Proof of sharpness of Bernoulli percolation on $\mathbb{Z}^d$ using an inequality for boolean functions

Victor Keil

School of Science

Bachelor's thesis
Espoo 25.9.2025

Supervisor

Prof. Kalle Kytölä

Advisor

Prof. Kalle Kytölä

Aalto University
School of Science

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the bachelor's thesis

Author Victor Keil	
Title Proof of sharpness of Bernoulli percolation on $\mathbb{Z}^d$ using an inequality for boolean functions	
Degree programme Bachelor's Programme in Science and Technology	
Major Mathematics and Systems Sciences	**Code of major** SCI3029
Teacher in charge Prof. Kalle Kytölä	
Advisor Prof. Kalle Kytölä	
Date 25.9.2025	**Number of pages** 34+3 **Language** English

Abstract

In 2018, Duminil-Copin published a paper containing a new proof of the sharpness of Bernoulli bond percolation on $\mathbb{Z}^d$. Notably, the proof utilizes a statistical inequality for boolean functions by O'Donnell, Schramm, Saks, and Servedio. This thesis reviews the proof and elaborates on its steps, making it more accessible to readers without prior knowledge of percolation theory and with limited exposure to probability theory. We also present visualizations of a simulation relating to a key step of the proof to give the reader an intuition of the underlying ideas.

Keywords Percolation, Bernoulli bond percolation, boolean functions, variable influences, sharpness

Aalto University
School of Science

Aalto-universitetet, PB 11000, 00076 AALTO
www.aalto.fi
Sammandrag av kandidatarbetet

Författare Victor Keil	
Titel Bevis av skarpheten av Bernoulliperkolation på $\mathbb{Z}^d$ med hjälp av en olikhet för booleska funktioner	
Utbildningsprogram Kandidatprogrammet i teknikvetenskap	
Huvudämne Matematik och systemvetenskaper	**Huvudämnets kod** SCI3029
Ansvarslärare Prof. Kalle Kytölä	
Handledare Prof. Kalle Kytölä	
Datum 25.9.2025	**Sidantal** 34+3 **Språk** Engelska

Sammandrag

År 2018 publicerade Duminil-Copin en artikel med ett nytt bevis på skarpheten av Bernoulliperkolation på $\mathbb{Z}^d$. Särskilt för beviset är dess användning av en statistisk olikhet för booleska funktioner av O'Donnell, Schramm, Saks och Servedio. Denna uppsats granskar beviset, och utvecklar dess delsteg med målet att göra det åtkomligt för läsare utan tidigare kunskap inom perkolationsteori och med mindre erfarenhet av sannolikhetsteori. Vi för även fram visualiseringar av en simulering av ett av bevisens nyckelsteg för att förse läsaren med en intuition för de underliggande idéerna.

Nyckelord Perkolation, Bernoulli-bindnings-perkolation, booleska funktioner, variabelinfluenser, skarphet

Contents

1 Introduction

The question of lattice percolation was first posed by Broadbent and Hammersley in 1957 [6, p. vii]. The purpose of the newly created *percolation model* was to study the random propagation of *fluid* in a *medium*, where the definition of fluid and medium were intentionally kept vague. For instance, the model might capture the behavior of a solute diffusing through a solvent, electrons migrating over an atomic lattice, or a gas or liquid penetrating a porous solid. [3]

While *diffusion models*, which ascribe the stochastic mechanism to the fluid (not the medium), were already studied at the time, Broadbent and Hammersley were interested in addressing the absence of models wherein the randomness presides in the *medium* [3]. The very general notion of *abstract crystals* introduced by Broadbent and Hammersley was subsequently specialized into different classes of percolation problems by other authors, a prominent one being lattice percolation and, principal to this thesis, Bernoulli bond percolation on the d-dimensional cubic lattice.

Consider the d-dimensional integer lattice $\mathbb{L}^d = (\mathbb{Z}^d, \mathbb{E}^d)$, where $\mathbb{E}^d = \{\{x, y\} : x, y \in \mathbb{Z}^d, \|y - x\| = 1\}$ is the set of edges between adjacent vertices. The random mechanism is then governed by a sequence of independent Bernoulli distributed random variables with parameter p, one for every edge in the lattice. If the value for an edge is 1, it is retained in the lattice, otherwise, it is removed. We call the retained edges *open*, and the removed ones *closed*; we refer to the generated sub-lattice as a *configuration*. A set of connected open (closed) edges is called an open (closed) *cluster* [6, pp. 1–14], [4].

A central question in percolation theory is the existence of an *infinite open cluster*. For instance, given that the sequence of Bernoulli variables are parameterized by p (meaning that the probability of an edge being open is p, and closed $1-p$), what is the probability that there exists an open cluster containing infinitely many edges? In fact, by the independence of the edges, and the fact that the existence of such a cluster is unaffected by changing the openness of any finite subset of edges, Kolmogorov's zero-one law ensures that an infinite open cluster exists with probability zero or one [8]. Heuristically, then, one might surmise that there is some probability $p_c \in [0, 1]$ for which

- the probability that there exists an infinite open cluster is 0 when $p < p_c$,
- the probability that there exists an infinite open cluster is 1 when $p > p_c$.

This is in fact true, and p_c is known as the critical point for Bernoulli percolation [6, p. 1–14], [4].

While the probability of the existence of an infinite connected open cluster is either one or zero, the probability that the origin is part of this cluster needs a more nuanced characterization. In this thesis, we will review a proof for such a characterization, namely the sharp phase transition of Bernoulli bond percolation on the cubic lattice. The theorem was originally proven by Men'shikov (1986) [9], and Aizenman and Barsky (1987) [1], but here we study a new proof by Duminil-Copin (2018) [4] which utilizes a theorem from computer science that relates the variance of a boolean function to its computational complexity. The paper is quite terse, and assumes a fair amount of experience with probability theory, so we take this thesis as an opportunity to elaborate on the steps of the proof, and make it more accessible to readers with less exposure to probability and statistics. Furthermore, we have simulated and visualized a key step of the proof to demonstrate the intuition behind it.

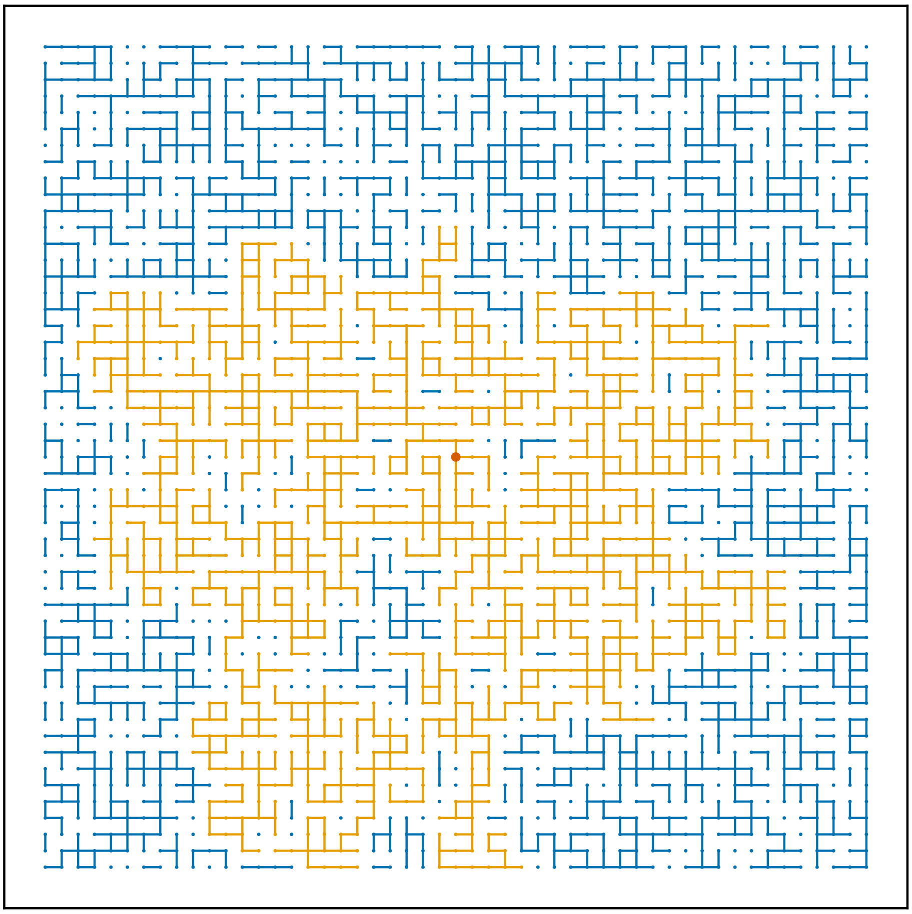

Figure 1: Example of a configuration with $p = 0.5$. The open cluster connected to the origin is shown in yellow.

2 Probability theory

2.1 Introduction

Though Bernoulli percolation is regarded as a comparatively simple percolation model, proving even rudimentary claims about it is nigh impossible, or at the very least exceedingly tedious, with only the basic tools of probability theory. Many proofs in this thesis use a variety of devices and notational tricks with which we do not expect all our readers to have prior familiarity. As such, we present here a short explanation of the necessary gadgetry.

Definition 2.1 (Probability space)**.** A probability space is a triple $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is an arbitrary set called the sample space; $\mathcal{F}$ is a σ-algebra on Ω[1] called the event space; and $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a measure on $(\Omega, \mathcal{F})$ satisfying $\mathbb{P}[\Omega] = 1$.

Elements in Ω are referred to as outcomes and elements in $\mathcal{F}$ (subsets of Ω) as events. The measure $\mathbb{P}$, sometimes called the distribution on the probability space, is simply a function, satisfying some necessary criteria, which maps events to probabilities [2]. Regarding σ-algebras, their purpose relates to the measurability of events, but their explicit mention is somewhat unwarranted for our objectives, so we will altogether drop them from the notation with the following caveat: It may be assumed that any and all subsets of the sample space that we examine are measurable. Henceforth, denote a probability space by $(\Omega, \mathbb{P})$, and assume that writing $A \subset \Omega$ implies that A is measurable.

Probability spaces enable us to work with random variables. Random variables model uncertainty in a system, but mathematically they are simply functions from the sample space to some set.

Definition 2.2 (Random variable)**.** For a probability space $(\Omega, \mathbb{P})$, a random variable $X : \Omega \to S$ on Ω is a (measurable) function from the sample space to some (measurable) space S.

Recall that an event is a subset of the sample space. Thus, to describe an event, one describes the properties satisfied by all outcomes contained in the event. Usually, this is done with the help of random variables.

Definition 2.3 (Events of random variables)**.** Given a random variable X on Ω with values in S, the event "X takes a value in E" for a given $E \subset S$ is denoted $\{X \in E\}$ and corresponds to the preimage of E under X, so $\{X \in E\} = X^{-1}(E) = \{\omega \in \Omega : X(\omega) = x\} \subset \Omega$.

In general, any predicate $p(X)$ can be used to denote an event in the same way; that is, writing $\{p(X)\}$ is equivalent to writing $\{\omega \in \Omega : p(X(\omega))\}$.

Seeing as random variables are simply functions, it seems intuitive that composing a random variable with another function should pose no issue; indeed, for a random variable $X : \Omega \to S$ and (measurable) function $f : S \to T$, $f \circ X$ is itself a random variable in the same probability space, which we denote as $f(X)$.

Finally, let us agree on some common notation for random variables. When denoting the random variable, we do not have to explicitly write out the argument of the function (X is equivalent to $X(\omega)$), because the context removes any ambiguity. In the same way, if we have defined a random variable $X : \Omega \to S$ and a function $f : S \to T$, we can write f instead of $f(X)$ if the context makes the intention clear.

2.2 Product spaces

In the real world, two different random events may occur independently from one another. Taken together, however, they can be interpreted as the single occurrence

[1] A σ-algebra on a set Ω is a set of subsets of Ω which is closed under complements and countable unions, and contains Ω itself.

of a composite event. This is precisely what is modeled by a *product probability space*. Recall that, for the rest of this thesis, we assume that all considered subsets of a sample space are measurable.

Definition 2.4 (Product probability space)**.** Given a sequence of probability spaces $(\Omega_k, \mathbb{P}_k)$ for $k \in \{1, \ldots, n\}$, we define the product probability space as $(\Omega, \mathbb{P})$, where $\Omega = \Omega_1 \times \ldots \times \Omega_n$ is the Cartesian product of the individual sample spaces and $\mathbb{P}$ is the product measure of the individual measures, which must satisfy $\mathbb{P}[A_1 \times \ldots \times A_n] = \mathbb{P}_1[A_1] \cdot \ldots \cdot \mathbb{P}_n[A_n]$ for all $A_1 \subset \Omega_1, \ldots, A_n \subset \Omega_n$.

Definition 2.5 (Independence of random variables)**.** A sequence of random variables $(X_1, \ldots, X_n)$, each X_i having codomain S_i, is called independent if, for all sequences $(E_1, \ldots, E_n)$ of subsets of $S_1, \ldots, S_n$, we have $\mathbb{P}[\bigcap_{i=1}^n \{X_i \in E_i\}] = \prod_{i=1}^n \mathbb{P}[X_i \in E_i]$.

It follows from the definition of independent random variables that a product probability space induces a sequence of independent random variables, where each variable X_i is the projection of the i:th component of the outcome.

2.3 Useful propositions

We collect here some results from probability theory which are necessary to understand the ideas presented in this thesis.

Definition 2.6 (Indicator function)**.** Given an arbitrary set U and subset $A \subset U$, define $\mathbb{1}_A : U \to \{0, 1\}$ as

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} .$$

In this thesis, A is always a set of configurations satisfying some predicate. For instance: $A = \{0 \leftrightarrow \infty\}$ is the set of configurations where the origin is part of an infinite cluster. The corresponding indicator function would then be denoted $\mathbb{1}_{0\leftrightarrow\infty}$.

In the following definitions, let $(\Omega, \mathbb{P})$ be a discrete probability space, meaning that Ω is countable and all subsets of Ω are events.

Definition 2.7 (Conditional probability)**.** For two events $A, B \subset \Omega$, with $\mathbb{P}[B] > 0$, define the conditional probability of A given B as

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$$

Remark 2.8. Although conditional probability is only defined for a conditioning event with non-zero probability, the following expression is still well defined even if, say, $\mathbb{P}[B] = 0$:

$$\mathbb{P}[A|B]\mathbb{P}[B].$$

This is because, by definition, the above expression is simply equal to $\mathbb{P}[A \cap B] = 0$.

Definition 2.9 (Conditional expectation)**.** For a random variable X with codomain S, function $f : S \to \mathbb{R}$ and an event $A \subset \Omega$ with $\mathbb{P}[A] > 0$, the conditional expectation is obtained using the measure conditioned on A. Define

$$\mathbb{E}[f(X)|A] = \sum_{x \in S} f(x)\mathbb{P}[X = x|A].$$

Proposition 2.10 (Law of total probability)**.** *Given a partition of the sample space $B = \{B_1, \ldots, B_n\}$ and an event $A \subset \Omega$, we have*

$$\mathbb{P}[A] = \sum_{i=1}^{n} \mathbb{P}[A|B_i]\mathbb{P}[B_i].$$

Proof. By definition, $\mathbb{P}[A|B_i]\mathbb{P}[B_i] = \mathbb{P}[A \cap B_i]$. Since B is a partition, we have $(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = \emptyset$ for all $i \neq j$. Countable additivity of measures then gives

$$\sum_{i=1}^{n} \mathbb{P}[A \cap B_i] = \mathbb{P}[\bigcup_{i=1}^{n} (A \cap B_i)] = \mathbb{P}[A \cap \bigcup_{i=1}^{n} B_i] = \mathbb{P}[A].$$

□

Theorem 2.11 (Bayes' theorem)**.** *Given two events A and B in the same probability space, the following relationship between conditional measures holds:*

$$\mathbb{P}[A|B]\mathbb{P}[B] = \mathbb{P}[B|A]\mathbb{P}[A]$$

Proof. Bayes' theorem follows from the definition of conditional probability and the commutativity of intersections. □

Proposition 2.12. *Given a random variable X and an event A in the same probability space, we have*

$$\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[X|A]\mathbb{P}[A]$$

Proof. Let S be the codomain of X. Starting from the definition of conditional expectation and using Bayes' theorem in the second equality, we get

$$\begin{aligned}
\mathbb{E}[X|A]\mathbb{P}[A] &= \mathbb{P}[A] \sum_{x \in S} x\mathbb{P}[X = x|A] = \mathbb{P}[A] \sum_{\omega \in \Omega} X(\omega)\mathbb{P}[\{\omega\}|A] \\
&= \mathbb{P}[A] \sum_{\omega \in \Omega} X(\omega) \underbrace{\mathbb{P}[A|\{\omega\}]}_{=\mathbb{1}_A(\omega)} \frac{\mathbb{P}[\{\omega\}]}{\mathbb{P}[A]} \\
&= \sum_{\omega \in \Omega} X(\omega)\mathbb{1}_A(\omega)\mathbb{P}[\{\omega\}] \\
&= \mathbb{E}[X\mathbb{1}_A]
\end{aligned}$$

□

Remark 2.13. This identity can be thought of as an analogue to conditional probability, meaning that

$$\mathbb{E}[X|A] = \frac{\mathbb{E}[X\mathbb{1}_A]}{\mathbb{P}[A]}$$

is to conditional expectation what

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$$

is to conditional probability.

Proposition 2.14 (Variance of Bernoulli random variable with values in $\{0,1\}$)**.** *Let $X \in \{0,1\}$ be a random variable such that $X \sim \mathrm{Ber}(p)$. Then, the variance of X with parameter p is given by*

$$\mathrm{Var}(X) = p(1-p) = \mathbb{E}[X](1 - \mathbb{E}[X]).$$

Proof. A Bernoulli distributed variable X takes value 1 with probability p and 0 with probability $1-p$. Hence, $\mathbb{E}[X] = 1 \cdot p = p$ and $X^2 = X$. Using the common formula for variance, we get

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \mathbb{E}[X](1 - \mathbb{E}[X]) = p(1-p).$$

□

2.4 Conditioning on random variables

Theorem 4.7 makes use of a shorthand notation for conditioning on random variables. It is a bit tricky, so we explain it here.

Let $(\Omega, \mathbb{P})$ be a discrete probability space. Let $X : \Omega \to S$ be a random variable and $A \subset \Omega$ an event. We wish to give meaning to the following notation:

$$\mathbb{P}[A|X].$$

Note first and foremost that $\mathbb{P}[A|X]$ is not a value, but a random variable. It can be viewed as the random variable X under the function $\gamma(x) = \mathbb{P}[A|X = x]$, where $\gamma : S \to [0,1]$. Formally, then, we have

$$\mathbb{P}[A|X] = \gamma \circ X,$$

or, using common notation, $\gamma(X)$. Importantly, when we are taking the measure of A, we are doing so under the condition of X having a constant value, even though we do not know what that value is. As an example, let $A = \{X = x\}$. Then, $\mathbb{P}[A|X]$ is a random variable with value 0 or 1 since, conditioned on X being some value, the predicate $X = x$ is either true or false for all $\omega \in \Omega$. In fact, $\mathbb{P}[X = x|X] = \mathbb{1}_{\{X=x\}}$. This, in and of itself, is not directly useful to our objective, but we will use it to define the expected value conditioned on a random variable, which we denote

$$\mathbb{E}[Y|X].$$

This conditional expectation is very useful, as it allows us to analyze the interplay between two random variables on a case by case basis, for every possible outcome of the conditioning variable (in this case X). In so doing, we might be able to discover a simplification in the expectation of Y *conditionally on the event* that X is some constant, even though we do not know for certain what that constant is. Using the conditional probability, we define conditional expectation on variables as (for random variables $X : \Omega \to S_X, Y : \Omega \to \mathbb{R}$)

$$\mathbb{E}[Y|X] = \sum_{y \in S_Y} y \mathbb{P}[Y = y|X],$$

and $\mathbb{E}[Y|X]$ is again a random variable, having value $\mathbb{E}[Y|X = x]$ on the event $\{X = x\}$. Let us give some examples. For $Z = \mathbb{E}[f(X)|X]$, assuming f is not dependent on any other random variables, its value is constant under the condition. Hence, the value of Z is equal to $f(x)$ on the event $\{X = x\}$, meaning $Z = f(X)$. For $Z = \mathbb{E}[X + Y|X]$, linearity of expectation together with our previous result gives us $Z = X + \mathbb{E}[Y|X]$. If X and Y are independent, $Z = X + \mathbb{E}[Y]$. The former case demonstrates how we can assume a constant X under the condition, whereas in the latter we see that conditioning on X preserves the non-determinism of X, while any variables independent of X (Y in this case) are reduced to a simple number.

Proposition 2.15 (Law of total expectation)**.** *Given random variables $X : \Omega \to S_X \subset \mathbb{R}$ and $Y : \Omega \to S_Y$ in the same probability space, the total expectation is given by*

$$\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[X].$$

Proof. From the definition of expected value and conditional expectation, a calculation gives

$$\begin{aligned}
\mathbb{E}[\mathbb{E}[X|Y]] &= \sum_{y \in S_Y} \mathbb{E}[X|Y = y]\mathbb{P}[Y = y] \\
&= \sum_{y \in S_Y} \sum_{x \in \Omega} x\mathbb{P}[X = x|Y = y]\mathbb{P}[Y = y] \\
&= \sum_{x \in \Omega} x\mathbb{P}[X = x] && \text{(total probability)} \\
&= \mathbb{E}[X]
\end{aligned}$$

□

3 Boolean functions

We will be studying boolean functions where the domain is taken to be configurations on the integer lattice restricted to $n \in \mathbb{N}$ edges, so there is a one to one correspondence between the edges of the configuration and the individual variables of the n-variate boolean function. A lot of the theory in this section applies to increasing boolean functions in general, but for our purposes we may limit ourselves to the case of configurations, which makes a lot of the intuitions easier to grasp, at the cost of some generality. For this reason, the word edge is used synonymously with its corresponding boolean variable.

For arbitrary sets A and E, denote by A^E the set of sequences $(a_e)_{e\in E}$ with $a_e \in A$ for all $e \in E$.

Definition 3.1. A boolean function is a function of the form $\mathbf{f} : \{0,1\}^E \to \{0,1\}$ for a set E.

In accordance with the notation in [4], denote an element in $\{0,1\}^E$ (a configuration) as $\omega = (\omega_e)_{e\in E}$, where we let E be the restriction of the cubic lattice to some finite set of edges. When considering random configurations, they are regarded as sequences of i.i.d. Bernoulli random variables with parameter p, using the product measure $\mathbb{P}_p$. For a boolean function $\mathbf{f} : \{0,1\}^E \to \{0,1\}$ define $f(p) := \mathbb{E}_p[\mathbf{f}]$.

Definition 3.2. A boolean function $\mathbf{f} : \{0,1\}^E \to \{0,1\}$ is increasing if for all $\omega, \tilde{\omega} \in \{0,1\}^E$ satisfying $\omega_e \le \tilde{\omega}_e$ for all $e \in E$, we have $\mathbf{f}(\omega) \le \mathbf{f}(\tilde{\omega})$.

Remark 3.3. For an increasing $\mathbf{f} : \{0,1\}^E \to \{0,1\}$ we have $f(p_1) \le f(p_2)$ for $p_1 \le p_2$. Lemma 3.10 in Section 3.3 proves this by showing that $f'(p)$ is non-negative.

For all $e \in E$, define the function $\mathsf{Flip}_e : \{0,1\}^E \to \{0,1\}^E$, which flips the value of ω_e, such that for all $b \in E$

$$\mathsf{Flip}_e(\omega)_b := \begin{cases} \omega_b, & \text{for } b \neq e \\ 1-\omega_b, & \text{for } b = e \end{cases}.$$

Then, for a boolean function $\mathbf{f}$, let

$$\nabla_e \mathbf{f}(\omega) := \mathbf{f}(\omega) - \mathbf{f}(\mathsf{Flip}_e(\omega)).$$

Definition 3.4. The *influence* of a boolean variable ω_e for the function $\mathbf{f}$ is defined as

$$\mathrm{Inf}_e(\mathbf{f}) := \mathbb{E}_p[|\nabla_e \mathbf{f}|].$$

It measures the extent to which the value of $\mathbf{f}$ depends on the boolean variable ω_e in ω [4, 10].

Note that the influence is dependent on the parameter p, even though it is omitted from the notation.

3.1 Percolation to the n-sphere

Definition 3.5. [4] Let the n-ball $\Lambda_n = [-n,n]^d \subset \mathbb{Z}^d$ be the set of vertices within (ℓ^∞-)distance n to the origin. Let the n-sphere $\partial\Lambda_n = \Lambda_n - \Lambda_{n-1}$ be the boundary of the n-ball.

Denote by $\{0 \leftrightarrow \partial\Lambda_n\}$ the set of configurations having an open connected path to the n-sphere. The main boolean function studied in this thesis is $\mathbb{1}_{0\leftrightarrow\partial\Lambda_n} : \{0,1\}^E \to \{0,1\}$, where E is a finite subset of edges of the lattice, and each variable of the function corresponds to an edge in the configuration. When $\mathbb{1}_{0\leftrightarrow\partial\Lambda_n}(\omega) = 1$ for some configuration ω, we say that there is percolation between the origin and the n-sphere [4].

3.2 Pivotality

Pivotality is a property of an edge in a given configuration which is closely related to influence. For a boolean function $\mathbf{f}$ and configuration ω, the edge e is called *pivotal* if flipping the value of ω_e changes the value of $\mathbf{f}(\omega)$. So e is a pivotal edge of ω if and only if $\nabla_e \mathbf{f}(\omega) \neq 0$.

Definition 3.6. Let the *pivotal set* for a boolean function $\mathbf{f} : \{0,1\}^E \to \{0,1\}$ and edge $e \in E$ be the set of configurations where e is pivotal, defined as

$$\pi_e(\mathbf{f}) = \{\omega \in \{0,1\}^E : \nabla_e \mathbf{f}(\omega) \neq 0\}.$$

Note that the statements "e is a pivotal edge in ω" and "ω is in the pivotal set of e" are equivalent. Hence, "e is pivotal in ω" is written $\omega \in \pi_e(\mathbf{f})$.

Remark 3.7. Something that follows immediately from the definition of pivotality is that, conditionally on the event $\{\omega \notin \pi_e(\mathbf{f})\}$, $\mathbf{f}(\omega)$ and ω_e are independent random variables since the value of $\mathbf{f}(\omega)$ does not change when ω_e changes.

Lemma 3.8. *The influence of an edge is equal to the probability of that edge being pivotal,*

$$\mathrm{Inf}_e(\mathbf{f}) = \mathbb{P}_p[\omega \in \pi_e(\mathbf{f})]$$

.

Proof. A straightforward calculation gives

$$\mathrm{Inf}_e(\mathbf{f}) = \mathbb{E}_p[|\nabla_e \mathbf{f}|] = \mathbb{P}_p[|\nabla_e \mathbf{f}(\omega)| = 1] = \mathbb{P}_p[\nabla_e \mathbf{f}(\omega) \neq 0].$$

□

Lemma 3.9. *An edge being pivotal in a configuration is independent of the edge's value. Thus, we have*

$$\mathbb{P}_p\left[\omega \in \pi_e(\mathbf{f}) | \omega_e\right] = \mathbb{P}_p[\omega \in \pi_e(\mathbf{f})].$$

Proof. For all $\omega \in \{0,1\}^E$, we have $\nabla_e \mathbf{f}(\omega) = -\nabla_e \mathbf{f}(\mathsf{Flip}_e(\omega))$. Hence, $\omega \in \pi_e(\mathbf{f})$ if and only if $\mathsf{Flip}_e(\omega) \in \pi_e(\mathbf{f})$. As such, $\mathbb{P}_p[\omega \in \pi_e(\mathbf{f}) | \omega_e]$ is the same for all values of ω_e, which proves the claim. □

3.3 The Margulis-Russo formula

Lemma 3.10 (Margulis-Russo formula [7, 11]). *For* $\mathbf{f} : \{0,1\}^E \to \{0,1\}$ *increasing, we have that*

$$f'(p) = \sum_{e \in E} \mathrm{Inf}_e(\mathbf{f}).$$

Recall that $f(p) = \mathbb{E}_p[\mathbf{f}]$*, and note that* $\mathrm{Inf}_e(\mathbf{f})$ *is implicitly dependent on* p.

Proof. [4] First, note that due to the independence of the boolean variables, the probability of any given configuration occurring depends only on the number of its open edges. Thus, we have

$$\mathbb{P}_p[\omega] = p^{|\omega|}(1-p)^{n-|\omega|}$$

and

$$\mathbb{E}_p[\mathbf{f}] = \mathbb{E}_p[\mathbf{f}(\omega)] = \sum_{\omega} \mathbf{f}(\omega)\mathbb{P}_p[\omega] = \sum_{\omega} \mathbf{f}(\omega)p^{|\omega|}(1-p)^{n-|\omega|},$$

where $|\omega| = \sum_{e\in E} \omega_e$ is the number of open edges in the configuration. We use the previous result to directly differentiate $f(p) = \mathbb{E}_p[\mathbf{f}]$, giving us

$$\begin{aligned}
f'(p) &= \sum_{\omega} \mathbf{f}(\omega)\frac{\mathrm{d}}{\mathrm{d}p}\left(p^{|\omega|}(1-p)^{n-|\omega|}\right) \\
&= \frac{1}{p}\sum_{\omega} \mathbf{f}(\omega)|\omega|p^{|\omega|}(1-p)^{n-|\omega|} - \frac{1}{1-p}\sum_{\omega} \mathbf{f}(\omega)(n-|\omega|)p^{|\omega|}(1-p)^{n-|\omega|} \\
&= \frac{1}{p}\mathbb{E}_p[\mathbf{f}(\omega)|\omega|] - \frac{1}{1-p}\mathbb{E}_p[\mathbf{f}(\omega)(n-|\omega|)] \\
&= \frac{1}{p(1-p)}\mathbb{E}_p[\mathbf{f}(\omega)(|\omega|-pn)] = \frac{1}{p(1-p)}\mathbb{E}_p\left[\mathbf{f}(\omega)\sum_{e\in E}(\omega_e - p)\right] \\
&= \frac{1}{p(1-p)}\sum_{e\in E}\mathbb{E}_p[\mathbf{f}(\omega)(\omega_e-p)]. \qquad (3.1)
\end{aligned}$$

Next, for each e, we split the space of possible configurations into the pivotal and non-pivotal subsets for e, $\pi_e = \pi_e(\mathbf{f})$ and π_e^c, respectively. Also, for two complementary sets in general, $1 = \mathbb{1}_A + \mathbb{1}_{A^c}$.

$$\begin{aligned}
\mathbb{E}_p[\mathbf{f}(\omega)(\omega_e-p)] &= \mathbb{E}_p[\mathbf{f}(\omega)(\omega_e-p)(\mathbb{1}_{\pi_e}+\mathbb{1}_{\pi_e^c})] \\
&= \mathbb{E}_p[\mathbf{f}(\omega)(\omega_e-p)\mathbb{1}_{\pi_e}] + \mathbb{E}_p[\mathbf{f}(\omega)(\omega_e-p)\mathbb{1}_{\pi_e^c}] \qquad (3.2)
\end{aligned}$$

We now have two cases for the value of the summand. We will start by showing that in the case where $\omega \in \pi_e^c$ the expected value is 0. By Proposition 2.12, we can write

$$\mathbb{E}_p[\mathbf{f}(\omega)(\omega_e-p)\mathbb{1}_{\pi_e^c}] = \mathbb{P}[\omega\in\pi_e^c]\mathbb{E}_p[\mathbf{f}(\omega)(\omega_e-p)|\omega\in\pi_e^c]$$

Now, notice that, conditioned on $\omega\in\pi_e^c$, $\mathbf{f}(\omega)$ and ω_e are independent (c.f. Remark 3.7). Furthermore, pivotality of an edge is independent of the edge's value (Lemma 3.9), so the expected value above becomes

$$\begin{aligned}
\mathbb{E}_p[\mathbf{f}(\omega)(\omega_e-p)|\omega\in\pi_e^c] &= \mathbb{E}_p[\mathbf{f}(\omega)|\omega\in\pi_e^c]\mathbb{E}_p[\omega_e-p|\omega\in\pi_e^c] \\
&= \mathbb{E}_p[\mathbf{f}(\omega)|\omega\in\pi_e^c]\mathbb{E}_p[\omega_e-p] \\
&= \mathbb{E}_p[\mathbf{f}(\omega)|\omega\in\pi_e^c]\left(\mathbb{E}_p[\omega_e]-p\right) \\
&= 0,
\end{aligned}$$

and thus the last term in (3.2) is 0. From that it follows that

$$\mathbb{E}_p[\mathbf{f}(\omega)(\omega_e-p)] = \mathbb{E}_p[\mathbf{f}(\omega)(\omega_e-p)\mathbb{1}_{\pi_e}]. \qquad (3.3)$$

Again, by Proposition 2.12,

$$\mathbb{E}_p[\mathbf{f}(\omega)(\omega_e - p)\mathbb{1}_{\pi_e}] = \mathbb{P}_p[\omega \in \pi_e]\mathbb{E}_p[\mathbf{f}(\omega)(\omega_e - p)|\omega \in \pi_e]. \tag{3.4}$$

We now have the condition that e is pivotal, so a change in the value of e must change the value of $\mathbf{f}(\omega)$. But since $\mathbf{f}$ is increasing, it must be that when $\omega_e = 0$, $\mathbf{f}(\omega) = 0$, and when $\omega_e = 1$, $\mathbf{f}(\omega) = 1$. It follows that

$$\begin{aligned}\mathbb{E}_p[\mathbf{f}(\omega)(\omega_e - p)|\omega \in \pi_e] &= 1 \cdot (1-p)\mathbb{P}_p[\omega_e = 1|\omega \in \pi_e] = (1-p)\mathbb{P}_p[\omega_e = 1] \\ &= (1-p)p,\end{aligned}$$

where the second to last equality again follows from independence of pivotality and value. Combining this result with (3.2), (3.3), and (3.4) yields

$$\mathbb{E}_p[\mathbf{f}(\omega)(\omega_e - p)] = \mathbb{P}_p[\omega \in \pi_e]p(1-p) = p(1-p)\mathrm{Inf}_e(\mathbf{f}).$$

Inserting this into (3.1) finally completes the proof. $\square$

3.4 Sharpness condition for boolean functions

Definition 3.11. [4] Let (E_n) be an increasing sequence of finite sets, $(\mathbf{f}_n)$ a sequence of boolean functions where $\mathbf{f}_n : \{0,1\}^{E_n} \to \{0,1\}$, and (p_n) a sequence of numbers in $(0,1)$. Then, we say that $(\mathbf{f}_n)$ undergoes a sharp threshold at (p_n) if there exists a sequence of positive real numbers (δ_n) tending to 0 such that $f_n(p_n - \delta_n) \to 0$ and $f_n(p_n + \delta_n) \to 1$ as $n \to \infty$.

A useful type of inequality for proving that a sequence of boolean functions undergoes a sharp threshold is

$$\sum_{e \in E} \mathrm{Inf}_e(\mathbf{f}) \geq C\ \mathrm{Var}_p(\mathbf{f}), \tag{3.5}$$

for some large constant C. That is to say, if we can show that the inequality holds for a sequence of $(\mathbf{f}_n)$ with (C_n) tending to infinity, it follows that $(\mathbf{f}_n)$ undergoes a sharp threshold.

Lemma 3.12. *Let $\mathbf{f} : \{0,1\}^E \to \{0,1\}$ be a boolean function on a sequence of i.i.d. Bernoulli variables with parameter p. Assume that*

$$\sum_{e \in E} \mathrm{Inf}_e(\mathbf{f}) \geq C\ \mathrm{Var}_p(\mathbf{f})$$

for some constant C. Then, there exists $p \in (0,1)$ such that for all $\delta > 0$, we have

$$f(p-\delta) \leq e^{-\delta C} \quad \textit{and} \quad f(p+\delta) \geq 1 - e^{-\delta C}. \tag{3.6}$$

Proof. By our assumption, and applying Lemma 3.10, we have

$$\sum_{e \in E} \mathrm{Inf}_e(\mathbf{f}) = f'(p) \geq C\ \mathrm{Var}_p(\mathbf{f}).$$

Since $\mathbf{f}$ takes values in $\{0, 1\}$, it is a Bernoulli random variable and thus

$$\mathrm{Var}_p(\mathbf{f}) = f(p)(1 - f(p)),$$

recalling that $f(p) = \mathbb{E}_p(\mathbf{f})$. This allows us to write the previous inequality as

$$f'(p) \geq Cf(p)(1 - f(p))$$

which is equivalent to

$$\frac{f'(p)}{f(p)(1 - f(p))} \geq C. \tag{3.7}$$

Observe that

$$\frac{f'(p)}{f(p)(1 - f(p))} = \frac{\mathrm{d}}{\mathrm{d}p}\left(\ln \frac{f(p)}{1 - f(p)}\right).$$

Integrating the left hand side of (3.7) from p_0 to p_1, we get

$$\int_{p_0}^{p_1} \frac{\mathrm{d}}{\mathrm{d}p}\left(\ln \frac{f(p)}{1 - f(p)}\right) \mathrm{d}p = \ln \frac{f(p_1)}{1 - f(p_1)} - \ln \frac{f(p_0)}{1 - f(p_0)} = \ln \frac{f(p_1)(1 - f(p_0))}{(1 - f(p_1))f(p_0)},$$

which together with the right hand side gives

$$\ln \frac{f(p_1)(1 - f(p_0))}{(1 - f(p_1))f(p_0)} \geq (p_1 - p_0)C.$$

By exponentiating and taking the reciprocal we have

$$\frac{(1 - f(p_1))f(p_0)}{f(p_1)(1 - f(p_0))} \leq e^{-(p_1 - p_0)C}. \tag{3.8}$$

Now, choose p such that $f(p) = \frac{1}{2}$ and let $p_0 = p - \delta$ and $p_1 = p$. Then, for any $\delta > 0$,

$$\frac{(1 - f(p))f(p - \delta)}{f(p)(1 - f(p - \delta))} = \frac{\frac{1}{2}f(p - \delta)}{\frac{1}{2}\underbrace{(1 - f(p - \delta))}_{\leq 1}} \leq e^{-\delta C},$$

which implies that

$$f(p - \delta) \leq e^{-\delta C}.$$

By setting $p_0 = p$ and $p_1 = p + \delta$, we get

$$\frac{(1 - f(p + \delta))f(p)}{f(p + \delta)(1 - f(p))} = \frac{(1 - f(p + \delta))\frac{1}{2}}{\underbrace{f(p + \delta)\tfrac{1}{2}}_{\leq 1}} \leq e^{-\delta C},$$

implying that

$$f(p + \delta) \geq 1 - e^{-\delta C}.$$

$\square$

In words, this means that for some $p \in [0, 1]$ such that $f(p) = \frac{1}{2}$, $f(p)$ decays exponentially towards 0 when the value of p is decreased, and exponentially towards 1 when the value of p is increased. If we let $\epsilon > e^{-\delta C}$, then

$$f(p - \delta) \leq \epsilon \quad \text{and} \quad f(p + \delta) \geq 1 - \epsilon,$$

where ϵ describes the maximum distance of $f(p)$ from 0 or 1 when p is outside the 2δ interval around p. We can see, from the relationship between ϵ, δ, and C, that for a given ϵ, increasing C allows δ to be smaller, narrowing the threshold interval. Figure 2 shows an area-graph of the values satisfying inequality (3.6).

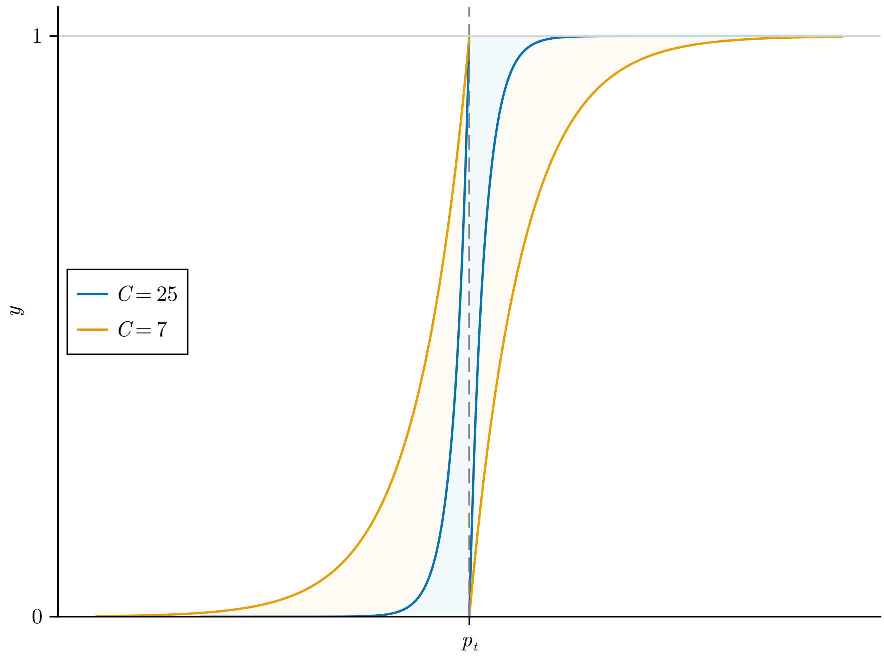

Figure 2: The bounds of inequality (3.6). For $\delta = |p - p_t|$, the shaded area satisfies $y \leq e^{-\delta C}$ when $p < p_t$ and $y \geq 1 - e^{-\delta C}$ when $p > p_t$. Note the different values of C for the two colored areas.

As an example of how the inequality (3.5) implies that a function exhibits a kind of sharpness, take f_C to be

$$f_C(x) = \frac{1}{1 + e^{-Cx}}.$$

A straightforward calculation gives that,

$$f'_C(x) = C f_C(x)(1 - f_C(x))$$

so f_C is actually a solution to the equation formed by the inequality (3.7) taken as a strict equality. Figure 3 shows a graph of f_C and its relation to ϵ and δ, as well as how it gets sharper for increasing C. When proving the sharpness of the percolation phase transition, we will use a similar inequality to (3.5), which (as will be shown) also gives rise to a differential inequality that implies sharpness.

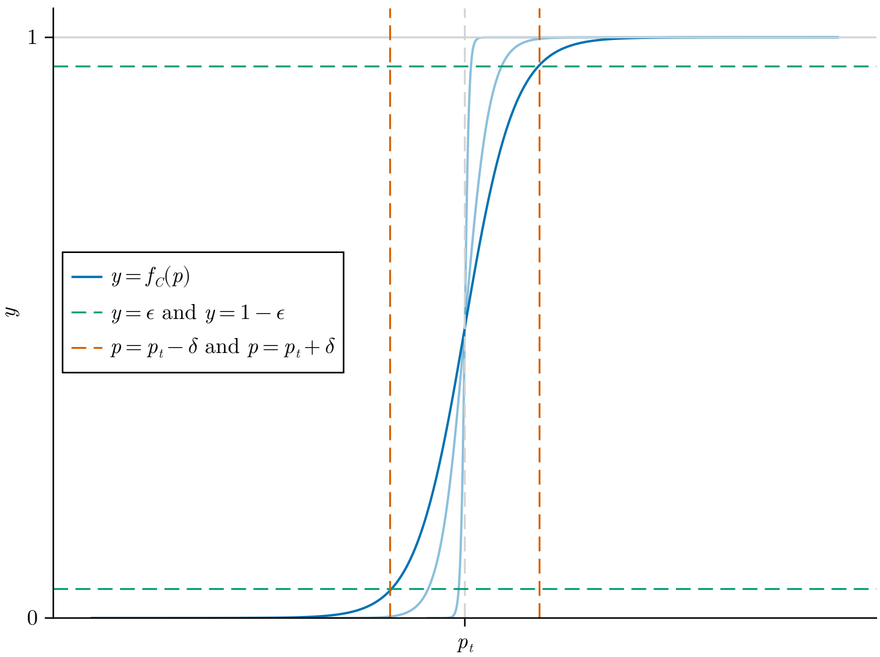

Figure 3: Visualization of how ϵ and δ relate to the sharpness of a function. Here $f_C(p) = \frac{1}{1+\exp(-Cp)}$ is a sigmoid function. The light blue curves are the result of increasing C, with the sharper curves having a higher value of C.

4 Studying percolation through algorithms

Viewing percolation as the value of a boolean function opens up the possibility of exploiting results from computer science to characterize it. One such result is the OSSS-inequality (Theorem 4.7), which relates the variance of a boolean function to its influences and something called the *revealment* of an *algorithm* for a boolean function.

Before we proceed, let us briefly reiterate some key concepts. First, recall that a configuration $\omega \in \{0,1\}^E$ is a sequence of boolean variables $(\omega_e)_{e \in E}$ where each variable corresponds to an edge in the set E, which is some finite subset of edges of the integer lattice. We call a connected set of open (closed) edges an open (closed) cluster. If two sets of vertices are connected by an open cluster, we say that the configuration percolates between these sets. In our case, the sets E will be the edge sets of n-balls (Definition 3.5) centered around the origin, and our goal is characterizing the probability of percolation between the origin and the outermost vertices of the n-ball.

4.1 Algorithms

When looking at a graph of a random configuration, it is clear that not every edge needs to be considered in order to determine percolation from the origin to some target set of points. For instance, it is enough to simply attempt to follow the connected open edges from the origin until either the target set is reached or the whole open cluster is discovered, disregarding all other edges (technically, the closed edges adjacent to the open cluster must also be considered in order to determine the boundary of the open cluster). This is a key property of percolation and is closely related to the monotonicity of the boolean functions we consider.

The process of evaluating edges connected to the origin in order to determine the value of $\mathbb{1}_{0\leftrightarrow\partial\Lambda_n}$ (cf. Definition 3.5) is an example of an *evaluation strategy* or *algorithm*. An algorithm sequentially queries or, *reveals*, the edges of a configuration with the goal of determining the value of a boolean function on the configuration. Which edge to reveal next is determined by the previously revealed edges and the *decision rules* of the algorithm.

Let us establish some notation to work with algorithms; we will use the same definitions as in [4]. First, write $[n]$ as a shorthand for $\{1, \ldots, n\}$. For an m-tuple $x = (x_1, \ldots, x_m)$, denote its restriction to the $t \leq m$ first elements by $x_{[t]} = (x_1, \ldots, x_t)$. Given a configuration $\omega \in \{0, 1\}^E$ and $x \in E^t$, denote $\omega_{x_{[t]}} = (\omega_{x_1}, \ldots, \omega_{x_t})$.

Definition 4.1. An *algorithm* $\mathbf{T} = (\varphi_t)_{t=1}^{|E|}$ takes in a configuration $\omega \in \{0, 1\}^E$ and gives out an ordered sequence of edges, constructed inductively as follows:

$$\mathbf{e}_t = \varphi_t(\mathbf{e}_{[t-1]}, \omega_{\mathbf{e}_{[t-1]}}) \in E \setminus \{\mathbf{e}_1, \ldots, \mathbf{e}_{t-1}\},$$

where φ_t is a function describing the decision rule at time step t.

This is to say, at time step t, $t-1$ edges in ω (chosen by the algorithm in previous steps) are known to the algorithm. Based on that knowledge, the algorithm chooses which edge to reveal next, which will be $\mathbf{e}_t$. Note that $\mathbf{e}_1$ is given as the result of an empty input to φ_1.

Definition 4.2. The *stopping time* $\tau_{\mathbf{T},\mathbf{f}}(\omega)$ of the algorithm $\mathbf{T}$ given a function $\mathbf{f}$ and configuration ω is defined as

$$\tau(\omega) = \tau_{\mathbf{f},\mathbf{T}}(\omega) = \min\{t \geq 1 : \forall x \in \{0, 1\}^E : x_{\mathbf{e}_{[t]}} = \omega_{\mathbf{e}_{[t]}} \implies \mathbf{f}(x) = \mathbf{f}(\omega)\}.$$

In words, what this means is that τ is the minimum number of edges that the algorithm needs to reveal in order to determine the value of $\mathbf{f}(\omega)$.

Definition 4.3. The *revealment* of $\mathbf{f}$, $\mathbf{T}$, and edge e is defined as

$$\delta_e(\mathbf{T}) = \delta_e(\mathbf{f}, \mathbf{T}) = \mathbb{P}_p[\exists t \leq \tau_{\mathbf{f},\mathbf{T}} : \mathbf{e}_t = e].$$

Thus, for an algorithm $\mathbf{T}$ and function $\mathbf{f}$, the revealment of e is the probability that e is queried by the algorithm when determining the value of $\mathbf{f}(\omega)$.

Lemma 4.4. *The revealment can also be written*

$$\delta_e(\mathbf{T}) = \sum_{t=1}^{|E|} \mathbb{P}_p\left[t \leq \tau, \mathbf{e}_t = e\right].$$

Proof. Let $\mathbf{e}$ be the sequence of edges constructed inductively by the algorithm $\mathbf{T}$. Then,

$$\delta_e(\mathbf{T}) = \mathbb{P}_p\left[\exists t \leq \tau : \mathbf{e}_t = e\right] = \mathbb{P}_p\left[\bigcup_{t=1}^{|E|}\{t \leq \tau, \mathbf{e}_t = e\}\right].$$

By the definition of an algorithm, every edge $e \in E$ is present at most once in $\mathbf{e}$. It follows that the events $\{\mathbf{e}_t = e\}$ are mutually exclusive for all t, so the probability of the union can be written as a sum. Hence,

$$\delta_e(\mathbf{T}) = \sum_{t=1}^{|E|}\mathbb{P}_p\left[t \leq \tau, \mathbf{e}_t = e\right]$$

□

Examples of algorithms

To acquaint ourselves with algorithms, let us take a look at two such instances. Consider the boolean function $\mathbf{f} = \mathbb{1}_{0 \leftrightarrow \partial\Lambda_m}$: Checking for percolation from the origin to the boundary of Λ_m. We wish to evaluate this function, starting with the trivial algorithm. First, let $E_m \subset \mathbb{E}^d$ be a subset of edges with both endpoints in Λ_m and fix an ordering of E such that $\mathbf{e}_1$ is the first edge, $\mathbf{e}_2$ the second, and so on. Let the sequence of edges produced by the algorithm be $\mathbf{e} = (\mathbf{e}_t)_{t \in [|E|]}$. This algorithm queries every edge, regardless of configuration.

An example of a slightly smarter algorithm is one which starts from the origin and finds all edges in the connected open cluster. Every time step reveals a new edge connected to the open cluster. The algorithm halts if either (1) it has found an open connected path to the boundary, or (2) the whole open cluster at the origin has been found.

The k-sphere algorithms

At this point, one might ask what we gain by formalizing these algorithms. In the next section, we will present the OSSS-inequality, which relates the variance of a boolean function to its influences and revealments for a given algorithm. Thus, if we can find an algorithm that provides a useful *uniform* upper bound for the revealments, we might find a similar inequality to (3.5), which would bring us closer to proving sharpness. The *origin-out* algorithm described above gives a good upper bound for the edges far from the origin, but the ones near the origin are still revealed with high probability. Duminil-Copin instead proposes what we will refer to as the *k-sphere* algorithms, which is a family of algorithms $\mathbf{T}_k$ that discover all open clusters in Λ_n intersecting $\partial\Lambda_k$. Then, for the algorithm to reveal an edge $e = \{x, y\} \in E_n$, where E_n is the set of edges with both endpoints in Λ_n, at least one of the endpoints of e must necessarily be connected to $\partial\Lambda_k$. Hence,

$$\delta_e(\mathbf{T}_k) \leq \mathbb{P}_p[x \leftrightarrow \partial\Lambda_k] + \mathbb{P}_p[y \leftrightarrow \partial\Lambda_k],$$

which is simply stating that the revealment is bounded from above by the probability of percolation between the k-sphere and at least one of the endpoints of e.

We would like to point out that the revealment-upper-bound for a single algorithm $\mathbf{T}_k$ is *still* not sufficient, but we will see in the proof of Theorem 5.2 that summing over all $k \in [n]$ algorithms *is* a good enough upper bound to prove sharpness. The sum-upper-bound together with the Margulis-Russo differentiation formula will allow us to use real analysis to prove the sharpness for Bernoulli percolation. [4]

Figures 4–5 show simulations of the revealments for the origin-out and k-sphere algorithms at the critical point. Compared to the origin-out algorithm, the simulated revealments for the k-sphere algorithms are indeed distributed more uniformly across edges in E_n. However, the upper bound seems, rather strikingly, far *worse* than that of the origin-out algorithm. While it may seem counter-productive at first, recall that we are not using a single instance of $\mathbf{T}_k$ to obtain an upper bound, but rather the sum over all $k \in [n]$.

Definition 4.5 (k-sphere algorithm, $\mathbf{T}_k$)**.** The k-sphere algorithm is defined formally as follows. Fix an ordering of E_n. Define recursively two growing sequences $\partial\Lambda_k = V_0 \subset V_1 \subset \cdots \subset V$ and $\emptyset = F_0 \subset F_1 \subset \cdots \subset E_n$, and observe the two cases:

- If there exists an edge $e = \{x, y\} \in E_n \setminus F_t$ such that $x \in V_t$ or $y \in V_t$, set $\mathbf{e}_{t+1} = e$ (picking the smallest one for the ordering if more than one exist), $F_{t+1} = F_t \cup \{e\}$, and

$$V_{t+1} = \begin{cases} V_t \cup \{x, y\}, & \text{if } \omega_e = 1 \\ V_t & \text{if } \omega_e = 0. \end{cases}$$

- If such an edge does not exist, take $e \in E_n \setminus F_t$ to be the smallest edge (for the ordering) of the undiscovered edges and set $\mathbf{e}_{t+1} = e$, $F_{t+1} = F_t \cup \{e\}$, and $V_{t+1} = V_t$.

The set F_t should be interpreted as the set of discovered edges by time step t, and V_t the set of vertices connected by an open path to $\partial\Lambda_k$, discovered by time step t. [4]

Remark 4.6. The meticulous reader might notice that our definition of the k-sphere algorithms differs very slightly from the one posed in [4, p. 14]. There seems to have been a minor technical oversight in the original definition which contradicts the logic used to derive the revealment upper bound. More on this can be found in the appendix, Section A.

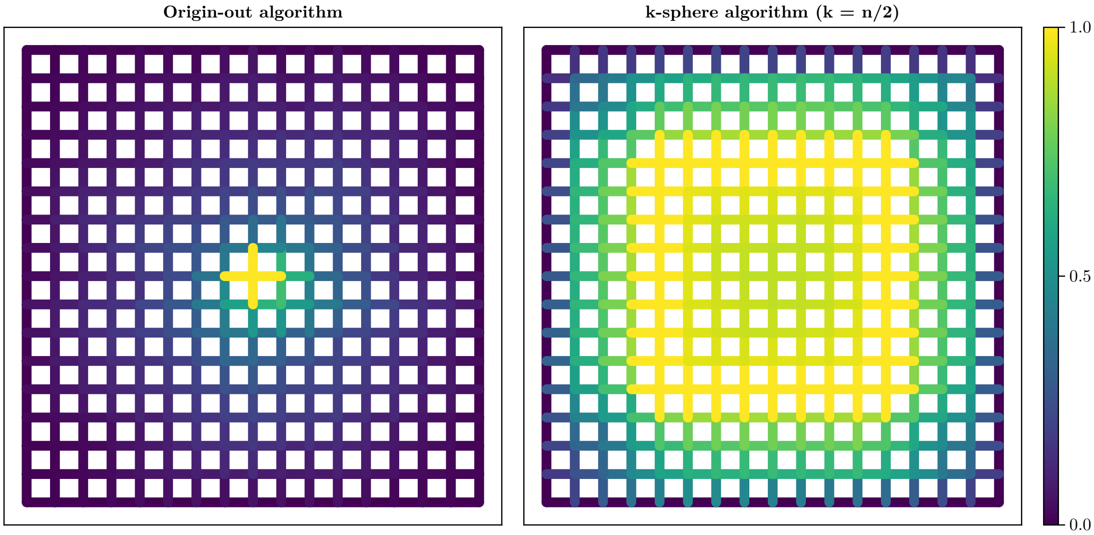

Figure 4: Revealment of edges for the origin-out (left) and the k-sphere (right) algorithms determining percolation to the n-sphere ($\mathbb{1}_{0\leftrightarrow\partial\Lambda_n}$). In both cases, $p = p_c$, and for k-sphere, $k = \frac{n}{2}$.

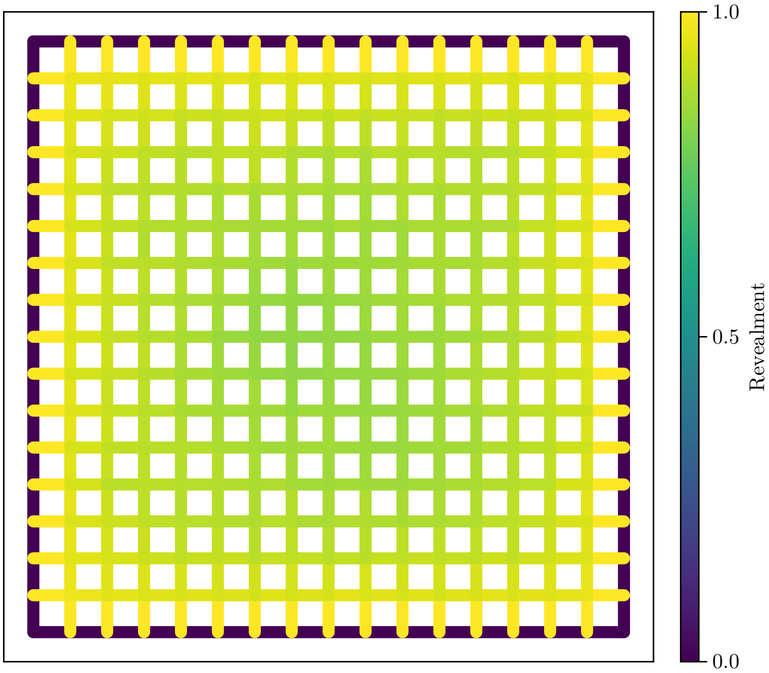

Figure 5: Revealment of edges for the k-sphere algorithm determining percolation to the n-sphere ($\mathbb{1}_{0\leftrightarrow\partial\Lambda_n}$). $p = p_c$ and $k = n$.

4.2 The O'Donnell-Schramm-Saks-Servedio inequality

We shall now present a theorem proven by O'Donnell, Saks, Schramm, and Servedio in [10]. The purpose of the paper was to attain a lower bound to the maximum influence of a boolean function, given by the function's evaluation complexity. Here we show the proof as presented by Duminil-Copin [4], which he adapted to account for a slightly different definition of influence and specialized to use the Euclidean metric on the values of boolean functions (where the original proof was for a general metric space).

Theorem 4.7 (OSSS inequality [10, 4])**.** *Consider $p \in [0,1]$ and a finite set E. Fix a boolean function $\mathbf{f} : \{0,1\}^E \to \{0,1\}$ and an algorithm $\mathbf{T}$. We have*

$$\mathrm{Var}_p(\mathbf{f}) \leq p(1-p) \sum_{e \in E} \delta_e(\mathbf{T}) \mathrm{Inf}_e(\mathbf{f})$$

Proof. Let ω and $\tilde{\omega}$ be two independent sequences of i.i.d. Bernoulli distributed random variables with parameter p. Let $\mathbf{T} = (\psi_t)$ be an algorithm with $\mathbf{e}$ constructed as

$$\mathbf{e}_{t+1} = \psi_t(\mathbf{e}_{[t]}, \omega_{\mathbf{e}_{[t]}}).$$

Note that $\mathbf{e}$ is independent of $\tilde{\omega}$. Also, let

$$\tau(\omega) = \min\{t \geq 1 : \forall x \in \{0,1\}^E, x_{\mathbf{e}_{[t]}} = \omega_{\mathbf{e}_{[t]}} \implies \mathbf{f}(x) = \mathbf{f}(y)\}$$

be the stopping time of algorithm $\mathbf{T}$ determining the value of $\mathbf{f}$ for ω (and hence τ also independent of $\tilde{\omega}$). Then, letting $m = |E|$, construct the hybrid input

$$\omega^t = (\tilde{\omega}_{\mathbf{e}_1}, \ldots, \tilde{\omega}_{\mathbf{e}_t}, \omega_{\mathbf{e}_{t+1}}, \ldots, \omega_{\mathbf{e}_\tau}, \tilde{\omega}_{\mathbf{e}_{\tau+1}}, \ldots, \tilde{\omega}_{\mathbf{e}_m}),$$

noticing that $\omega^t = \tilde{\omega}$ when $t \geq \tau$, and $\mathbf{f}(\omega^0) = \mathbf{f}(\omega)$.

Since $\mathbf{f}$ is a Bernoulli distributed random variable with values in $\{0,1\}$, we have

$$\mathrm{Var}_p(\mathbf{f}) = f(p)(1 - f(p)),$$

by Proposition 2.14. We also have

$$\mathbb{E}_p[|\mathbf{f} - f(p)|] = f(p) \cdot |1 - f(p)| + (1 - f(p)) \cdot |0 - f(p)| = 2f(p)(1 - f(p)),$$

so

$$2\mathrm{Var}_p(\mathbf{f}) = \mathbb{E}_p[|\mathbf{f} - f(p)|]. \tag{4.1}$$

From here on, use $\mathbb{P}$ as the product measure of $\omega, \tilde{\omega}$ (and likewise with $\mathbb{E}$). Now, write $\mathbf{f}(\omega) = \mathbb{E}[\mathbf{f}(\omega)|\omega] = \mathbb{E}[\mathbf{f}(\omega^0)|\omega]$, and using the fact that ω and $\tilde{\omega}$ are independent and $\tilde{\omega} = \omega^m$, write $f(p) = \mathbb{E}[\mathbf{f}(\omega^m)] = \mathbb{E}[\mathbf{f}(\omega^m)|\omega]$. We rewrite the above as

$$\begin{aligned}
\mathbb{E}\left[|\mathbf{f} - f(p)|\right] &= \mathbb{E}\left[\left|\mathbb{E}[\mathbf{f}(\omega^0)|\omega] - \mathbb{E}[\mathbf{f}(\omega^m)|\omega]\right|\right] \\
&= \mathbb{E}\left[\left|\mathbb{E}[\mathbf{f}(\omega^0) - \mathbf{f}(\omega^m)|\omega]\right|\right] && \text{(linearity of expectation)} \\
&\leq \mathbb{E}\left[\mathbb{E}\left[\left|\mathbf{f}(\omega^0) - \mathbf{f}(\omega^m)\right| \middle| \omega\right]\right] && (|\mathbb{E}[X]| \leq \mathbb{E}[|X|]) \\
&= \mathbb{E}\left[\left|\mathbf{f}(\omega^0) - \mathbf{f}(\omega^m)\right|\right]. && \text{(total expectation)}
\end{aligned}$$

We expand into a telescopic sum

$$\mathbf{f}(\omega^m) - \mathbf{f}(\omega^0) = \sum_{t=1}^{m} \mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1}),$$

which combined with the triangle inequality yields

$$\mathbb{E}\left[\left|\mathbf{f}(\omega^m) - \mathbf{f}(\omega^0)\right|\right] \leq \mathbb{E}\left[\left|\sum_{t=1}^{m} \mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1})\right|\right] \leq \sum_{t=1}^{m} \mathbb{E}\left[\left|\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1})\right|\right],$$

so now we have

$$\mathbb{E}[|\mathbf{f} - f(p)|] \leq \sum_{t=1}^{m} \mathbb{E}\left[\left|\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1})\right|\right]. \tag{4.2}$$

Since we have $\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1}) = 0$ when $t > \tau$, the summand above can be written

$$\mathbb{E}\left[\left|\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1})\right|\right] = \mathbb{E}\left[\left|\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1})\right| \mathbb{1}_{t\leq\tau}\right] = \mathbb{E}\left[\mathbb{E}\left[\left|\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1})\right| \mathbb{1}_{t\leq\tau} \middle| \omega_{\mathbf{e}_{[t-1]}}\right]\right],$$

where the last equality follows from the law of total expectation (Proposition 2.15). Now, notice that the event $\{t \leq \tau\}$ is entirely determined by $\omega_{\mathbf{e}_{[t-1]}}$, so in the inner expectation, which is conditional on $\omega_{\mathbf{e}_{[t-1]}}$, $\mathbb{1}_{t\leq\tau}$ is constant. Hence, we can write

$$\mathbb{E}\left[\left|\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1})\right|\right] = \mathbb{E}\left[\mathbb{E}\left[\left|\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1})\right| \middle| \omega_{\mathbf{e}_{[t-1]}}\right] \mathbb{1}_{t\leq\tau}\right]. \tag{4.3}$$

Conditional on $\omega_{\mathbf{e}_{[t-1]}}$, the sequences ω^t and ω^{t-1} are i.i.d. sequences of Bernoulli variables whose values are independent of the values of $\omega_{\mathbf{e}_{[t-1]}}$ since they only involve $\tilde{\omega}$ on edges $\mathbf{e}_{[t-1]}$. They are also necessarily equal on every edge except for $\mathbf{e}_t$, where they might differ since $\omega^t_{\mathbf{e}_t} = \tilde{\omega}_{\mathbf{e}_t}$ and $\omega^{t-1}_{\mathbf{e}_t} = \omega_{\mathbf{e}_t}$. Conditionally on $\omega_{\mathbf{e}_{[t-1]}}$ and $t \leq \tau$, in the case where $\omega_{\mathbf{e}_t} = \tilde{\omega}_{\mathbf{e}_t}$ we have $\omega^{t-1} = \omega^t$ and so $\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1}) = 0$; and in the case where $\omega_{\mathbf{e}_t} \neq \tilde{\omega}_{\mathbf{e}_t}$, the difference between ω^{t-1} and ω^t is that the edge $\mathbf{e}_t$ is flipped; thus we can write $|\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1})| = |\nabla_{\mathbf{e}_t}\mathbf{f}(\omega^t)|\mathbb{1}_{\omega_{\mathbf{e}_t}\neq\tilde{\omega}_{\mathbf{e}_t}}$. In light of this, the inner expectation of (4.3) becomes

$$\mathbb{E}\left[\left|\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1})\right| \middle| \omega_{\mathbf{e}_{[t-1]}}\right] = \mathbb{E}\left[\left|\nabla_{\mathbf{e}_t}(\mathbf{f}(\omega^t))\right| \mathbb{1}_{\omega_{\mathbf{e}_t}\neq\tilde{\omega}_{\mathbf{e}_t}} \middle| \omega_{\mathbf{e}_{[t-1]}}\right],$$

which by Proposition 2.12 can be written

$$\mathbb{P}[\omega_{\mathbf{e}_t} \neq \tilde{\omega}_{\mathbf{e}_t} | \omega_{\mathbf{e}_{[t-1]}}]\mathbb{E}\left[\left|\nabla_{\mathbf{e}_t}(\mathbf{f}(\omega^t))\right| \middle| \omega_{\mathbf{e}_{[t-1]}}, \omega_{\mathbf{e}_t} \neq \tilde{\omega}_{\mathbf{e}_t}\right].$$

Since $\omega_{\mathbf{e}_t}$ and $\tilde{\omega}_{\mathbf{e}_t}$ are independent and their values independent of $\omega_{\mathbf{e}_{[t-1]}}$, the above probability is $p(1-p)+(1-p)p = 2p(1-p)$, recalling that p is probability of an edge being open. In the expectation, the value of $\nabla_{\mathbf{e}_t}(\mathbf{f}(\omega^t))$ is independent of the event $\omega_{\mathbf{e}_t} \neq \tilde{\omega}_{\mathbf{e}_t}$ by Lemma 3.9. Also, due to the fact that $\omega_{\mathbf{e}_{[t-1]}}$ fully determines $\mathbf{e}_t$ and is independent of the value of ω^t, the expectation becomes $\mathbb{E}\left[|\nabla_{\mathbf{e}_t}(\mathbf{f}(\omega^t))| \middle| \omega_{\mathbf{e}_{[t-1]}}\right] = \mathrm{Inf}_{\mathbf{e}_t}(\mathbf{f})$, noting that $\mathbf{e}_t$ is a random variable. Consequently,

$$\mathbb{E}\left[\left|\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1})\right| \middle| \omega_{\mathbf{e}_{[t-1]}}\right] = 2p(1-p)\mathrm{Inf}_{\mathbf{e}_t}(\mathbf{f}),$$

which when inserted into (4.3) gives

$$\mathbb{E}\left[\left|\mathbf{f}(\omega^t) - \mathbf{f}(\omega^{t-1})\right|\right] = \mathbb{E}[2p(1-p)\mathrm{Inf}_{\mathbf{e}_t}(\mathbf{f})\mathbb{1}_{t\leq\tau}].$$

We insert this into (4.2) which yields

$$\mathbb{E}\left[|\mathbf{f} - f(p)|\right] \leq \sum_{t=1}^{m} \mathbb{E}\left[2p(1-p)\mathrm{Inf}_{\mathbf{e}_t}(\mathbf{f})\mathbb{1}_{t\leq\tau}\right]. \tag{4.4}$$

Next, notice that at a given time step t, exactly one new edge is revealed. Therefore, the events $\{\mathbf{e}_t = e\}$ over $e \in E$ form a partition of the sample space, so $\sum_{e\in E} \mathbb{1}_{\mathbf{e}_t=e} = 1$ and $\sum_{e\in E} \mathbb{1}_{\mathbf{e}_t=e,t\leq\tau} = \mathbb{1}_{t\leq\tau}$. Substituting into (4.4):

$$\mathbb{E}\left[|\mathbf{f} - f(p)|\right] \leq \sum_{t=1}^{m} \mathbb{E}\left[2p(1-p)\mathrm{Inf}_{\mathbf{e}_t}(\mathbf{f}) \sum_{e\in E} \mathbb{1}_{t\leq\tau,e=\mathbf{e}_t}\right],$$

then using linearity of expectation and the presence of $\mathbb{1}_{e=\mathbf{e}_t}$, we bring $\mathrm{Inf}_{\mathbf{e}_t}(\mathbf{f}) = \mathrm{Inf}_e(\mathbf{f})$ out of the expectation, resulting in

$$\begin{aligned}\mathbb{E}\left[|\mathbf{f} - f(p)|\right] &\leq 2p(1-p) \sum_{e\in E} \sum_{t=1}^{m} \mathrm{Inf}_e(\mathbf{f})\mathbb{E}[\mathbb{1}_{t\leq\tau,e=\mathbf{e}_t}] \\ &= 2p(1-p) \sum_{e\in E} \mathrm{Inf}_e(\mathbf{f}) \sum_{t=1}^{m} \mathbb{P}[t \leq \tau, e = \mathbf{e}_t].\end{aligned}$$

Finally, recalling that $\sum_{t=1}^{m} \mathbb{P}[t \leq \tau, e = \mathbf{e}_t] = \delta_e$ (Lemma 4.4) and applying (4.1) completes the proof. □

5 Sharpness of the phase transition for Bernoulli percolation on $\mathbb{Z}^d$

Let us now return our attention to this thesis' central point of inquiry, namely the sharp phase transition of Bernoulli percolation on $\mathbb{Z}^d$. The sharpness of Bernoulli percolation phase transition was originally proven by Men'shikov (1986) [9], and Aizenman and Barsky (1987) [1]. In 2015, Duminil-Copin published a new proof of the theorem in collaboration with Tassion [5], and in 2018 he published a paper containing yet a new proof, the one we are about to examine.

Definition 5.1. [9] Define the critical point of Bernoulli percolation as

$$p_c := \sup\{p : \mathbb{P}_p[0 \leftrightarrow \infty] = 0\},$$

where $\mathbb{P}[0 \leftrightarrow \infty] = \lim_{n\to\infty} \mathbb{P}[0 \leftrightarrow \partial\Lambda_n]$ denotes the probability that the origin is part of an infinite open cluster.

In the proof [4], Duminil-Copin shows that there exists a certain point $\tilde{p}_c \in [0, 1]$, where $\mathbb{P}_p[0 \leftrightarrow \partial\Lambda_n]$ behaves differently for $p < \tilde{p}_c$ and $p > \tilde{p}_c$. It then follows from the proven behavior of $\mathbb{P}_p[0 \leftrightarrow \partial\Lambda_n]$ and the definition of p_c that $\tilde{p}_c = p_c$.

Theorem 5.2. [9, 1, 4] *Consider Bernoulli percolation on $\mathbb{Z}^d$.*

1. *For $p < p_c$, there exists $c_p > 0$ such that for all $n \geq 1, \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_n] \leq \exp(-c_p n)$.*
2. *There exists $c > 0$ such that for $p > p_c$, $\mathbb{P}_p[0 \leftrightarrow \infty] \geq c(p - p_c)$.*

The proof of the theorem will be structured in two parts. First, we prove two general lemmas for convergent sequences of functions, a supporting one and a main one. The main lemma (Lemma 5.4) gives a result for the limiting behavior of a sequence of functions satisfying a certain differential inequality. Then, we will show that $p \mapsto \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_n]$ satisfies this inequality up to a constant, from which phase transition around the critical point p_c follows.

Lemma 5.3. *For a convergent sequence of real numbers $(a_n)_{n\in\mathbb{N}}$ tending to a, we have the following result:*

$$a = \lim_{n\to\infty} \frac{1}{\ln n} \sum_{i=1}^{n} \frac{a_i}{i}.$$

Proof. Observe that for $n \in \mathbb{N}$, there exists some $C_n \in \mathbb{R}$ such that

$$\sum_{i=1}^{n} \frac{1}{i} = \ln n + C_n, \tag{5.1}$$

where $|C_n|$ is bounded by a real number C_∞ for all $n \in \mathbb{N}$.

Now, define the sequence $(\epsilon_n)_{n\in\mathbb{N}}$ such that $a_n = a + \epsilon_n$ for all $n \in \mathbb{N}$. Note that the convergence of (a_n) implies that (ϵ_n) tends to 0. Then, we have

$$\frac{1}{\ln n} \sum_{i=1}^{n} \frac{a_i}{i} = \frac{1}{\ln n} \sum_{i=1}^{n} \frac{a + \epsilon_i}{i} = \frac{a}{\ln n} \sum_{i=1}^{n} \frac{1}{i} + \frac{1}{\ln n} \sum_{i=1}^{n} \frac{\epsilon_i}{i}.$$

By using (5.1) we see that the first term tends to a as n tends to ∞. The second term tends to 0, which we will show by splitting the sum into two parts. To that end, let

$$T_n = \frac{1}{\ln n} \sum_{i=1}^{n} \frac{\epsilon_i}{i}$$

and note that there exists, for every positive number ϵ, a natural number N such that $|\epsilon_n| < \epsilon$ for every $n \geq N$. We fix $\epsilon > 0$ and $N \in \mathbb{N}$ to get

$$|T_n| = \frac{1}{\ln n} \left| \sum_{i=1}^{N-1} \frac{\epsilon_i}{i} + \sum_{i=N}^{n} \frac{\epsilon_i}{i} \right| \leq \frac{1}{\ln n} \left| \sum_{i=1}^{N-1} \frac{\epsilon_i}{i} \right| + \frac{1}{\ln n} \sum_{i=N}^{n} \frac{\epsilon}{i}. \tag{5.2}$$

Using (5.1) in the second case below, fix any $\delta > 0$ and $M \in \mathbb{N}$ such that for all $n \geq M$, we have

$$\frac{1}{\ln n} \left| \sum_{i=1}^{N-1} \frac{\epsilon_i}{i} \right| \leq \delta \quad \text{and} \quad \frac{1}{\ln n} \sum_{i=1}^{n} \frac{1}{i} = 1 + \frac{C_n}{\ln n} \leq 1 + \delta.$$

Noting that $\sum_{i=N}^{n} i^{-1} \leq \sum_{i=1}^{n} i^{-1}$ and applying the above to (5.2) gives $|T_n| \leq \delta + \epsilon + \epsilon\delta$ for all $n \geq \max\{M, N\}$. Since ϵ and δ can be chosen arbitrarily small, T_n tends to 0 as n tends to ∞. $\square$

Lemma 5.4. [4] *Consider a convergent sequence of non-negative, increasing, continuously differentiable functions $f_n : [0, x_0] \to [0, M]$ bounded by some constant M, which satisfy*

$$f_n' \geq \frac{n}{\Sigma_n} f_n, \tag{5.3}$$

for all $n \geq 1$, where $\Sigma_n = \sum_{k=0}^{n-1} f_k$. Then, there exists $x_1 \in [0, x_0]$ such that

P1 *For any $x < x_1$, there exists $c_x > 0$ such that $f_n(x) \leq Me^{-c_x n}$.*

P2 *For $x > x_1$, $f = \lim_{n\to\infty} f_n$ satisfies $f(x) \geq x - x_1$.*

Proof. Define

$$x_1 = \inf\left\{x : \limsup_{n\to\infty} \frac{\ln \Sigma_n(x)}{\ln n} \geq 1\right\}.$$

Observe that the definition of x_1 together with the increasingness of Σ_n give the result that $\limsup_{n\to\infty} \frac{\ln \Sigma_n(x)}{\ln n} < 1$ when $x < x_1$ and $\limsup_{n\to\infty} \frac{\ln \Sigma_n(x)}{\ln n} \geq 1$ when $x > x_1$. We will prove the cases $x < x_1$ and $x > x_1$ separately.

Assume $x < x_1$. Fix $\delta > 0$ such that $x + 2\delta < x_1$ and set $x' = x + \delta$ and $x'' = x + 2\delta$ (so $x < x' < x'' < x_1$). We will show that $f(x)$ decays exponentially with respect to n. By our assumption,

$$\limsup_{n\to\infty} \frac{\ln \Sigma_n(x'')}{\ln n} < 1,$$

so there exists $\alpha > 0$ and an integer N such that

$$\sup_{n\geq N} \frac{\ln \Sigma_n(x'')}{\ln n} < 1 - \alpha,$$

which implies that

$$\Sigma_n(x'') < n^{1-\alpha} \quad \text{for all } n \geq N,$$

and the same upper bound holds on the interval $[0, x'']$ due to the monotonicity of Σ_n. Using this in (5.3) gives

$$\frac{f_n'}{f_n} \geq \frac{n}{\Sigma_n} > \frac{n}{n^{(1-\alpha)}} = n^\alpha.$$

Integrating the differential inequality between x' and x'' yields

$$\ln(f_n(x'')) - \ln(f_n(x')) \geq \delta n^\alpha,$$

and after multiplying by -1 and exponentiating each side, we get

$$\frac{f_n(x')}{f_n(x'')} \leq e^{-\delta n^\alpha}.$$

Finally, multiplying by $f_n(x'')$ and using the fact that f_n is bounded by M, we have

$$f_n(x') \leq Me^{-\delta n^\alpha} \quad \text{for all } n \geq N.$$

This implies that $\Sigma_\infty(x') = \sum_{n=0}^{\infty} f_n(x') \leq \sum_{n=0}^{\infty} Me^{-\delta n^\alpha} < \infty$ converges, which implies that the partial sums $\Sigma_n(x')$ for all $n \in \mathbb{N}$ are bounded by a constant real number Σ. The same upper bound holds on $[0, x']$, again due to the monotonicity of Σ_n. Applying inequality (5.3) once more to the previous result, we get

$$\frac{f_n'}{f_n} \geq \frac{n}{\Sigma_n} \geq \frac{n}{\Sigma},$$

and integrating that between x and x' gives

$$f_n(x) \leq f_n(x')e^{-\delta n/\Sigma} \leq Me^{-\delta n/\Sigma} \quad \text{for all } n \geq 0.$$

Since x is any number in $[0, x_1)$, **P1** follows as a result.

Assume $x > x_1$. For $n \geq 1$, define $T_n = \frac{1}{\ln n} \sum_{i=1}^{n} \frac{f_i}{i}$. The following inequality follows from (5.3):

$$T_n' = \frac{1}{\ln n} \sum_{i=1}^{n} \frac{f_i'}{i} \geq \frac{1}{\ln n} \sum_{i=1}^{n} \frac{f_i}{\Sigma_i}. \tag{5.4}$$

We also have the following general formula for $a, x \geq 0$:

$$\frac{x}{a} = \int_a^{a+x} \frac{1}{a} \mathrm{d}t \geq \int_a^{a+x} \frac{1}{t} \mathrm{d}t = \ln(a+x) - \ln a$$

since $\frac{1}{a} \geq \frac{1}{t}$ for $t \in [a, a+x]$. Because $\Sigma_i + f_i = \Sigma_{i+1}$, we can use this formula to get

$$\frac{f_i}{\Sigma_i} \geq \ln \Sigma_{i+1} - \ln \Sigma_i.$$

Applying this to (5.4) gives

$$T_n' \geq \frac{1}{\ln n} \sum_{i=1}^{n} (\ln \Sigma_{i+1} - \ln \Sigma_i) = \frac{\ln \Sigma_{n+1} - \ln \Sigma_1}{\ln n} \geq \frac{\ln \Sigma_{n+1} - \ln M}{\ln n},$$

Assuming $x' \in (x_1, x)$, we integrate this from x' to x, using the monotonicity of Σ_n together with $\Sigma_{n+1} \geq \Sigma_n$, to obtain

$$T_n(x) - T_n(x') \geq \int_{x'}^{x} \frac{\ln \Sigma_{n+1}(t) - \ln M}{\ln n} \mathrm{d}t \geq \frac{\ln \Sigma_n(x') - \ln M}{\ln n}(x - x').$$

Since the inequality holds for all n, we have

$$\sup_{m \geq n} (T_m(x) - T_m(x')) \geq (x - x') \left(\sup_{m \geq n} \frac{\ln \Sigma_m(x') - \ln M}{\ln m} \right).$$

By applying Lemma 5.3 pointwise, T_n tends to f as $n \to \infty$, so taking the limit as $n \to \infty$ on both sides gives

$$f(x) - f(x') \geq (x - x') \left(\limsup_{n \to \infty} \frac{\ln \Sigma_n(x') - \ln M}{\ln n} \right) = (x - x') \underbrace{\left(\limsup_{n \to \infty} \frac{\ln \Sigma_n(x')}{\ln n} \right)}_{\geq 1}.$$

By the definition of x_1 and our assumption of $x' > x_1$ this implies that

$$f(x) - f(x') \geq x - x'.$$

Letting $x' \to x_1$ from above (and recalling that f is non-negative) yields $f(x) \geq x - x_1$, which is precisely the proposition **P2**. $\square$

Proof of Theorem 5.2. Let

$$\mathbf{\Theta}_n = \mathbb{1}_{0 \leftrightarrow \partial\Lambda_n} \quad \text{and} \quad S_n = \sum_{k=0}^{n-1} \theta_k,$$

where θ_n is just another notation for $\mathbb{E}_p[\mathbf{\Theta}_n]$, so $\theta_n(p) = \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_n]$. We will use the OSSS-inequality (Theorem 4.7) to find a sufficient upper bound for θ_n such that we can apply Lemma 5.4. In particular, we will show that the existence of a certain family of algorithms $\mathbf{T}_k$ determining $\mathbf{\Theta}_n$ with a sufficient revealment upper bound *proves* that θ_n satisfies (5.3). This family of algorithms is the k-sphere algorithms introduced in Section 4.1 which have, for all $k \in [n]$, the upper bound for the revealment

$$\delta_e(\mathbf{\Theta}_n, \mathbf{T}_k) \leq \mathbb{P}_p[u \leftrightarrow \partial\Lambda_k] + \mathbb{P}_p[v \leftrightarrow \partial\Lambda_k] \tag{5.5}$$

for all edges $e = \{u, v\} \in E_n$, recalling that E_n is the set of edges with both endpoints in Λ_n.

Let $e = \{u, v\}$ be an edge in E_n. First, $\{u \leftrightarrow \partial\Lambda_k\}$ is the event where a vertex $u \in \Lambda_n$ is connected to the sphere $\partial\Lambda_k$ centered around the *origin*. For this event to occur, there must also be an open path connecting u to the sphere $\partial\Lambda_{|d(0,u)-k|}(u)$[2] centered around u, so $\{u \leftrightarrow \partial\Lambda_k\} \subset \{u \leftrightarrow \partial\Lambda_{|d(0,u)-k|}(u)\}$ and hence

$$\mathbb{P}_p[u \leftrightarrow \partial\Lambda_k] \leq \mathbb{P}_p[u \leftrightarrow \partial\Lambda_{|d(0,u)-k|}(u)] = \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_{|d(0,u)-k|}],$$

where translation invariance is used for the last equality. We sum over k to get, for all $u \in \Lambda_n$,

$$\begin{aligned}\sum_{k=1}^{n} \mathbb{P}_p[u \leftrightarrow \partial\Lambda_k] &\leq \sum_{k=1}^{n} \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_{|d(0,u)-k|}] \\ &= \sum_{k=0}^{d(0,u)-1} \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_k] + \sum_{k=1}^{n-d(0,u)} \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_k].\end{aligned}$$

Next, we use the fact that $\mathbb{P}_p[0 \leftrightarrow \partial\Lambda_{k+1}] \leq \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_k]$ implies that $\sum_{k=1}^{n} \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_k] \leq \sum_{k=0}^{n-1} \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_k]$ to obtain

$$\begin{aligned}\sum_{k=1}^{n} \mathbb{P}_p[u \leftrightarrow \partial\Lambda_k] &\leq \sum_{k=0}^{n-1} \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_k] + \sum_{k=0}^{n-1} \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_k] \\ &= 2S_n.\end{aligned}$$

[2] $d(u, v) = \max\{|u_i - v_i| : i \in [d]\}$

Assume that the family of algorithms $\mathbf{T}_k$ satisfies (5.5). Then, sum over all $k \in [n]$ and apply the previous result to get

$$\sum_{k=1}^{n} \delta_e(\mathbf{\Theta}, \mathbf{T}_k) \le \sum_{k=1}^{n} (\mathbb{P}_p[u \leftrightarrow \partial\Lambda_k] + \mathbb{P}_p[v \leftrightarrow \partial\Lambda_k]) \le 4S_n. \tag{5.6}$$

Now, recall that $\mathrm{Var}_p(\mathbf{\Theta}_n) = \theta_n(1-\theta_n)$. Theorem 4.7 gives the inequality

$$\theta_n(1-\theta_n) \le p(1-p) \sum_{e \in E_n} \delta_e(\mathbf{\Theta}, \mathbf{T}_k)\mathrm{Inf}_e(\mathbf{\Theta}_n).$$

Summing both sides over $k \in [n]$ and utilizing (5.6), we get

$$\begin{aligned} n\theta_n(1-\theta_n) &\le p(1-p) \sum_{k=1}^{n} \sum_{e \in E} \delta_e(\mathbf{\Theta}, \mathbf{T}_k)\mathrm{Inf}_e(\mathbf{\Theta}_n) \\ &\le p(1-p)4S_n \sum_{e \in E} \mathrm{Inf}_e(\mathbf{\Theta}_n), \end{aligned}$$

so

$$\sum_{e \in E} \mathrm{Inf}_e(\mathbf{\Theta}_n) \ge \frac{n}{4S_n p(1-p)} \theta_n(1-\theta_n) \ge \frac{n}{S_n}\theta_n(1-\theta_n),$$

where the last inequality follows from $4p(1-p) \le 1$ for all $p \in [0,1]$. By Lemma 3.10, $\sum_{e \in E} \mathrm{Inf}_e(\mathbf{\Theta}_n) = \theta_n'(p)$, so

$$\theta_n' \ge \frac{n}{S_n}\theta_n(1-\theta_n).$$

If we can find a constant lower bound for $1-\theta_n$, we will have shown that (θ_n) is, up to a constant, a sequence of functions of the kind used in Lemma 5.4. To that end, fix $p_0 \in (p_c, 1)$. Since

1. θ_n is increasing and
2. $\theta_n = \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_n] \le \mathbb{P}_p[0 \leftrightarrow \partial\Lambda_1] = \theta_1 \quad$ for all $n \ge 1$,

we have, for $p \le p_0$, $\theta_n(p) \le \theta_1(p_0)$ and thus

$$1 - \theta_n(p) \ge 1 - \theta_1(p_0). \tag{5.7}$$

If we let $f_n = \frac{1}{1-\theta_1(p_0)}\theta_n$, then $\Sigma_n = \sum_{i=0}^{n-1} f_i = \frac{1}{1-\theta_1(p_0)} S_n$, and we see that

$$f_n' = \frac{1}{1-\theta_1(p_0)}\theta_n' \ge \frac{1}{1-\theta_n(p_0)} \frac{n}{S_n}\theta_n(1-\theta_n).$$

Given (5.7), we have

$$\frac{1-\theta_n}{1-\theta_n(p_0)} \ge 1$$

for $p \in [0, p_0]$, so

$$f_n' \ge \frac{n}{S_n}\theta_n = \frac{n}{\Sigma_n} f_n,$$

meaning that f_n satisfies (5.3) on $[0, p_0]$. We apply Lemma 5.4 to obtain the following result. There exists $\tilde{p}_c \in [0, p_0]$ such that

1. For any $p < \tilde{p}_c$, there exists c_p such that $f_n(p) \leq \frac{1}{1-\theta_1(p_0)} e^{-c_p n}$.

   Since $f_n = \frac{1}{1-\theta_1(p_0)}\theta_n$, we have $\mathbb{P}_p[0 \leftrightarrow \partial\Lambda_n] \leq e^{-c_p n}$.

2. For $p > \tilde{p}_c$, $f = \lim_{n\to\infty} f_n$ satisfies $f(p) \geq p - \tilde{p}_c$.

   There exists $c > 0$ such that $f_n = \frac{1}{c}\theta_n$, and hence there exists $c > 0$ such that $\mathbb{P}_p[0 \leftrightarrow \infty] \geq c(p - \tilde{p}_c)$.

Recalling the definition $p_c := \sup\{p : \mathbb{P}_p[0 \leftrightarrow \infty] = 0\}$, we must have $\tilde{p}_c \leq p_c$ since $p < \tilde{p}_c$ implies that $\mathbb{P}_p[0 \leftrightarrow \infty] = 0$; and simultaneously $\tilde{p}_c \geq p_c$ since $p > \tilde{p}_c$ implies that $\mathbb{P}_p[0 \leftrightarrow \infty] > 0$. It follows that $\tilde{p}_c = p_c$, completing the proof. □

6 Discussion

6.1 Considerations when simulating revealment

Calculating the precise revealment of edges is infeasible for non-trivial algorithms. The only computationally viable way to get data on the revealment is to sample random configurations and measure the proportion of configurations in which edges are revealed. This is how the simulations in this thesis were conducted, with the main goal of visualizing the intuition behind the k-sphere algorithms and their role in providing an upper bound to the revealment. While Theorem 5.2 applies to all dimensions $d \geq 2$, the simulations in this thesis were only done in two dimensions at the critical point; the two-dimensional case is most useful for our visualizations, and the critical point for $d = 2$ is known, having value $p_c = \frac{1}{2}$.

The simulations were done at the critical point for a few reasons. First and foremost: We are interested in the percolation behavior near the critical point. But in addition to that, running the simulations with p far from the critical point has other unintended consequences. Simulating the revealment at subcriticality would bias the edges near the starting points of the algorithm, since edges far away from the starting points are unlikely to be connected to them by an open path and hence unlikely to be revealed. On the other hand, running the simulations at supercriticality would excessively favor the edges which come first in the fixed ordering of the algorithm, since many open connected paths likely exist in the configuration. It should be noted that there will *always* be a slight bias toward the edges that come first in the ordering, but its extent is insignificant for the purpose of visualization. The number of samples for each edge was chosen as 2000, so the statistical error (sampling standard deviation) for each revealment value is at most $\frac{1}{2\sqrt{2000}} \approx 0.011$.

6.2 Interpretation of the theorem and its proof

The earlier proofs by Men'shikov [9], Aizenman and Barsky [1], and Duminil-Copin and Tassion [5] involved defining two different phase transitions for percolation with regards to p; one with certain behavior in the subcritical region and the other with certain behavior in the supercritical. By then showing that the two critical points coincide, the implication that there is a sharp transition from one phase to the other across the coincident critical point follows.

The proof discussed in this thesis follows a slightly different structure. Using the differential inequality obtained by the OSSS-inequality and Margulis-Russo formula, Duminil-Copin shows that there exists *some* point $\tilde{p}_c$ which is a critical point for the properties described in Lemma 5.4. Those properties are very closely linked to the ones defining the critical point, so it follows directly that $\tilde{p}_c$ is in fact the critical point for Bernoulli percolation. Exploiting the OSSS-inequality, the proof of sharpness of percolation distills into the choice of a good algorithm followed by the characterization of a differential inequality.

The implications of Theorem 5.2 are perhaps not as immediately obvious as, for instance, the sharpness of the existence of an infinite cluster. One possible portrayal is as follows. In the subcritical region, the theorem characterizes the expected size of clusters on the lattice, stating that the probability of a cluster with radius n decays exponentially with n. A direct consequence is that the expected radius of open clusters is finite. Furthermore, the c_p parameter gives an upper bound to the *scale* at which one can expect to find clusters at a given parameter p. In the supercritical region, the probability of a given point belonging to the infinite cluster is linearly bounded from below. Importantly, the transition between these two phases occurs across a single critical point, p_c. Equivalently, the phase transition is sharp.

References

[1] Michael Aizenman and David J. Barsky. Sharpness of the phase transition in percolation models. *Communications in Mathematical Physics*, 108(3):489 – 526, 1987.

[2] Heinz Bauer. *Probability Theory*. Walter de Gruyter GmbH, Berlin/Boston, 1995.

[3] S. R. Broadbent and J. M. Hammersley. Percolation processes: I. crystals and mazes. *Mathematical Proceedings of the Cambridge Philosophical Society*, 53(3): 629–641, 1957. doi:10.1017/S0305004100032680.

[4] Hugo Duminil-Copin. Sharp threshold phenomena in statistical physics, 2018. URL `https://arxiv.org/abs/1810.03384`.

[5] Hugo Duminil-Copin and Vincent Tassion. A new proof of the sharpness of the phase transition for bernoulli percolation and the ising model. *Communications in Mathematical Physics*, 343(2):725–745, November 2015. ISSN 1432-0916. doi:10.1007/s00220-015-2480-z. URL `http://dx.doi.org/10.1007/s00220-015-2480-z`.

[6] Geoffrey Grimmett. *Percolation*, volume 321 of *Grundlehren der mathematischen Wissenschaften*. Springer Berlin, Heidelberg, second edition, 1999. doi:10.1007/978-3-662-03981-6.

[7] G. A. Margulis. Probabilistic characteristics of graphs with large connectivity. *Probl. Peredachi Inf.*, 10:101 – 108, 1974. URL https://www.mathnet.ru/eng/ppi1034.

[8] Declan McNamara. Kolmogorov's zero-one law with applications. 2017. URL https://api.semanticscholar.org/CorpusID:21731442.

[9] M. V. Men'shikov. Coincidence of critical points in percolation problems. *Doklady Akademii Nauk SSSR*, 288:1308 – 1311, 1986. URL https://www.mathnet.ru/eng/dan8543.

[10] Ryan O'Donnell, Michael E. Saks, Oded Schramm, and Rocco A. Servedio. Every decision tree has an influential variable. *CoRR*, abs/cs/0508071, 2005. URL http://arxiv.org/abs/cs/0508071.

[11] L. Russo. On the critical percolation probabilities. *Z. Wahrscheinlichkeitstheorie verw Gebiete*, 56:229 – 237, 1981. doi:10.1007/BF00535742.

A k-sphere algorithms

The original definition of the k-sphere algorithms in the source material seems to suffer from an issue when encountering configurations wherein percolation is dependent on the edges on the k-sphere.

Recall that the algorithm is constructed with the intention of guaranteeing the upper bound for the revealment of any edge $e = \{x, y\}$,

$$\delta_e(\mathbf{f}, \mathbf{T}_k) \leq \mathbb{P}_p[\partial\Lambda_k \leftrightarrow x] + \mathbb{P}_p[\partial\Lambda_k \leftrightarrow y],$$

where $\mathbf{T}_k$ is the family of algorithms as defined in [4, p. 14]. This upper bound follows from the assumption that

$$\{\exists t \leq \tau : \mathbf{e}_t = e\} \subset \{\partial\Lambda_k \leftrightarrow x\} \cup \{\partial\Lambda_k \leftrightarrow y\}. \tag{A1}$$

However, due to a small technicality, this subset relation does not always hold.

One edge at a time, $\mathbf{T}_k$ will begin to reveal all open clusters connected to the k-sphere. When all the clusters have been revealed, we say that the *halting condition* is reached. In [4], it is asserted that this never occurs before the stopping time, τ (Definition 4.2), which justifies (A1). However, this assertion presupposes that the algorithm is able to surely determine percolation before the halting condition, which we maintain is not the case.

As edges are revealed, the algorithm accumulates all vertices which are contained in the open clusters. Notably, $\mathbf{T}_k$ is defined not to reveal any edges with *both* endpoints in the accumulated open clusters before the halting condition is reached, so the k-sphere itself is not revealed until after the halting condition. As shown in Figure A1, there are configurations for which percolation cannot be surely determined before revealing the edges on the k-sphere. For such configurations, the halting condition occurs *before* the stopping time, which violates the assertion necessary to obtain (A1).

Resolving this issue is the motivation for our slight alteration of the algorithm, which simply amounts to also revealing edges with both endpoints in the hitherto discovered connected open clusters.

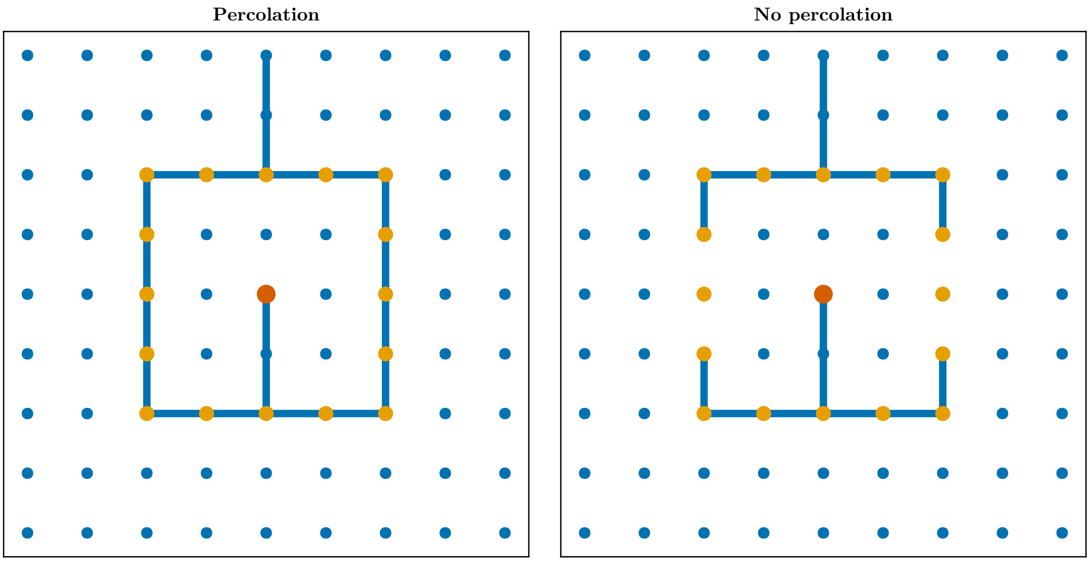

Figure A1: Example of configurations which pose a problem to the original definition of the k-sphere algorithms found in [4, p. 14]. The yellow vertices are those of the k-sphere. All open clusters connected to the k-sphere (excluding the k-sphere itself) will be discovered before the stopping time, τ.

B Maturity test (in Swedish)

År 2018 publicerade Duminil-Copin en artikel med ett nytt bevis för skarpheten av Bernoulliperkolation på det d-dimensionella heltalsgittret $\mathbb{Z}^d$. Satsen bevisades ursprungligen år 1986 av Men'shikov, och sedan dess har ett antal andra nya bevis publicerats. Duminil-Copins bevis, som behandlas i denna avhandling, bidrar med ett nytt perspektiv till undersökningen av perkolation.

I denna avhandling granskas beviset, och dess delsteg utvecklas med målet att göra det åtkomligt för läsare utan tidigare kunskap inom perkolationsteori. Som hjälpmedel läggs det även fram visualiseringar av en simulering av ett av bevisens nyckelsteg för att förse läsaren med en intuition för de underliggande idéerna.

Eftersom ämnet i fråga nödgar tillämpningen av sannolikhetsteori på en kunskapsnivå som något överskrider den en kandidatstuderande förväntas besitta innefattar avhandlingen en kort introduktion till dessa mer avancerade koncept. I samma anda har hänsyn tagits till att från grundläggande begrepp konstruera de verktyg som brukas för att analysera Bernoulliperkolation.

En av konstruktionerna som presenteras i avhandlingen är så kallade *konfigurationer* på heltalsgittret. Dessa konfigurationer kan återges som sekvenser av oberoende och identiskt Bernoullifördelade variabler, och frågan om det förekommer perkolation på en konfiguration reduceras således till resultatet av en boolesk funktion. Ett annat viktigt verktyg i beviset är en formel av Margulis och Russo, som erbjuder ett mycket nyttigt sätt att uttrycka derivatan av en boolesk funktions väntevärde med avseende på Bernoulli-parametern p.

Säreget för det nya beviset är dess användning av en statistisk olikhet för booleska funktioner av O'Donnell, Schramm, Saks och Servedio. Olikheten använder sig av koncept från datavetenskapen som kan tillämpas för att dra sannolikhetsteoretiska slutsatser om en funktion som är svår att karaktärisera som sådan. Istället definieras en välvald deterministisk *algoritm* som bevisligen förmår att evaluera funktionen. Algoritmen är utformad så att den har en kritisk egenskap vilken tillsammans med den ovannämnda olikheten och Margulis-Russo-formeln lägger grunden för det i avhandlingen centrala beviset, det vill säga skarpheten av Bernoulliperkolation på $\mathbb{Z}^d$ kring en kritisk punkt.