, 2, \ldots, n\}$, how do we "sample" from $A \cap \mathscr{S}$? A nice way to proceed is as follows: choose a random configuration outside of A, then look at the induced function on A and sample from the induced function's spectral measure. The following proposition says exactly this in precise terms. Its proof is just an extension of the proof of Proposition IX.3.

Proposition IX.4. *Fix a Boolean function f on Ω_n. For $A \subseteq \{1, 2, \ldots, n\}$ and $y \in \{\pm 1\}^{A^c}$, that is a configuration on A^c, let g_y be the function defined on $\{\pm 1\}^A$ obtained by using f but fixing the configuration to be y outside of A. Then for any $S \subseteq A$, we have*

$$\mathbb{Q}(\mathscr{S}_f \cap A = S) = \mathbb{E}[\mathbb{Q}(\mathscr{S}_{g_y} = S)] = \mathbb{E}[\hat{g}_y^2(S)].$$

Proof. Using the first line of the proof of Proposition IX.3, it is easy to check that for any $S \subseteq A$, we have that

$$\mathbb{E}\big[f\, \chi_S \;\big|\; \mathcal{F}_{A^c}\big] = \sum_{S' \subseteq A^c} \widehat{f}(S \cup S')\, \chi_{S'}\,.$$

This gives

$$\mathbb{E}\Big[\mathbb{E}\big[f\, \chi_S \;\big|\; \mathcal{F}_{A^c}\big]^2\Big] = \sum_{S' \subseteq A^c} \widehat{f}(S \cup S')^2 = \mathbb{Q}[\mathscr{S} \cap A = S]$$

which is precisely the claim. □

Corollary IX.5. *If f is a Boolean function mapping into $\{\pm 1\}$, then for all i and j,*

$$\mathbb{P}(i, j \in \mathcal{P}) = \mathbb{Q}(i, j \in \mathscr{S}).$$

Proof. Although it has been done, we first show Proposition IX.4 yields equality of the 1-dimensional marginals.

$$\mathbb{Q}(i \in \mathscr{S}) = \mathbb{Q}(\mathscr{S} \cap \{i\} = \{i\}) = \mathbb{E}[\hat{g}_y^2(\{i\})].$$

Note g_y is $\pm\omega_i$ if i is pivotal and constant if i is not pivotal. Hence the above is $\mathbb{P}(i \in \mathcal{P})$.

For the 2-dimensional marginals, one first checks this by hand when $n = 2$. For general n, taking $A = \{i, j\}$ in Proposition IX.4, we have

$$\mathbb{Q}(i, j \in \mathscr{S}) = \mathbb{P}(\mathscr{S} \cap \{i, j\} = \{i, j\}) = \mathbb{E}[\hat{g}_y^2(\{i, j\})].$$

For fixed y, the $n = 2$ case tells us that $\hat{g}_y^2(\{i, j\}) = \mathbb{P}(i, j \in \mathcal{P}_{g_y})$. Now a little thought shows that $\mathbb{E}[\mathbb{P}(i, j \in \mathcal{P}_{g_y})] = \mathbb{P}(i, j \in \mathcal{P})$ completing the proof. □

3 Nontrivial spectrum near the upper bound for percolation

We now return to our central event of percolation crossings of the rectangle R_n and f_n denotes this event. At this point, we know that for $\mathbb{Z}^2$, (most of) the spectrum lies between n^{ϵ_0} (for some $\epsilon_0 > 0$) and $n^2\alpha_4(n)$ while for $\mathbb{T}$, it sits between $n^{1/8+o(1)}$ and $n^{3/4+o(1)}$. In this section, we show that there is a nontrivial amount of spectrum near the upper bound $n^2\alpha_4(n)$. For $\mathbb{T}$, in terms of quantitative noise sensitivity, this tells us that if our noise sequence ϵ_n is equal to $1/n^{3/4-\delta}$ for fixed $\delta > 0$, then in the limit, we know that the two variables $f(\omega)$ and $f(\omega_{\epsilon_n})$ are not perfectly correlated; there is some degree of independence. (See the exercises for understanding such arguments.) However, we cannot conclude that there is full independence since we don't know that 'all' of the spectrum is near $n^{3/4+o(1)}$ (yet!).

Theorem IX.6. *Consider our percolation crossing events f_n of the rectangle R_n for $\mathbb{Z}^2$ or $\mathbb{T}$. There exists $c > 0$ such that for all n,*

$$\mathbb{Q}(|\mathscr{S}_n| \geq cn^2\alpha_4(n)) \geq c.$$

The key lemma for proving this is the following second moment bound on the number of pivotals which we prove afterwards. It has a similar flavor to Exercise 6 in Chapter VI.

Lemma IX.7. *Consider our percolation crossing events f_n for $\mathbb{Z}^2$ or $\mathbb{T}$ and let R'_n be the box concentric with R_n with half the radius. If $X_n = |\mathcal{P}_n \cap R'_n|$ is the cardinality of the set of pivotal points in R'_n, then there exists a constant C such that for all n we have that*

$$\mathbb{E}[|\mathcal{P}_n|^2] \leq C\mathbb{E}[|\mathcal{P}_n|]^2.$$

Proof of Theorem IX.6. Since $\mathcal{P}_n$ and $\mathscr{S}_n$ have the same 1 and 2 dimensional marginals, we also have that for all n

$$\mathbb{Q}[|\mathscr{S}_n \cap R'_n|^2] \leq C\mathbb{Q}[|\mathscr{S}_n \cap R'_n|]^2.$$

Recall now the Paley-Zygmund inequality which states that if $Z \geq 0$, then for all $\theta \in (0,1)$,

$$P(Z \geq \theta\mathbb{E}[Z]) \geq (1-\theta)^2\frac{\mathbb{E}[Z]^2}{\mathbb{E}[Z^2]}.$$

The two above inequalities (take $\theta = 1/2$) imply that for all n

$$\mathbb{Q}(|\mathscr{S}_n \cap R'_n| \geq \mathbb{Q}[|\mathscr{S}_n \cap R'_n|]/2) \geq 1/(4C).$$

(A slight variant of) Proposition VI.9 now completes the proof. □

We are now left with

Proof of Lemma IX.7. For $x, y \in R'_n$, a picture shows that

$$\mathbb{P}(x, y \in \mathcal{P}_n) \leq \alpha_4^2(|x-y|/2)\alpha_4(2|x-y|, n/2)$$

since we need to have the 4 arm event around x to distance $|x-y|/2$, the same for y, and the 4 arm event in the annulus centered at $(x+y)/2$ from distance $2|x-y|$ to distance $n/2$ and finally these three events are independent. This is by quasi-multiplicity at most

$$O(1)\alpha_4^2(n)/\alpha_4(|x-y|, n)$$

and hence

$$\mathbb{E}[|\mathcal{P}_n|^2] \leq O(1)\alpha_4^2(n) \sum_{x,y} \frac{1}{\alpha_4(|x-y|, n)}.$$

Since for a given x, there are at most $O(1)2^{2k}$ y's with $|x-y| \in [2^k, 2^{k+1}]$, using quasi-multiplicity, the above sum is at most

$$O(1)n^2\alpha_4^2(n) \sum_{k=0}^{\log_2(n)} \frac{2^{2k}}{\alpha_4(2^k, n)}.$$

Using

$$\frac{1}{\alpha_4(r, R)} \leq (R/r)^{2-\epsilon}$$

(this is the fact that the four arm exponent is strictly less than 2), the sum becomes at most

$$O(1)n^{4-\epsilon}\alpha_4^2(n) \sum_{k=0}^{\log_2(n)} 2^{k\epsilon}.$$

Since the last sum is at most $O(1)n^\epsilon$, we are done. □

In terms of the consequences for quantitative noise sensitivity, Theorem IX.6 implies the following corollary; see the exercises. We state this only for the triangular lattice.

Corollary IX.8. *For $\mathbb{T}$, if $\epsilon_n = 1/n^{3/4-\delta}$ for a fixed $\delta > 0$, then there exists $c > 0$ such that for all n,*

$$\mathbb{P}(f_n(\omega) \neq f_n(\omega_{\epsilon_n})) \geq c.$$

Note importantly, this does *not* say that $f_n(\omega)$ and $f_n(\omega_{\epsilon_n})$ become asymptotically uncorrelated, only that they are not completely correlated. To insure that they are asymptotically uncorrelated is significantly harder and requires showing that "all" of the spectrum is near $n^{3/4}$. This more difficult task is the subject of the next chapter.

4 Further discussion

It is our understanding that it was Gil Kalai who suggested that thinking of the spectrum as a random set could shed some light on these types of questions.

Exercise sheet on chapter IX

Exercise 1. Let f_n be an arbitrary sequence of Boolean functions with corresponding spectral samples $\mathcal{S}_n$.
(i) Show that $\mathbb{P}(0 < |\mathcal{S}_n| \leq A_n) \to 0$ implies that $\mathbb{E}[(1-\epsilon_n)^{|\mathcal{S}_n|} I_{\mathcal{S}_n \neq \emptyset}] \to 0$ if $\epsilon_n A_n \to \infty$.
(ii) Show that $\mathbb{E}[(1-\epsilon_n)^{|\mathcal{S}_n|} I_{\mathcal{S}_n \neq \emptyset}] \to 0$ implies that $\mathbb{P}(0 < |\mathcal{S}_n| \leq A_n) \to 0$ if $\epsilon_n A_n = O(1)$.

Exercise 2. Prove Corollary IX.8

Exercise 3. For the iterated 3-Majority sequence, recall that the total influence is n^α where $\alpha = 1 - \log 2/\log 3$. Show that for $\epsilon_n = 1/n^\alpha$, $\mathbb{P}(f_n(\omega) \neq f_n(\omega_{\epsilon_n}))$ does not tend to 0.

Exercise 4. Construct an example of a sequence of Boolean functions with mean 0 such that for $\epsilon_n = 1/n^{1/2}$, $\mathbb{P}(f_n(\omega) \neq f_n(\omega_{\epsilon_n}))$ does not tend to 0 nor to 1/2. (Not going to 1/2 means that they are not asymptotically independent.)

Exercise 5. Assume that f_n is a sequence of monotone Boolean functions on n bits with total influence equal to $n^{1/2}$ up to constants. Show that the sequence cannot be noise sensitive.

Exercise 6. Assume that f_n is a sequence of monotone Boolean functions on n bits. Show that one cannot have noise sensitivity when using noise level $\epsilon_n = 1/n^{1/2}$

Exercise 7. (challenging problem) Do you expect exercise 5 is sharp meaning that if 1/2 is replaced by $\alpha < 1/2$, then one can find noise sensitive examples?

Chapter X

Sharp noise sensitivity of percolation

Under construction ...

Chapter XI

Applications to dynamical percolation

In this section, we present a very natural model where percolation undergoes a time-evolution: this is the model of **dynamical percolation** described below. The study of the "dynamical" behavior of percolation as opposed to its "static" behavior turns out to be very rich: interesting phenomena arise especially at the phase transition point. We will see that in some sense, dynamical **planar** percolation at criticality is a very unstable (or chaotic) process. In order to understand this instability, sensitivity of percolation (and therefore its Fourier analysis) will play a key role. In fact, the original motivation for the paper [BKS99] on noise sensitivity was to solve a particular problem in the subject of dynamical percolation. [Ste09] provides a recent survey on the subject of dynamical percolation.

1 The model of dynamical percolation

This model was introduced by Häggström, Peres and Steif [HPS97] inspired by a question that Paul Malliavin asked at a lecture at the Mittag Leffler Institute in 1995. This model was invented independently by Itai Benjamini.

In the general version of this model as it was introduced, given an arbitrary graph G and a parameter p, the edges of G switch back and forth according to independent 2-state continuous time Markov chains where closed switches to open at rate p and open switches to closed at rate $1-p$. Clearly, the product measure with density p, denoted by π_p, is the unique stationary distribution for this Markov process. The general question studied in dynamical percolation is whether, when we start with distribution π_p, there exist atypical times at which the percolation structure looks markedly different than that at a fixed time. In almost all cases, the term "markedly different" refers to the existence or nonexistence of an infinite connected component. Dynamical percolation

on site percolation models, which includes our most important case of the hexagonal lattice, is defined analogously.

We very briefly summarize a few early results in the area. It was shown in [HPS97] that below criticality, there are no times at which there is an infinite cluster and above criticality, there is an infinite cluster at all times. See the exercises. In this paper, examples of graphs which do not percolate at criticality but for which there exist exceptional times where percolation occurs were given. (Also given were examples of graphs which do percolate at criticality but for which there exist exceptional times where percolation does not occur.) A fairly refined analysis of the case of so called *spherically symmetric* trees was given. See the exercises.

Given the above results, it is natural for the reader to ask what happens on the standard graphs that we work with. Recall that for $\mathbb{Z}^d$ it is only known for $d = 2$ and $d \geq 19$ that we don't percolate at criticality.

2 What's going on in high dimensions: $\mathbb{Z}^d, d \geq 19$?

For the high dimensional case, $\mathbb{Z}^d, d \geq 19$, it was shown in [HPS97] that there are no exceptional times of percolation at criticality.

Theorem XI.1. *[HPS97] For the integer lattice $\mathbb{Z}^d$ with $d \geq 19$, dynamical critical percolation has no exceptional times of percolation.*

The key reason for this is a highly nontrivial result due to work of Hara and Slade ([HS94]), using earlier work of Aizenman and Barsky ([BA91]), that says that if $\theta(p)$ is the probability that the origin percolates when the parameter is p, then for $p \geq p_c$

$$\theta(p) = O(p - p_c)\,. \tag{XI.1}$$

In fact, this is the only thing which is used in the proof.

Outline of Proof. By countable additivity, it suffices to show that there are no times at which the origin percolates during $[0, 1]$. We use a first moment argument. We break the time interval $[0, 1]$ into m intervals each of length $1/m$. If we fix one of these intervals, the set of edges which are open at *some time* during this interval has density about $p_c + 1/m$. Hence the probability that the origin percolates with respect to these set of edges is by (XI.1) at most $O(1/m)$. It follows that if N_m is the number of intervals where this occurs, then $\mathbb{E}[N_m]$ is at most $O(1)$. It is not hard to check that $N \leq \liminf_m N_m$ where N is the number of times during $[0, 1]$ at which the origin percolates. Fatou's lemma now yields that $\mathbb{E}(N) < \infty$ and hence there are at most finitely many exceptional times during $[0, 1]$ at which the origin percolates. To go from there to no exceptional times can either be done by using some rather abstract Markov process theory or by a more hands on approach as was done in [HPS97]. $\square$

Remark XI.1. It is known that (XI.1) holds for any homogeneous tree (see [Gri99] for the binary tree case) and hence there are no exceptional times of percolation in this case also.

Remark XI.2. It is was proved by Kesten and Zhang [KZ87], that (XI.1) fails for $\mathbb{Z}^2$ and hence the proof method above fails. This infinite derivative in this case might suggest that there are in fact exceptional times for critical dynamical percolation on $\mathbb{Z}^2$, a question left open in [HPS97].

3 $d = 2$ and BKS

One of the questions posed in [HPS97] was whether there are exceptional times of percolation for the 2-dimensional lattice. It was this question which was one of the main motivations for the paper [BKS99]. While they did not prove the existence of exceptional times of percolation, they did obtain the following very interesting result which has a very similar flavor.

Theorem XI.2. *Consider an $R \times R$ on which we run critical dynamical percolation. Let S_R be the number of times during $[0, 1]$ at which the configuration changes from having a percolation crossing to not have one. Then*

$$S_R \to \infty \text{ in probability as } R \to \infty.$$

Noise sensitivity of percolation as well as the above theorem tells us that certain large scale connectivity properties decorrelate very fast. This suggests that in some vague sense $\omega_t^{p_c}$ "changes" very quickly as time goes on and hence there might be some chance that an infinite cluster appears.

In the next section, we begin our study of exceptional times for $\mathbb{Z}^2$ and the hexagonal lattice.

4 The second moment method and the spectrum

In this section, we reduce the question of exceptional times to a 'second moment method' computation which in turn reduces to questions concerning the spectral behavior for specific Boolean functions involving percolation. Since $p = 1/2$, our dynamics can be equivalently defined by having each edge or hexagon be rerandomized at rate 1.

The key random variable which one needs to look at is

$$X = X_R := \int_0^1 1_{0 \overset{\omega_t}{\longleftrightarrow} R} \, dt$$

where $0 \overset{\omega_t}{\longleftrightarrow} R$ is of course the event that at time t there is an open path from the origin to distance R away and the above integral is simply the Lebesgue measure of the set of times in $[0, 1]$ at which this occurs.

We want to apply the second moment method here. We isolate the easy part of the argument so that the reader who is not familiar with this method understands it in a more general context. However, the reader should keep in mind that the difficult part is always to prove the needed bound on the second moments.

Proposition XI.3. *If there exists a constant C such that for all R*

$$\mathbb{E}(X_R^2) \leq C\mathbb{E}(X_R)^2, \tag{XI.2}$$

then there are exceptional times of percolation.

Proof. For any nonnegative random variable Y, the Cauchy-Schwarz inequality applied to $YI_{\{Y>0\}}$ yields

$$\mathbb{P}(Y > 0) \geq \mathbb{E}(Y)^2/\mathbb{E}(Y^2).$$

Hence we have that for all R,

$$\mathbb{P}(X_R > 0) \geq 1/C$$

and hence by countable additivity

$$\mathbb{P}(\cap_R\{X_R > 0\}) \geq 1/C.$$

Had the set of times that a fixed edge is on been a closed set, then the above would have yielded by compactness that there is an exceptional time of percolation with probability at least $1/C$. However, this is not a closed set. On the other hand, this point is very easily fixed by modifying the process so that the times each edge is on is a closed set and observing that a.s. no new times of percolation are introduced by this modification. The details are left to the reader. □

The first moment of X_R is, due to Fubini's Theorem, simply the probability of our 1 arm event α_R. The second moment of X_R is easily seen to be

$$\mathbb{E}(X^2) = \mathbb{E}(\int_0^1 \int_0^1 1_{0\overset{\omega_s}{\longleftrightarrow}R} 1_{0\overset{\omega_t}{\longleftrightarrow}R}\, ds\, dt) = \int_0^1 \int_0^1 \mathbb{P}(0 \overset{\omega_s}{\longleftrightarrow} R, 0 \overset{\omega_t}{\longleftrightarrow} R)\, ds\, dt \tag{XI.3}$$

which is, by time invariance, at most

$$2\int_0^1 \mathbb{P}(0 \overset{\omega_s}{\longleftrightarrow} R, 0 \overset{\omega_0}{\longleftrightarrow} R)\, ds. \tag{XI.4}$$

The key observation now, which brings us back to noise sensitivity, is that the integrand $\mathbb{P}(0 \overset{\omega_s}{\longleftrightarrow} R, 0 \overset{\omega_0}{\longleftrightarrow} R)$ is precisely $\mathbb{E}[f_R(\omega)f_R(\omega_\epsilon)]$ where f_R is indicator of the event that there is an open path from the origin to distance R away and $\epsilon = 1 - e^{-s}$ since looking at our process at two different times is exactly looking at a configuration and a noisy version.

What we have seen in this subsection is that proving the existence of exceptional times comes down to proving a second moment estimate and furthermore that the integrand in this second moment estimate concerns noise sensitivity, something for which we have already developed a fair number of tools.

5 Proof of existence of exceptional times for the hexagonal lattice via randomized algorithms

In [SS10b], exceptional times were shown to exist for the hexagonal lattice; this was the first transitive graph for which this result was obtained. However the methods in this paper did not allow the authors to prove that $\mathbb{Z}^2$ had exceptional times.

Theorem XI.4. *For dynamical percolation on the hexagonal lattice $\mathbb{T}$ at the critical point $p_c = 1/2$, one has Almost surely, there exist exceptional times $t \in [0, \infty]$ such that ω_t has an infinite cluster.*

Proof. We first note that, since two different times of our model can be viewed as "noising" and the probability that a hexagon is rerandomized within t units of time is $1 - e^{-t}$, we have that

$$\mathbb{P}\big[0 \overset{\omega_0}{\longleftrightarrow} R,\, 0 \overset{\omega_t}{\longleftrightarrow} R\big] = \mathbb{E}\big[f_R\big]^2 + \sum_{\emptyset \neq S \subseteq B(0,R)} \hat{f}_R(S)^2 \exp(-t|S|) \tag{XI.5}$$

where $B(0, R)$ are the set of hexagons involved in f_R. We see in this expression, that for small times t, the frequencies contributing in the correlation between $\{0 \overset{\omega_0}{\longleftrightarrow} R\}$ and $\{0 \overset{\omega_t}{\longleftrightarrow} R\}$ are of 'small' size $|S| \lesssim 1/t$. In particular, in order to detect the existence of exceptional times,one needs to achieve a good control on the **Lower tail** of the **Fourier Spectrum** of f_R.

What remains to be done is to find an algorithm minimizing the revealment as much as possible. However there is a difficulty here, since our algorithm might have to look near the origin in which case it is difficult to keep the revealment small. There are other reasons for a potential problem. If R is very large and t very small, then if one conditions on the event $\{0 \overset{\omega_0}{\longleftrightarrow} R\}$, since few sites are updated, the open path in ω_0 from 0 to distance R will still be preserved in ω_t at least up to some distance $L(t)$ (further away, large scale connections start being "noise sensitive"). In some sense the geometry associated to the event $\{0 \overset{\omega}{\longleftrightarrow} R\}$ is "frozen" on a certain scale between time 0 and time t. Therefore, it is natural to divide our correlation analysis into two scales: the ball of radius $r = r(t)$ and the annulus from $r(t)$ to R. Obviously the "frozen radius" $r = r(t)$ increases as $t \to 0$. We proceed as follows instead.

$$\begin{aligned} \mathbb{P}\big[0 \overset{\omega_0}{\longleftrightarrow} R,\, 0 \overset{\omega_t}{\longleftrightarrow} R\big] &\leq \mathbb{P}\big[0 \overset{\omega_0}{\longleftrightarrow} r\big]\, \mathbb{P}\big[r \overset{\omega_0}{\longleftrightarrow} R,\, r \overset{\omega_t}{\longleftrightarrow} R\big] \\ &\leq \alpha_1(r)\, \mathbb{E}\big[f_{r,R}(\omega_0) f_{r,R}(\omega_t)\big], \end{aligned} \tag{XI.6}$$

where $f_{r,R}$ is the indicator function of the event denoted by $r \overset{\omega}{\longleftrightarrow} R$ that there is an open path from distance r away to distance R away. Now, as above, we have

$$\mathbb{E}\big[f_{r,R}(\omega_0) f_{r,R}(\omega_t)\big] \leq \mathbb{E}\big[f_{r,R}\big]^2 + \sum_{k=1}^{\infty} \sum_{|S|=k} \hat{f}_{r,R}(S)^2 \exp(-t|S|). \tag{XI.7}$$

The Boolean function $f_{r,R}$ somehow avoids the singularity at the origin, and it is possible to find algorithms for this function with small revealments. In any case, letting $\delta = \delta_{r,R}$ be the revealment of $f_{r,R}$, it follows Theorem VIII.1 and the fact that $\sum_k k \exp(-tk) \leq O(1)/t^2$ that

$$\mathbb{E}\big[f_{r,R}(\omega_0) f_{r,R}(\omega_t)\big] \leq \alpha_1(r,R)^2 + O(1)\delta\alpha_1(r,R)/t^2. \tag{XI.8}$$

The following proposition gives a bound on δ. We will sketch why it is true afterwards.

Proposition XI.5. *Let $2 \leq r < R$. There is a randomized algorithm A for f_r^R satisfying*

$$\delta_A \leq O(1)\alpha(r,R)\,\alpha_2(r)\,. \tag{XI.9}$$

Putting together (XI.6), (XI.8), Proposition XI.5 and using quasimultiplicativity of α_1 yields

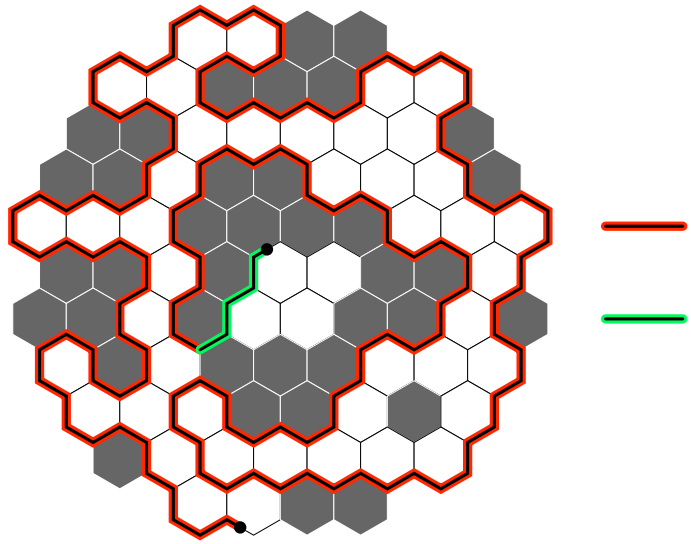
$$\mathbb{P}\big[0 \overset{\omega_0}{\longleftrightarrow} R,\, 0 \overset{\omega_t}{\longleftrightarrow} R\big] \leq O(1)\frac{\alpha_1(R)^2}{\alpha_1(r)}\left(1 + \frac{\alpha_2(r)}{t^2}\right).$$

This is true for all r and t. If we choose $r = r(t) = (1/t)^8$ and ignore $o(1)$ terms in the critical exponents (which can easily be handled anyway), we obtain, using the values of the critical exponents, that

$$\mathbb{P}\big[0 \overset{\omega_0}{\longleftrightarrow} R,\, 0 \overset{\omega_t}{\longleftrightarrow} R\big] \leq t^{-5/6}\alpha_1(R)^2\,. \tag{XI.10}$$

Now, since $\int_0^1 t^{-5/6}dt < \infty$, by integrating the above correlation bound over the unit interval, one obtains that $\mathbb{E}\big[X_R^2\big] \leq C\mathbb{E}\big[X_R\big]^2$ for some universal $C > 0$ as desired. $\square$

Outline of Proof of Proposition XI.5.
We use an algorithm that mimics the one we used for percolation crossings except the present setup is "radial". As in the chordal case, we randomize the starting point of our exploration process by choosing a site uniformly on the 'circle' of radius R. Then, we explore the picture with an exploration path γ directed towards the origin; this means that as in the case of crossings, when the interface encounters an open (resp closed) site, it turns say on the right (resp left), the only difference being that when the exploration path closes a loop around the origin, it continues its exploration inside the connected component of the origin. (It is known that this discrete curve converges towards **radial** SLE_6 on $\mathbb{T}$, when the mesh goes to zero.) It turns out that the so-defined exploration path gives all the information we need. Indeed, if the exploration path closes a clockwise loop around the origin, this means that there is a closed circuit around the origin making $f_{r,R}$ equal to zero. On the other hand, if the exploration path does not close any clockwise loop until it reaches radius r, it means that $f_{r,R} = 1$. Hence, we run the exploration path until either it closes a clockwise loop or it reaches radius r. This is our algorithm. Neglecting boundary issues (points near radius r or R),



if x is a point at distance u from 0, with $2r < u < R/2$, in order for x to be examined by the algorithm, it is needed that there is an open path from $2u$ to R and the 2 arm event holds in the ball centered at u with radius $u/2$. Hence the probability $\mathbb{P}\big[x \in J\big]$, for $|x| = u$ is of order $\alpha_2(u)\alpha_1(u, R)$. Due to the value of the 1 and 2 arm exponents, this expression is decreasing in u. Hence, ignoring the boundary, the revealment is at most $O(1)\alpha_2(r)\alpha_1(r, R)$. $\square$

6 Proof of existence of exceptional times via the geometric approach of the spectrum

Recall that our third approach for proving the noise sensitivity of percolation crossings was based on a geometrical analysis of the spectrum, viewing the spectrum as a random set. This approach yielded the exact noise sensitivity exponent for percolation crossings for the hexagonal lattice. This approach can also be used here as we will now explain. Two big advantages of this approach are that it proved the existence of exceptional times for percolation crossings on $\mathbb{Z}^2$, something which [SS10b] was not able to do, as well as proved the exact Hausdorff dimension for the set of exceptional times.

TO BE DONE!

7 Hausdorff dimensions

In this section, we assume that the reader is familiar with the notion Hausdorff dimension. We let $\mathcal{E} \subset [0, \infty]$ denote the (random) set of these exceptional times at which percolation occurs. It is an immediate consequence of Fubini's Theorem that $\mathcal{E}$ has

Lebesgue measure zero and hence we should look at its Hausdorff dimension if we want to measure its 'size'. The first result is the following

Theorem XI.6. *[SS10b] The Hausdorff dimension of $\mathcal{E}$ is an almost sure constant in* $[1/6, 31/36]$.

It was conjectured that the dimension of these exceptional times is a.s. 31/36.

Outline of Proof. The fact that it the dimension is an almost sure constant follows from easy 0-1 Laws. The lower bounds are obtained by placing a random $\mathcal{E}$ with finite so-called α–energies for any $\alpha < 1/6$ and using a result called Frostman's Theorem. Basically, the 1/6 comes from the fact that for any $\alpha < 1/6$, one can multiply the integrand in $\int_0^1 t^{-5/6} dt$ by $(1/t)^\alpha$ and still be integrable. It is the amount of 'room to spare' you have. If one could obtain better estimates on the correlations, one could thereby improve the lower bounds on the dimension. The upper bound is obtained via a first moment argument similar to the proof of Theorem XI.1. □

Using the geometric approach to the spectrum, it was proved that the upper bound in the previous result is the exact dimension as conjectured.

Theorem XI.7. *[GPS08] The Hausdorff dimension of $\mathcal{E}$ is almost surely* 31/36.

8 Some "super exceptional" events

It turns out that there are other types of exceptional events which occur although less 'often' than what we have been looking at and hence might be called "super exceptional" events. For example, it was shown in [GPS08] that the Hausdorff dimension of the set of times at which the origin percolates in the upper half plane is 5/9; that this was an upper bound was established in [SS10b].

Exercise sheet of Chapter XI

Exercise 1. Prove that on any graph below criticality, there are no times at which there is an infinite cluster while above criticality, there is an infinite cluster at all times.

Exercise 2. (Somewhat hard). A *spherically symmetric* tree is one where all vertices at a given level have the same number of children, although this number may depend on the given level. Let T_n be the number of vertices at the nth level. Show that there is percolation at p if

$$\sum_n \frac{1}{p^{-n} T_n} < \infty$$

Hint: Let X_n be the number of vertices in the nth level which are connected to the root. Apply the second moment method to the sequence of X_n's.

The convergence of the sum is also necessary for percolation but this you are not asked to show. This theorem is due to Russell Lyons.

Exercise 3. Show that if T_n is $n^2 2^n$ up to multiplicative constants, then the critical value of the graph is $1/2$ and we percolate at the critical value. (This yields a graph which percolates at the critical value.)

Exercise 4. (Quite a bit harder). Consider dynamical percolation on a spherically symmetric tree. Show that there for the parameter p, there are **exceptional times** at which one percolates if

$$\sum_n \frac{1}{n p^{-n} T_n} < \infty$$

Hint: Find an appropriate random variable X_n to which the second moment method can be applied.

Exercise 5. Find a spherically symmetric tree which does not percolate at criticality but for which there are exceptional times at which percolation occurs.

Bibliography

[BA91] D. J. Barsky and M. Aizenman. Percolation critical exponents under the triangle condition. *Ann. Probab.*, 19(4):1520–1536, 1991.

[BR08] Michel Benaïm and Raphaël Rossignol. Exponential concentration for first passage percolation through modified Poincaré inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(3):544–573, 2008.

[BKS99] Itai Benjamini, Gil Kalai, and Oded Schramm. Noise sensitivity of Boolean functions and applications to percolation. *Inst. Hautes Études Sci. Publ. Math.*, (90):5–43 (2001), 1999.

[BKS03] Itai Benjamini, Gil Kalai, and Oded Schramm. First passage percolation has sublinear distance variance. *Ann. Probab.*, 31(4):1970–1978, 2003.

[BSW05] Itai Benjamini, Oded Schramm, and David B. Wilson. Balanced Boolean functions that can be evaluated so that every input bit is unlikely to be read. In *STOC'05: Proceedings of the 37th Annual ACM Symposium on Theory of Computing*, pages 244–250. ACM, New York, 2005.

[BR06] Béla Bollobás and Oliver Riordan. A short proof of the Harris-Kesten theorem. *Bull. London Math. Soc.*, 38(3):470–484, 2006.

[BKK+92] Jean Bourgain, Jeff Kahn, Gil Kalai, Yitzhak Katznelson, and Nathan Linial. The influence of variables in product spaces. *Israel J. Math.*, 77(1-2):55–64, 1992.

[FK96] Ehud Friedgut and Gil Kalai. Every monotone graph property has a sharp threshold. *Proc. Amer. Math. Soc.*, 124(10):2993–3002, 1996.

[Fri98] Ehud Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):27–35, 1998.

[Fri04] Ehud Friedgut. Influences in product spaces: KKL and BKKKL revisited. *Combin. Probab. Comput.*, 13(1):17–29, 2004.

[Gar10] Christophe Garban. Oded Schramm's contributions to noise sensitivity, 2010. to appear in the *Annals of Probability* volume in honor of Oded Schramm.

[GPS08] Christophe Garban, Gábor Pete, and Oded Schramm. The Fourier spectrum of critical percolation. *Acta Math.*, to appear, 2008.

[Gri99] Geoffrey Grimmett. *Percolation*. Grundlehren der mathematischen Wissenschaften 321. Springer-Verlag, Berlin, second edition, 1999.

[Gro75] Leonard Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.

[HPS97] Olle Häggström, Yuval Peres, and Jeffrey E. Steif. Dynamical percolation. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(4):497–528, 1997.

[HS94] Takashi Hara and Gordon Slade. Mean-field behaviour and the lace expansion. In *Probability and phase transition (Cambridge, 1993)*, volume 420 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 87–122. Kluwer Acad. Publ., Dordrecht, 1994.

[Joh00] Kurt Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.*, 209(2):437–476, 2000.

[KKL88] Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on boolean functions. *29th Annual Symposium on Foundations of Computer Science*, (68–80), 1988.

[Kes80] Harry Kesten. The critical probability of bond percolation on the square lattice equals $\frac{1}{2}$. *Comm. Math. Phys.*, 74(1):41–59, 1980.

[Kes87] Harry Kesten. Scaling relations for 2D-percolation. *Comm. Math. Phys.*, 109(1):109–156, 1987.

[KZ87] Harry Kesten and Yu Zhang. Strict inequalities for some critical exponents in two-dimensional percolation. *J. Statist. Phys.*, 46(5-6):1031–1055, 1987.

[LSW02] Gregory F. Lawler, Oded Schramm, and Wendelin Werner. One-arm exponent for critical 2D percolation. *Electron. J. Probab.*, 7:no. 2, 13 pp. (electronic), 2002.

[Mar74] G. A. Margulis. Probabilistic characteristics of graphs with large connectivity. *Problemy Peredači Informacii*, 10(2):101–108, 1974.

[Nel66] Edward Nelson. A quartic interaction in two dimensions. In *Mathematical Theory of Elementary Particles (Proc. Conf., Dedham, Mass., 1965)*, pages 69–73. M.I.T. Press, Cambridge, Mass., 1966.

[NP95] Charles M. Newman and Marcelo S. T. Piza. Divergence of shape fluctuations in two dimensions. *Ann. Probab.*, 23(3):977–1005, 1995.

[Nol09] Pierre Nolin. Near-critical percolation in two dimensions. *Electron. J. Probab.*, 13(55):1562—1623, 2009.

[O'D] Ryan O'Donnell. History of the hypercontractivity theorem.

[O'D03] Ryan O'Donnell. *Computational Applications Of Noise Sensitivity.* PhD thesis, M.I.T., 2003.

[OS07] Ryan O'Donnell and Rocco A. Servedio. Learning monotone decision trees in polynomial time. *SIAM J. Comput.*, 37(3):827–844 (electronic), 2007.

[OSSS05] Ryan O'Donnell, Michael Saks, Oded Schramm, and Rocco Servedio. Every decision tree has an influential variable. *FOCS*, 2005.

[PP94] R. Pemantle and Y. Peres. Planar first-passage percolation times are not tight. In *Probability and phase transition (Cambridge, 1993)*, volume 420 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 261–264. Kluwer Acad. Publ., Dordrecht, 1994.

[Rus81] Lucio Russo. On the critical percolation probabilities. *Z. Wahrsch. Verw. Gebiete*, 56(2):229–237, 1981.

[Rus82] Lucio Russo. An approximate zero-one law. *Z. Wahrsch. Verw. Gebiete*, 61(1):129–139, 1982.

[Sch00] Oded Schramm. Scaling limits of loop-erased random walks and uniform spanning trees. *Israel J. Math.*, 118:221–288, 2000.

[SS10a] Oded Schramm and Stanislav Smirnov. On the scaling limits of planar percolation. *Ann. Probab.*, to appear, 2010.

[SS10b] Oded Schramm and Jeffrey Steif. Quantitative noise sensitivity and exceptional times for percolation. *Ann. Math.*, 171(2):619–672, 2010.

[Smi01] Stanislav Smirnov. Critical percolation in the plane: conformal invariance, Cardy's formula, scaling limits. *C. R. Acad. Sci. Paris Sér. I Math.*, 333(3):239–244, 2001.

[SW01] Stanislav Smirnov and Wendelin Werner. Critical exponents for two-dimensional percolation. *Math. Res. Lett.*, 8(5-6):729–744, 2001.

[Ste09] Jeffrey Steif. A survey of dynamical percolation. *Fractal geometry and stochastics, IV, Birkhauser*, pages 145–174, 2009.

[Tal94] Michel Talagrand. On Russo's approximate zero-one law. *Ann. Probab.*, 22(3):1576–1587, 1994.

[Tal96] Michel Talagrand. How much are increasing sets positively correlated? *Combinatorica*, 16(2):243–258, 1996.

[Wer07] Wendelin Werner. *Lectures on two-dimensional critical percolation.* IAS Park City Graduate Summer School, 2007.