Differential Inequalities for the Random Cluster Model

Sunay Joshi

December 4, 2023

§1 Introduction

Statistical mechanics studies how microscopic interactions give rise to macroscopic phenomena. Two of the most well-known models in the subject are Bernoulli percolation and the Ising model.

Bernoulli bond percolation, which we refer to simply as percolation, was invented by Broadbent and Hammersley [DC18]. In the model, each edge e of the lattice $\mathbb{Z}^d$ is included (open) or excluded (closed) independently according to a Bernoulli random variable with parameter $p_e = 1 - e^{-\beta J_e}$. The resulting random graph is known as the percolation configuration, and we seek to understand properties of its connected clusters. Although percolation is a simple model, it possesses a number of remarkable properties. The main question regarding the model is whether the cluster C_0 of the origin is infinite with positive probability. If this occurs, then we say that the configuration percolates, in the sense that the origin is connected to the point at infinity. It turns out that the model exhibits a phase transition at a certain critical parameter β_c, so that below β_c, there is almost surely no infinite cluster containing 0, while above β_c the cluster of the origin is infinite with positive probability [DC18].

In contrast to Bernoulli percolation, the Ising model introduces dependency between individual components of the model. The Ising model originated in physics as a model of ferromagnetism [DC17]. Ferromagnetism is a property that we are all well acquainted with; ordinary metals can become magnetized when exposed to a strong external magnetic field, thereby becoming magnets in their own right. Consider a ferromagnetic metal placed in an external field h at some temperature T. The Ising model posits that a magnet consists of many interacting dipoles, each having spin up $(+1)$ or spin down (-1). The energy (Hamiltonian) of the configuration has two determinants: first, the amount of agreement between neighboring sites of $\mathbb{Z}^d$, and second, the amount of agreement with the external field. Configurations with lower energies are more likely to be realized, so that the configuration can be thought of as a sample from a certain probability distribution. The magnetization is defined as the expected spin of the origin. The key question is at which temperatures the phenomenon of spontaneous magnetization occurs, when the material retains nonzero magnetization even as the external field is decreased to zero. As it turns out, the Ising model exhibits a phase transition at a critical temperature $T = T_c$ called the Curie temperature. Equivalently, the spontaneous magnetization is positive above a critical inverse temperature $\beta = \beta_c$ [DC17].

The phase transitions in both percolation and the Ising model have an additional feature. Below the critical point, the models exhibit exponential decay of correlations. However, above the critical point, the correlations are bounded below by a positive constant, so that there is long-range order [DC18] [DC17]. Stated formally, in either percolation or the Ising model, if β_T denotes the largest β for which the two-point function exhibits exponential decay, and if β_H denotes the smallest value of β for which there is long-range order, then we have the following result [AB87] [ABF87].

Theorem 1.1 (Sharpness of percolation and Ising)
For all $d \geq 2$, both Bernoulli percolation and the Ising model with free boundary condition on $\mathbb{Z}^d$ satisfy $\beta_T = \beta_H$.

This theorem, known as the sharpness of the phase transition, is remarkable; although it is not hard to establish exponential decay for sufficiently small β and long-range order for sufficiently large β by means of perturbative methods, it is not at all clear that the decaying regime should extend up to the start of long-range interactions. The common technique behind the proofs of the sharpness of the phase transition is the concept of differential inequalities [AB87] [ABF87].

The aim of this report is to exposit results on the sharpness of the phase transition for another model, the random cluster model. The random cluster model, introduced by Fortuin and Kasteleyn, is a generalization of percolation and the Ising model that is less well understood, with sharpness only proved recently [Gri06] [DCRT18]. In this report we present a number of differential inequalities that have been derived for the random cluster model, including tail bounds on the radius and volume of the cluster of the origin. The ultimate goal of differential inequalities is to gain an understanding of critical exponents, and we touch on this subject at the end of the paper.

The structure of the report is as follows: in Section 2 we present background; in Section 3 we derive partial differential inequalities for percolation; in Section 4 we present the theory of decision trees and forests; and in Section 5 we present three applications of the theory to the random cluster model.

§2 Background

In this section, we quickly recall the features of Bernoulli percolation, the Ising model, and the random cluster model that are relevant to the body of the report.

§2.1 Bernoulli percolation

In this section, we follow the papers [DC18] and [AB87].

Consider the integer lattice $\mathbb{Z}^d$, where $d \geq 1$. We equip $\mathbb{Z}^d$ with the nearest-neighbor graph structure, where two lattice points x and y are joined by an edge if and only if $\|x - y\| = 1$, where $\|\cdot\|$ denotes the ℓ_1-norm. We write $x \sim y$ if $\|x - y\| = 1$. In keeping with the terminology of statistical mechanics, we also refer to vertices and edges as sites and bonds, respectively. We now take a finite connected subgraph Λ of $\mathbb{Z}^d$ with edges E and a parameter $\beta > 0$ and construct a random graph on Λ in the following manner: for each bond $e = xy \in E$, e is included in the graph with probability $p_e = 1 - e^{-\beta J_e}$, independent of all other edges, where J_e is a nonnegative coupling constant. A bond e is called occupied if $\omega_e = 1$. In other words, if ω_e the indicator random variable that e is occupied, then ω_e follows a Bernoulli distribution with parameter p_e.

The resulting random graph ω is called a configuration, and the probability measure on the set of configurations is a product measure over all so-called bond variables ω_e. Explicitly, the measure can be written as

$$\mathbb{P}_\beta[\omega] = \prod_{e \in E} p_e^{\omega_e}(1-p_e)^{1-\omega_e}.$$

If $p_e = p$ for all bonds e, then the measure can be rewritten as $p^{o(\omega)}(1-p)^{c(\omega)}$, where $o(\omega)$ and $c(\omega)$ denote the number of open and closed bonds, respectively. See Figure 1 for an depiction of a percolation configuration.

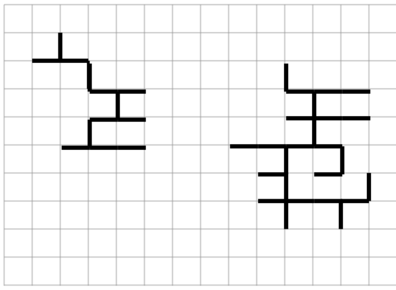

Figure 1: A percolation configuration ω on $\Lambda \subseteq \mathbb{Z}^2$.

The model can be extended to $\Lambda = \mathbb{Z}^d$ by a standard argument. Formally, the infinite volume measure is the infinite product measure over all bond variables ω_e. In infinite volume, we require the coupling constants to be translation invariant and we require that $|J| = \sum_x J_{0x}$ is finite.

To fix notation, given sites x and y, let $x \overset{S}{\leftrightarrow} y$ denote the event that x and y are connected by a path of open bonds in a given subset S of Λ; if $S = \Lambda$, then we typically drop the label. Given a site x, let C_x denote the connected component of x. We refer to C_x as the cluster containing x. For percolation on $\mathbb{Z}^d$ with $d \geq 2$, let $\theta(\beta) = \mathbb{P}_\beta[|C_0| = \infty]$. The existence of a phase transition states that there exists $\beta_c > 0$ such that $\theta(\beta) = 0$ for $\beta < \beta_c$, while $\theta(\beta) > 0$ for $\beta > \beta_c$. The critical points β_T and β_H for percolation on $\mathbb{Z}^d$ are defined as $\beta_T = \sup\{\beta : \exists c_\beta > 0 \text{ s.t. } \mathbb{P}_\beta[x \leftrightarrow y] \leq e^{-c_\beta \|x-y\|} \text{ for all } x, y \in \mathbb{Z}^d\}$ and $\beta_H = \inf\{\beta : \theta(\beta) > 0\}$.

Before continuing, we collect a few standard tools from probability theory that are indispensable in the study of percolation.

Consider the ordering on configurations given by $\omega \leq \omega'$ if and only if $\omega_e \leq \omega'_e$ for all $e \in E$. A random variable $X : \Omega \to \mathbb{R}$ is called increasing if $\omega \leq \omega'$ implies that $X(\omega) \leq X(\omega')$. An event A is called increasing if the indicator random variable $\mathbf{1}_A$ is increasing. The Fourtuin–Kasteleyn–Ginibre (FKG) inequality states that increasing random variables are positively correlated with respect to the percolation measure.

Proposition 2.1 (FKG inequality)
For Bernoulli percolation on $\mathbb{Z}^d$, $\mathbb{E}_\beta[fg] \geq \mathbb{E}_\beta[f]\mathbb{E}_\beta[g]$ for all increasing random variables f and g. In particular, $\mathbb{P}_\beta[A \cap B] \geq \mathbb{P}_\beta[A]\mathbb{P}_\beta[B]$ for all increasing events A and B.

Next, recall that given an event A and a configuration ω, a set of edges I is called a witness for A if $\omega'(e) = \omega(e)$ for all $e \in I$ implies that $\omega' \in A$. Two events A and B are disjointly-realized if there exist disjoint witnesses for A and B. The event that A and B are disjointly-realized is denoted $A \circ B$. The van der Berg–Kesten (BK) inequality provides a sort of converse to FKG.

Proposition 2.2 (BK inequality)

For Bernoulli percolation on $\mathbb{Z}^d$, if A and B are events that depend on finitely many edges, then $\mathbb{P}_\beta[A \circ B] \leq \mathbb{P}_\beta[A]\mathbb{P}_\beta[B]$.

Finally, we state Russo's formula, which allows us to compute the rate of change of expectations and probabilities with respect to bond probabilities. Given an increasing event A and a configuration ω, an edge e is pivotal for A if A occurs when e is open but A does not occur when e is closed.

Proposition 2.3 (Russo's formula for percolation)

Given a finite subset Λ of $\mathbb{Z}^d$ and a function $f : \{0,1\}^\Lambda \to \mathbb{R}$, we have

$$\frac{\partial}{\partial p_e}\mathbb{P}_\beta[f(\omega)] = \mathrm{Cov}[f(\omega), \omega_e],$$

where covariance is taken with respect to the measure $\mathbb{P}_\beta$ on Λ. Thus, for Bernoulli percolation on $\mathbb{Z}^d$, if A is an increasing event that depends on finitely many edges and if e is an edge, then

$$\frac{\partial}{\partial p_e}\mathbb{P}_\beta[A] = \mathbb{P}_\beta[e \text{ pivotal for } A].$$

Although stated in the context of percolation, Russo's formula holds for any product measure on a finite sample space, and we will make use of this fact below.

§2.2 Ising model

In this section, we follow the papers [DC17] and [ABF87].

Fix an inverse temperature $\beta = \frac{1}{T} > 0$ and a constant external magnetic field $h > 0$. As in percolation, we work on a finite connected subgraph Λ of the lattice $\mathbb{Z}^d$. We assign a spin $\sigma_x \in \{\pm 1\}$ to each $x \in \Lambda$, and we let σ denote the resulting configuration. (Since we do not make use of arbitrary boundary conditions on the Ising model in this report, we simply use the free boundary condition f; however, this will not be the case for the random cluster model.) The Hamiltonian of the configuration σ is given by the sum

$$H^f_{\Lambda,h}(\sigma) = -\sum_{x\sim y} J_{xy}\sigma_x\sigma_y - h\sum_{x\in\Lambda}\sigma_x,$$

where J_{xy} are nonnegative coupling constants between adjacent sites. The partition function of the model is defined as the sum of the so-called Gibbs factors over all configurations:

$$Z^f_{\Lambda,\beta,h} = \sum_{\sigma\in\{\pm1\}^\Lambda} e^{-\beta H^f_{\Lambda,h}(\sigma)}.$$

The Ising probability measure with free boundary condition from which the configuration σ is sampled is given as

$$\mu^f_{\Lambda,\beta,h}[\sigma] = \frac{1}{Z^f_{\Lambda,\beta,h}} e^{-\beta H^f_{\Lambda,h}(\sigma)}.$$

For convenience, we drop subscripts and superscripts when they are clear from context.

As with percolation, the model can be extended to $\Lambda = \mathbb{Z}^d$. Assume that the coupling constants are translation invariant and that $|J| = \sum_{x\in\mathbb{Z}^d} J_{0x} < \infty$. Formally, the infinite volume Ising measure with free boundary condition is given as

$$\mu^f_{\beta,h}[\sigma] = \frac{1}{Z^f_{\beta,h}} e^{-\beta H^f_h(\sigma)} d\rho(\sigma),$$

where ρ is the uniform product measure on all configurations on $\mathbb{Z}^d$.

We now introduce our notions of interaction. Consider the Ising model on $\mathbb{Z}^d$ with $d \geq 2$ with free boundary condition. Given a random variable X, we denote its expectation with respect to the measure $\mu^f_{\beta,h}$ as $\langle X\rangle_{\beta,h}$, and we drop labels for convenience. The spin-spin correlation between two sites x and y is defined as the expected value $\langle \sigma_x\sigma_y\rangle$. The magnetization $M(\beta, h)$ of the model is defined as the expected spin $\langle\sigma_0\rangle$ at the origin. As explained above, we are interested in the spontaneous magnetization, which is given by sending the external field h to zero: $M(\beta) = \lim_{h\downarrow 0} M(\beta, h)$. The magnetic susceptibility is defined as the sum $\chi(\beta, h) = \sum_{x\in\mathbb{Z}^d}\langle\sigma_0\sigma_x\rangle$, and we write $\chi(\beta) = \lim_{h\downarrow 0}\chi(\beta, h)$. The existence of a phase transition states that for $d \geq 2$, there exists $\beta_c > 0$ such that $M(\beta) = 0$ for $\beta < \beta_c$, while $M(\beta) > 0$ for $\beta > \beta_c$. The critical points β_T and β_H for the Ising model on $\mathbb{Z}^d$ are defined as

$$\beta_T = \sup\{\beta : \exists c_\beta > 0 \text{ s.t. } \langle\sigma_x\sigma_y\rangle \leq e^{-c_\beta\|x-y\|} \text{ for all } x, y \in \mathbb{Z}^d\}$$

and $\beta_H = \inf\{\beta : M(\beta) > 0\}$.

§2.3 Random cluster model

In this section, we follow the papers [DC17] and [DCRT18].

We now introduce the random cluster model. Although not obvious at first glance, the model is a unification of a wide swath of statistical mechanics models, including both Bernoulli percolation and the Ising model. Although the set-up is rather similar to that of Bernoulli percolation, there is a key difference in the probability measure that causes the states of distinct bonds to in general be dependent.

Fix $\beta > 0$ and $q \geq 1$, and let $p_e = 1 - e^{-\beta J_e}$ for nonnegative coupling constants J_e. Consider a finite connected subgraph Λ of the lattice $\mathbb{Z}^d$ with edges E. A boundary condition ξ is a partition $P_1 \sqcup \ldots \sqcup P_m$ of the boundary $\partial\Lambda = \{x \in \Lambda : \exists y \in \mathbb{Z}^d \setminus \Lambda \text{ s.t. } x \sim y\}$. The free boundary condition is $f = \bigsqcup_{x\in\partial\Lambda}\{x\}$, and the wired boundary condition is $w = \{\partial\Lambda\}$. Given a configuration of bonds ω, the graph ω^ξ is ω with all sites $x \in P_i$ identified for each $1 \leq i \leq m$. The random cluster measure on Λ with boundary condition ξ can be written as

$$\phi^\xi_{\Lambda,\beta,q}[\omega] = \frac{1}{Z^\xi_{\Lambda,\beta,q}} q^{k(\omega^\xi)} \prod_{e\in E} p_e^{\omega_e}(1-p_e)^{1-\omega_e},$$

where $k(\omega^\xi)$ denotes the number of clusters in ω^ξ. Note that if $q = 1$, then we recover bond percolation on Λ. However, when $q > 1$, the states of distinct bonds are dependent due to the presence of the factor $q^{k(\omega)}$ that favors configurations with more clusters. In particular, when $q = 2$, the Edwards–Sokal coupling couples the random cluster model to the Ising model. When $\xi \in \{f, w\}$, the model can be extended to $\Lambda = \mathbb{Z}^d$ via limits of finite graphs.

The notation used for percolation carries over to the random cluster model. In particular, for the random cluster model on $\mathbb{Z}^d$ with $d \geq 2$ with boundary condition $\xi \in \{f, w\}$, $\theta^\xi(\beta)$ is the probability that C_0 is infinite. Given $\xi \in \{f, w\}$, the existence of a phase transition states that there exists $\beta_c^\xi > 0$ such that $\theta^\xi(\beta) = 0$ for $\beta < \beta_c^\xi$, while $\theta^\xi(\beta) > 0$ for $\beta > \beta_c^\xi$. It is known that $\beta_c^f = \beta_c^w$, so that we may drop the superscript without ambiguity [Gri06].

We now present the probabilistic tools that generalize to the random cluster model. These properties are most easily stated in the general context of monotonic measures. Given a graph $G = (V, E)$ with edges E and vertices V, a measure μ on $\{0,1\}^E$ is monotonic if for all $F \subseteq E$, for all $\xi, \zeta \in \{0,1\}^F$ with $\xi \leq \zeta$, and for all increasing subsets A of $\{0,1\}^F$, we have that $\mu[A|\omega_e = \xi_e \forall e \in F] \leq \mu[A|\omega_e = \zeta_e \forall e \in F]$ whenever $\mu[\omega_e = \xi_e \forall e \in F] > 0$ and $\mu[\omega_e = \zeta_e \forall e \in F] > 0$ [DCRT18]. The random cluster model is monotonic for $q \geq 1$ and inherits the following properties.

Proposition 2.4 (Positive association)
Given a finite connected subgraph Λ of $\mathbb{Z}^d$ and $q \geq 1$, we have:

- (Comparison between boundary conditions) If A is an increasing event and $\xi \leq \xi'$ are boundary conditions, then $\phi^\xi_{\Lambda,\beta,q}[A] \leq \phi^{\xi'}_{\Lambda,\beta,q}[A]$.
- (Monotonicity) If A is an increasing event and $\beta \leq \beta'$, then $\phi^\xi_{\Lambda,\beta,q}[A] \leq \phi^\xi_{\Lambda,\beta',q}[A]$.
- (FKG inequality) If A and B are increasing events, then

$$\phi^\xi_{\Lambda,\beta,q}[A \cap B] \geq \phi^\xi_{\Lambda,\beta,q}[A]\phi^\xi_{\Lambda,\beta,q}[B].$$

It is often useful to condition on the states of a set of bonds. The following proposition ensures that the resulting conditional distribution is again a random cluster model with a modified boundary condition.

Proposition 2.5 (Domain Markov Property)
Given a finite connected subgraph $\Lambda = (V, E)$ of $\mathbb{Z}^d$, a subgraph $\Lambda' = (V', E')$ of Λ, a boundary condition ξ, and configurations ψ and ψ' on $E \setminus E'$ and E' respectively, we have

$$\phi^\xi_{\Lambda,\beta,q}[\omega_{|_{E'}} = \psi' | \omega_{|_{E\setminus E'}} = \psi] = \phi^{\psi^\xi}_{\Lambda',\beta,q}[\psi'].$$

Here $\omega_{|_E}$ denotes the restriction of ω to the set E.

Finally, we state two analogues of Proposition 2.3. The finite volume analogue is a direct generalization [Hut20].

Proposition 2.6 (Russo's formula for random cluster model)
Given a finite connected subgraph Λ of $\mathbb{Z}^d$ and a function $f : \{0,1\}^\Lambda \to \mathbb{R}$, we have

$$\frac{d}{d\beta}\phi^{\xi}_{\Lambda,\beta,q}[f(\omega)] = \sum_{e\in E} \frac{J_e}{e^{\beta J_e}-1}\mathrm{Cov}[f(\omega),\omega_e],$$

where covariance is taken with respect to the measure $\phi^{\xi}_{\Lambda,\beta,q}$.

There is also the following infinite volume generalization, first proven in [Hut20]. Unlike the finite volume case, it takes the form of an inequality. Given a function $f : [a,b] \to \mathbb{R}$, we define its lower-right Dini derivative at $x \in [a,b)$ as

$$\left(\frac{d}{dx}\right)_+ f(x) = \liminf_{\varepsilon\downarrow 0} \frac{f(x+\varepsilon)-f(x)}{\varepsilon}.$$

As an example, the function $f(x) = \begin{cases} x\sin(1/x) & x\neq 0 \\ 0 & x=0 \end{cases}$ has lower-right Dini derivative -1 at $x=0$. Note also that

$$\left(\frac{d}{dx}\right)_+ \log f(x) = \frac{1}{f(x)}\left(\frac{d}{dx}\right)_+ f(x).$$

Proposition 2.7 (Infinite volume Russo's formula for random cluster model)
Given an increasing function $f : \{0,1\}^{\mathbb{Z}^d} \to \mathbb{R}$ and boundary condition $\xi \in \{f,w\}$, we have

$$\left(\frac{d}{d\beta}\right)_+ \phi^{\xi}_{\beta,q}[f(\omega)] \geq \sum_{e\in E} \frac{J_e}{e^{\beta J_e}-1}\mathrm{Cov}[f(\omega),\omega_e]$$

for all $\beta \geq 0$, where covariance is taken with respect to the measure $\phi^{\xi}_{\beta,q}$.

Note that there is no analog of the BK inequality for the random cluster model. Historically, the BK inequality has been at the heart of proofs of the sharpness of percolation [AB87] [Men86]. As we will see below, the random cluster model requires additional tools to analyze.

§3 Sharpness of Bernoulli percolation

We begin by presenting the key result in the Aizenman–Barsky proof of the sharpness of the phase transition in Bernoulli bond percolation on $\mathbb{Z}^d$: a pair of partial differential inequalities for the model. While other approaches exist – most notably, Menshikov's proof in [Men86] – the advantage of the partial differential inequalities of Aizenman–Barsky is the fact that they yield critical exponent bounds that match the mean-field limit. The quantity at the heart of our differential inequalities is a certain notion of magnetization M for Bernoulli percolation. We will arrive at the definition of M by way of an analogy with the Ising model. This is one instance of the fruitful interplay between percolation and the Ising model. In this section, we follow the paper [AB87].

As stated above, the quantity at the heart of our percolation inequalities is a certain analogue of the magnetization $M(\beta, h)$. We proceed formally. Recall that the Hamiltonian of the Ising model on $\mathbb{Z}^d$ with free boundary condition is $H(\sigma) = -\sum_{x\sim y} J_{xy}\sigma_x\sigma_y - h\sum_{x\in\mathbb{Z}^d}\sigma_x$. We now augment the system by introducing a Griffiths "ghost site" $\hat{y}$ with spin $\hat{\sigma}$ sampled uniformly from $\{\pm 1\}$ and coupling constants $J_{x\hat{y}} = 1 - e^{-h}$ for all sites $x \in V = \mathbb{Z}^d$. We modify the spins of the remaining sites by setting $\tilde{\sigma}_x = \sigma_x\hat{\sigma}$. The Hamiltonian of the modified system with *zero* external field is $\tilde{H} = -\sum_{x,y\in V'} J_{xy}\hat{\sigma}_x\hat{\sigma}_y$, where $V' = V \cup \{\hat{y}\}$. Splitting the sum, we find

$$\tilde{H} = -\sum_{x,y\in V, x\sim y} J_{xy}\hat{\sigma}_x\hat{\sigma}_y - \sum_{x\in V} J_{x\hat{y}}\hat{\sigma}_x\hat{\sigma}.$$

Since $\hat{\sigma}_x\hat{\sigma}_y = \sigma_x\sigma_y$ and $J_{x\hat{y}}\hat{\sigma}_x\hat{\sigma} = (1 - e^{-h})\sigma_x$, we have

$$\tilde{H} = -\sum_{x\sim y} J_{xy}\sigma_x\sigma_y - (1 - e^{-h})\sum_{x\in V}\sigma_x,$$

which as $h \downarrow 0$ is precisely the original Hamiltonian H. We see that the ghost site trick allows us to convert the Ising model *with* an external field to an Ising model *without* an external field, simply by adding an additional site. Not only that, but the two-point function $\langle\tilde{\sigma}_x\hat{\sigma}\rangle$ is equal to the magnetization $\langle\sigma_x\rangle$. Therefore we can convert questions about magnetization to questions about the two-point function, and vice versa.

Having established this dictionary for Ising, we define $M(\beta, h)$ for percolation. Append a ghost site g such that the bond xg is open with probability $1 - e^{-h}$ for all x. By the above, the two-point function of the modified system should correspond to the magnetization of the original system. Thus we define $M(\beta, h) = \mathbb{P}[0 \leftrightarrow g]$ for percolation.

Note that there is a way to define M that avoids the introduction of a ghost site. At each site x, independently assign a site variable $\eta_x \in \{0, 1\}$, where $\eta_x = 1$ with probability $1 - e^{-h}$. We let $\mathcal{G}$ denote the set of sites x with $\eta_x = 1$, and we call $\mathcal{G}$ the ghost field of green sites. The correspondence to the ghost site construction is clear: the set of green sites corresponds to the set of sites x for which the bond xg is occupied. As a result, $M(\beta, h) = \mathbb{P}[0 \leftrightarrow \mathcal{G}]$, where $\{0 \leftrightarrow \mathcal{G}\} = \{C_x \cap \mathcal{G} \neq \emptyset\}$. In the proof of the differential inequalities below, we primarily utilize the perspective of $\mathcal{G}$. We abuse notation and let $\mathbb{P}$ denote the product measure.

We now compute certain functions of M that are relevant to our inequalities. Let us condition on the set of bonds in the connected component C_0. Note that 0 is not connected to $\mathcal{G}$ if and only if none of the sites $x \in C_0$ are green. Each site x is not green with probability e^{-h}, hence all sites are not green with probability $e^{-h|C_0|}$. It follows that $M(\beta, h) = \mathbb{P}[1 - e^{-h|C_0|}]$. Sending $h \downarrow 0$,

$$M(\beta) = \lim_{h\downarrow 0} M(\beta, h) = \mathbb{P}[\mathbf{1}_{|C_0|=\infty}] = \mathbb{P}[|C_0| = \infty] = \theta(\beta).$$

In other words, $M(\beta) > 0$ is equivalent to percolation to infinity, and when $h > 0$, the event $\{0 \leftrightarrow \mathcal{G}\}$ is a surrogate for percolation to infinity. Next, we define the susceptibility

$$\chi(\beta, h) = \frac{\partial M}{\partial h} = \mathbb{P}[|C_0|e^{-h|C_0|}].$$

As $h \to 0$, $\chi(\beta, h)$ tends to $\mathbb{P}[|C_0| \cdot \mathbf{1}_{|C_0|<\infty}]$. Hence by definition of β_H, we have $\chi(\beta, 0) = \mathbb{E}[|C_0|]$ for $\beta < \beta_H$. In fact, for $h > 0$, we can write $\chi(\beta, h) = \mathbb{P}[|C_0| \cdot \mathbf{1}_{0\not\leftrightarrow\mathcal{G}}]$.

We are ready to state our differential inequalities.

Theorem 3.1 (Partial differential inequalities)
The following two inequalities hold for percolation on $\mathbb{Z}^d$ with $d \geq 2$:

$$\frac{\partial M}{\partial \beta} \leq |J| M \frac{\partial M}{\partial h} \qquad \text{(Burgers' inequality)}$$

$$M \leq h \frac{\partial M}{\partial h} + \beta M \frac{\partial M}{\partial \beta} + M^2 \qquad (\phi^3 \text{ inequality})$$

The first inequality is a Burgers' inequality. The Burgers' equation $\frac{\partial v}{\partial t} = v \frac{\partial v}{\partial x}$ is the simplest partial differential equation with shock formation. The equation, which models the velocity $v(x,t)$ of a fluid moving in one dimension, exhibits shocks when the initial condition $v(x,0)$ is decreasing in x. As a result, the solution becomes discontinuous at some time t_c [Lan11]. The correspondence to the Burgers' inequality is via $v \leftrightarrow M$, $t \leftrightarrow \beta$, and $x \leftrightarrow h$. The second inequality is known as a ϕ^3 inequality due to its relation to field theories in physics.

We now prove the two inequalities. Note that we must consider a finite volume approximation on the box $\Lambda_L = (-L, L]^d \cap \mathbb{Z}^d$ with periodic boundary conditions to ensure translation invariance. Sending $L \to \infty$ yields the inequalities in infinite volume. For details, see [AB87].

In addition, throughout the proof, we apply Russo's formula (Proposition 2.3) to functions on the space of configurations (ω, η) of bonds and green sites. This is valid because the configuration (ω, η) is sampled from a product measure on the set $E \cup \mathbb{Z}^d$ of edges and sites, and Russo's formula applies to functions on any product space.

Proof. We begin with the Burgers' inequality. As we will see, our work will be useful in the proof of the second inequality. We claim that

$$\frac{\partial M}{\partial \beta} = \sum_{e \in E} J_e \mathbb{P}[\omega_e = 0, e \text{ pivotal for } 0 \leftrightarrow \mathcal{G}].$$

To see this, recall that Russo's formula (Proposition 2.3) states that

$$\frac{\partial}{\partial p_e} \mathbb{P}[E] = \mathbb{P}[e \text{ pivotal for } E] = \frac{1}{1-p_e} \mathbb{P}[\omega_e = 0, e \text{ pivotal for } E].$$

Note that $\frac{\partial p_e}{\partial \beta} = J_e e^{-\beta J_e}$ and $1 - p_e = e^{-\beta J_e}$. Therefore by the chain rule,

$$\frac{\partial}{\partial \beta} \mathbb{P}[E] = \sum_{e \in E} \frac{\partial p_e}{\partial \beta} \cdot \frac{\partial}{\partial p_e} \mathbb{P}[E] = \sum_{e \in E} J_e \mathbb{P}[\omega_e = 0, e \text{ pivotal for } E],$$

as desired.

Next, we interpret the event in the sum geometrically. If $e = xy$, then the event that e is pivotal for E is the event that (1) any connection between 0 and $\mathcal{G}$ must pass through the edge e, and (2) if bond e is opened, then there exists a path from 0 to $\mathcal{G}$. When we set $\omega_e = 0$, the connected cluster of the origin is disconnected from $\mathcal{G}$, but we retain a path from y to $\mathcal{G}$. Therefore, the event E may be rewritten as $\{x \in C_0, C_0 \cap \mathcal{G} = \emptyset, y \leftrightarrow \mathcal{G}\}$.

We now perform the key step in the argument: conditioning on the cluster $C_0 = A$ for a given set of sites A. Conditioning gives us additional geometric control over the events in question, and this technique will reappear throughout this report. Our sum becomes

$$\sum_{e=xy}\sum_{A,x\in A} J_e\mathbb{P}[C_0 = A, A\cap\mathcal{G} = \emptyset, y \overset{\Lambda_L\setminus A}{\leftrightarrow} \mathcal{G}].$$

Note that the event $\{y \overset{\Lambda_L\setminus A}{\leftrightarrow} \mathcal{G}\}$ that y is connected to $\mathcal{G}$ in the complement of A is clearly independent of $\{C_0 = A, A\cap G = \emptyset\}$, since the states of the bonds in $\Lambda_L \setminus A$ are independent of bonds and sites in A. We may therefore split the probability and apply the trivial inequality $\mathbb{P}[y \overset{\Lambda_L\setminus A}{\leftrightarrow} \mathcal{G}] \le \mathbb{P}[y \leftrightarrow \mathcal{G}] = M$ to find an upper bound of

$$M\sum_{e=xy}\sum_{A,x\in A} J_e\mathbb{P}[C_0 = A, A\cap\mathcal{G} = \emptyset].$$

Removing the conditioning, this simplifies to

$$M\sum_{e=xy} J_e\mathbb{P}[x\in C_0, C_0\cap\mathcal{G} = \emptyset].$$

We now separate into a sum over x and y to find

$$M\sum_{x\in\mathbb{Z}^d}\mathbb{P}[x\in C_0, C_0\cap\mathcal{G} = \emptyset]\cdot\sum_y J_{xy}.$$

We recognize the sum on x as $\chi(\beta, h)$, the expected size of $|C_0|$ conditional on $C_0\cap\mathcal{G} = \emptyset$, and by translation invariance the sum on y equals $|J|$. Thus our bound reads $\frac{\partial M}{\partial\beta} \le |J|M\chi$, establishing the first partial differential inequality.

We continue by proving a weak version of the second inequality. Specifically, we show that

$$M \le \tilde{h}\chi + M^2 + \beta QM\frac{\partial M}{\partial\beta},$$

where we use the notation $\tilde{x} = e^x - 1$, and where $Q = \sup_e \frac{\widetilde{\beta J_e}}{\beta J_e}$. One can show that this implies the second inequality with a little extra effort; for details, see [AB87].

We begin with the left-hand side of our inequality, the probability M that 0 is connected to the green set $\mathcal{G}$. We decompose the event $\{0 \leftrightarrow \mathcal{G}\}$ into three cases:

- Case 1: there is *exactly* one green site in C_0;
- Case 2: 0 is doubly-connected to $\mathcal{G}$; or
- Case 3: there are at least two green sites in C_0 but 0 is *not* doubly-connected to $\mathcal{G}$.

Here, 0 is doubly-connected to $\mathcal{G}$ if there are two paths between 0 and $\mathcal{G}$ that only intersect at the starting site 0. Note that Case 1 is not the same as a single path from 0 to g. For a visual depiction, see Figure 2.

We bound the probability of each case in turn. Cases 1 and 2 are relatively straightforward. Case 1 is equivalent to the existence of a site x for which the variable η_x is pivotal for the event

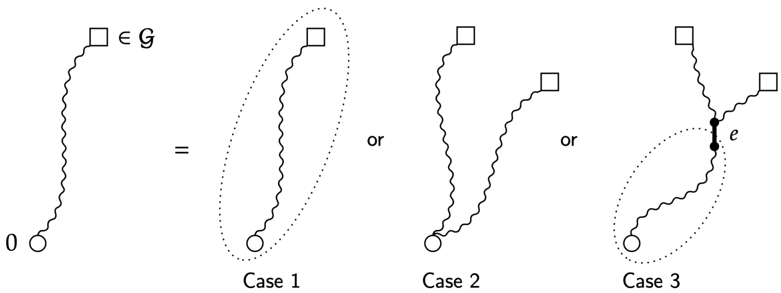

Figure 2: The three cases in the proof of Theorem 3.1; reproduced from Figure 2 of [AB87].

$\{0 \leftrightarrow \mathcal{G}\}$ and for which $\eta_x = 1$. By Russo's formula, this is given by

$$\begin{aligned}\sum_{x\in\mathbb{Z}^d} \mathbb{P}[\eta_x \text{ pivotal for } \{0 \leftrightarrow \mathcal{G}\}, \eta_x = 1] &= (1-e^{-h}) \sum_{x\in\mathbb{Z}^d} \mathbb{P}[\eta_x \text{ pivotal for } \{0 \leftrightarrow \mathcal{G}\}] \\ &= (1-e^{-h}) \frac{\partial M}{\partial(1-e^{-h})} \\ &= (1-e^{-h}) \frac{1}{\partial(1-e^{-h})/\partial h} \frac{\partial M}{\partial h} \\ &= (e^h - 1)\frac{\partial M}{\partial h},\end{aligned}$$

yielding the first term in the upper bound. In Case 2, the two paths are disjointly-realized. Since each path occurs with probability M, the BK inequality (Proposition 2.2) implies that the probability that they are disjointly-realized is bounded by M^2.

Case 3 is the most tricky, but as we will see, the conditioning arguments we utilized in establishing the first partial differential inequality will serve us well. The geometric picture to bear in mind is the following: since 0 is not doubly-connected to $\mathcal{G}$, there exists an ordered set of pivotal edges for the event $\{0 \leftrightarrow \mathcal{G}\}$ such that every connection must pass through these edges. Between consecutive pivotal edges, we have "bubbles" consisting of at least two edge-disjoint paths. Beyond the last pivotal edge xy, there must exist at least two disjoint paths from y to $\mathcal{G}$. The BK inequality will allow us to "peel off" one of these disjoint paths in such a way that the resulting event is identical to one encountered in the first partial differential inequality.

Let us make this precise. Let F denote the event in question. Let $e = xy$ be the last bond common to all paths from 0 to $\mathcal{G}$, so that any such path contains xy. Then y must be doubly-connected to $\mathcal{G}$ in the exterior of $C_0^{\sim e}$, the set of sites connected to 0 through paths that do not include $e = xy$. By our choice of xy, the set $C_0^{\sim e}$ cannot intersect $\mathcal{G}$. As a result, we have the upper bound

$$\mathbb{P}[F] \le \sum_{e=xy} \mathbb{P}[\omega_e = 1, x \in C_0, C_0^{\sim e} \cap \mathcal{G} = \emptyset, y \text{ doubly-connected to } \mathcal{G} \text{ in } \Lambda_L \setminus C_0^{\sim e}].$$

We now apply the key conditioning: in order to manipulate the sum, we condition on the set $C_0^{\sim e} = A$. Note that the bond random variables in $\Lambda_L \setminus A$ remain independent after conditioning. The bound becomes

$$\sum_{e=xy} \sum_{A, x\in A} \mathbb{P}[\omega_e = 1] \cdot \mathbb{P}[C_0^{\sim e} = A, A \cap \mathcal{G} = \emptyset] \cdot \mathbb{P}[y \text{ doubly-connected to } \mathcal{G} \text{ in } \Lambda_L \setminus A].$$

We now "peel off" one of the two connections from y to $\mathcal{G}$. Specifically, note that by the BK inequality, we have $\mathbb{P}[y \text{ doubly-connected to } \mathcal{G} \text{ in } \Lambda_L \setminus A] \le \mathbb{P}[y \overset{\Lambda_L \setminus A}{\leftrightarrow} \mathcal{G}]^2$, which is clearly at most $M \cdot \mathbb{P}[y \overset{\Lambda_L \setminus A}{\leftrightarrow} \mathcal{G}]$. Therefore our bound becomes

$$M \sum_{e=xy} \sum_{A, x \in A} \mathbb{P}[\omega_e = 1] \cdot \mathbb{P}[C_0^{\sim e} = A, A \cap \mathcal{G} = \emptyset] \cdot \mathbb{P}[y \overset{\Lambda_L \setminus A}{\leftrightarrow} \mathcal{G}].$$

Since $\mathbb{P}[\omega_e = 1] = 1 - e^{-\beta J_e}$, we have $\mathbb{P}[\omega_e = 1] = \widetilde{\beta J_e} \cdot \mathbb{P}[\omega_e = 0]$, which by definition is at most $\beta Q J_e \mathbb{P}[\omega_e = 0]$. Also, it is clear that $\mathbb{P}[y \overset{\Lambda_L \setminus A}{\leftrightarrow} \mathcal{G}] \leq \mathbb{P}[y \leftrightarrow \mathcal{G}]$. Plugging these bounds into the above and removing the conditioning on A, we find

$$\mathbb{P}[F] \leq \beta Q M \sum_{e=xy} J_e \mathbb{P}[\omega_e = 0, x \in C_0, C_0^{\sim e} \cap \mathcal{G} = \emptyset, y \leftrightarrow \mathcal{G}].$$

Note that the event in the sum states that e is closed and that e is pivotal for the event $0 \leftrightarrow \mathcal{G}$. However, this very expression appeared in the proof of the Burgers' inequality! As above, Russo's formula and the chain rule imply that the sum is precisely $\frac{\partial M}{\partial \beta}$. It follows that $\mathbb{P}[F] \leq \beta Q M \frac{\partial M}{\partial \beta}$, as desired. Summing Cases 1, 2, and 3, we find the ϕ^3 inequality. □

It turns out that these inequalities hold for the Ising model as well. In fact, using the random current and random walk representations of the Ising model, Aizenman–Barsky–Fernandez obtain an improvement on the second inequality [ABF87].

Since Bernoulli percolation and the Ising model can be identified with the random cluster model with weight $q = 1$ and $q = 2$ respectively, it is natural to ask whether any of the above can be generalized to to the random cluster measure. Indeed, we may define the magnetization $M(\beta, h) = \mathbb{E}[1 - e^{-h|C_0|}]$ for the random cluster measure on $\mathbb{Z}^d$ as above. We conjecture the following.

Conjecture 3.2. For $d \geq 2$, the partial differential inequalities in Theorem 3.1 hold for the random cluster model on $\mathbb{Z}^d$ with boundary condition $\xi \in \{f, w\}$ and $q \in [1, 2]$.

Note that existing proofs of the partial differential inequalities for percolation and the Ising model are quite different, and it is desirable for the proof of Conjecture 3.2 to apply for all $q \in [1, 2]$. Below, we present a recent proof of a weak version of the ϕ^3 inequality that applies for all $q \geq 1$ due to Aizenman–El Bahri. However, there does not exist progress on the Burgers' inequality for general q.

§4 Decision tree and decision forest inequalities

We now shift our attention to the random cluster model. As we have seen, the Aizenman–Barsky proof of the sharpness of percolation uses the independence of the model in a fundamental way. For instance, the BK inequality is used in the proof of both the Burgers' and ϕ^3 inequalities, but there is no generalization of BK to the random cluster model [Gri06]. Similar issues prevent the proof of sharpness of the Ising model from generalizing. In order to circumvent these obstacles, a new input was needed. Surprisingly, the key tool came from the theory of boolean functions and decision tree inequalities in theoretical computer science.

§4.1 One-function DRT inequality

Decision trees are models of computation in computer science. Informally, a decision tree for a function $f : \{0,1\}^E \to [0,1]$ processes the input vector from left-to-right and terminates once the value of the function is uniquely determined. A natural question is how the depth or complexity of the decision tree is related to the function f. Intuitively, the larger the variance of the function f, the more complex the tree must be. The OSSS inequality formalized this notion in the case that the input vector is sampled from a product space [OSSS05]. The OSSS inequality states that a certain weighted sum of the influences of the edges is lower bounded by the variance of f. This has a natural connection to percolation theory: think of $\{0,1\}^E$ as the bond configuration, f as the indicator function of an event of interest, and let our decision tree be formed by "exploring" the edges of the configuration from a given starting site. For the sake of brevity, we directly present a generalization of the OSSS inequality to monotonic measures due Duminil-Copin–Raoufi–Tassion, which we dub the DRT inequality. In this section, we follow the paper [DCRT18], with some statements generalized according to [Hut20].

We begin with generalities on decision trees, following the conventions of [DCRT18] and [Hut20]. Given an n-tuple $(z_1, \ldots, z_n)$, let $z_{[t]}$ denote its first t elements. Consider a countable set of edges E. Fix an edge $e_1 \in E$. A decision tree is a function $T : \{0,1\}^E \to E^{\mathbb{N}}$ sending configurations ω to sequences of edges $(T_1, T_2, \ldots) = (e_1, e_2, \ldots)$ that start with e_1, along with a collection of decision rules $\varphi_t : E^{t-1} \times \{0,1\}^{t-1} \to E$ for all $t \geq 2$ satisfying the following property: for each $t \geq 2$, we have $e_t = \varphi_t(e_{[t-1]}, \omega_{[t-1]})$, where $\omega_{[t-1]}$ denotes $(\omega_{e_1}, \ldots, \omega_{e_{t-1}})$. Stated simply, the decision tree begins by querying the state ω_{e_1} of edge e_1, and at step $t \geq 2$, it uses the decision rule φ_t to decide which edge e_t to query next. For a visual depiction, see Figure 3.

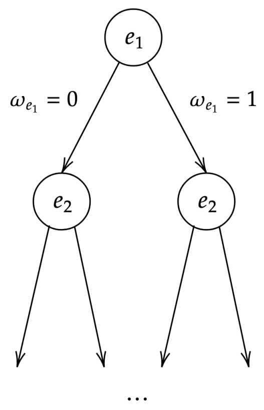

Figure 3: A visual depiction of a decision tree querying a configuration ω.

Now suppose we are given a measure μ on $\{0,1\}^E$ and let ω be a random configuration distributed as μ. In this case, the sequence $(e_1, e_2, \ldots)$ of edges visited by our decision tree T consists of random variables. For $t \geq 1$, let $\mathcal{F}_t(T) = \sigma(\{e_1, \ldots, e_t\})$ denote the σ-algebra generated by $e_1, \ldots, e_t$. Let $\mathcal{F}(T) = \sigma(\bigcup_{t \geq 1} \mathcal{F}_t(T))$ denote the smallest σ-algebra containing $\mathcal{F}_t(T)$ for all $t \geq 1$. Given a measurable function $f : \{0,1\}^E \to \mathbb{R}$, we say that T computes f if $f(\omega)$ is measurable with respect to the μ-completion of $\mathcal{F}(T)$. If f is μ-integrable, then since $\mathcal{F}_t(T) \nearrow \mathcal{F}(T)$, the martingale convergence theorem implies that as $t \to \infty$, the Doob martingale $\mu[f(\omega)|\mathcal{F}_t(T)]$ converges to $f(\omega)$ μ-almost surely [Dur19]. The revealment of an edge e is defined as the probability that e is queried by the tree: $\delta_e(T, \mu) = \mu[\exists t \geq 1 \text{ s.t. } e_t = e]$. When E is finite, we define the stopping time τ of T as the minimum number of steps until the decision tree is guaranteed to have computed the value of f: $\tau(\omega) = \min\{t \geq 1 : \forall \omega' \in \{0,1\}^E, \omega'_{e_{[t]}} = \omega_{e_{[t]}} \implies f(\omega') = f(\omega)\}$. When E is finite, $\delta_e(T, \mu) = \mu[\exists t \leq \tau(\omega) \text{ s.t. } e_t = e]$.

Theorem 4.1 (One-function DRT inequality)

Let E be a finite or countably infinite set and let μ be a monotonic measure on $\{0,1\}^E$. Then for every measurable, μ-integrable, increasing function $f : \{0,1\}^E \to [0,1]$ and every decision tree T computing f we have that

$$\mathrm{Var}_\mu f \leq \sum_{e \in E} \delta_e(T,\mu)\mathrm{Cov}_\mu[f,\omega_e].$$

Proof. We begin by proving the result for E finite; at the end of the proof, we will bootstrap the result to countable E. The first step in the proof is a clever method of sampling the configuration ω from μ using independent and identically distributed (iid) uniform random variables. Let U be a random vector whose coordinates are n iid Unif$[0,1]$ variables, and let $e_{[n]} = (e_1,\ldots,e_n)$ be an n-tuple of edges. For $t \geq 1$, we inductively construct a random configuration X on $e_{[n]}$ according to

$$X_{e_t} = \begin{cases} 1 & \text{if } U_t \geq \mu[\omega_{e_t} = 0 | \omega_{e_{[t-1]}} = x_{e_{[t-1]}}] \\ 0 & \text{else} \end{cases}$$

In other words, at step t, we sample U_t from Unif$[0,1]$, and if U_t is larger than the conditional probability that $\omega_{e_t} = 0$ given the states of the previous edges $e_{[t-1]}$, we set X_{e_t} to 1. The output X of this procedure is denoted $F_e(U)$. The following lemma formalizes the intuition that X is distributed as μ.

Lemma 4.2

Suppose that U is a random vector whose coordinates are n iid Unif$[0,1]$ random variables, and suppose that $e_{[n]} = (e_1,\ldots,e_n)$ is a random n-tuple of edges. Suppose that for all $1 \leq t \leq n$, U_t is independent of the variables $(e_{[t]}, U_{[t-1]})$. Then $X = F_e(U)$ is distributed as μ.

Proof. See Lemma 2.1 in [DCRT18]. □

We continue with the proof of Theorem 4.1. Let U be a sequence of n iid Unif$[0,1]$ random variables. We use our decision tree T and the sequence U to construct an n-tuple $e_{[n]}$ of edges and a configuration X on these edges. Our construction of $e_{[n]}$ and X is inductive. Let e_1 be the first edge visited by T. For $t \geq 1$, let

$$X_{e_t} = \begin{cases} 1 & \text{if } U_t \geq \mu[\omega_{e_t} = 0 | \omega_{e_{[t-1]}} = x_{e_{[t-1]}}] \\ 0 & \text{else} \end{cases}$$

and for $t > 1$, let $e_t = \varphi_t(e_{[t-1]}, X_{e_{[t-1]}})$. We see that $e_{[n]}$ is the order in which the decision tree T traverses the edges, and X is the configuration obtained from sampling U. Let τ denote the stopping time associated to T for the configuration X.

Consider another sequence V of n iid Unif$[0,1]$ random variables, independent of U. We utilize a Lindeberg swapping argument. For $0 \leq t \leq n$, define the random vector

$$W^t = (V_1,\ldots,V_t,U_{t+1},\ldots,U_\tau,V_{\tau+1},\ldots,V_n)$$

obtained by swapping entries $t+1$ through τ of V with those of U. Let $Y^t = F_e(W^t)$ be the configuration obtained by the decision tree when T follows the prescribed order $e_{[n]}$ and draws uniform random variables according to the entries of W^t. Note that when $t \geq \tau$, we have $Y^t = Y^n$. Also, when $t = 0$, the entries of the random vectors W^0 and U agree up to the stopping time τ, so that Y^0 and X agree up to τ and $f(Y^0) = f(X)$.

By Lemma 4.2, X and Y^n are distributed as μ. Note that X is measurable with respect to U, while Y^n is independent of U. Below, we let $\mathbb{E}$ denote the expectation with respect to the joint measure $\mu \otimes \mu$ of X and Y^n.

In order to apply the Lindeberg method, we must rewrite the left-hand side of the desired inequality as a difference of expectations. Note that

$$\begin{aligned}
\mathrm{Var} f &\leq \frac{1}{2}\mu[|f - \mu[f]|] \\
&= \frac{1}{2}\mathbb{E}[|\mathbb{E}[f(X)|U] - \mathbb{E}[f(Y^n)|U]|] \\
&= \frac{1}{2}\mathbb{E}[|\mathbb{E}[f(X) - f(Y^n)|U]|] \\
&\leq \frac{1}{2}\mathbb{E}[|f(X) - f(Y^n)||U],
\end{aligned}$$

where in the second step we use the fact that X is measurable with respect to U and Y^n is independent of U, and where in the last step we apply Jensen's inequality. Since $f(Y^0) = f(X)$, we in fact have $\mathrm{Var} f \leq \frac{1}{2}\mathbb{E}[|f(Y^0) - f(Y^n)|]$. We now apply the triangle inequality to find

$$\mathrm{Var} f \leq \frac{1}{2}\sum_{t=1}^{n} \mathbb{E}[|f(Y^t) - f(Y^{t-1})|].$$

Since $Y^t = Y$ for all $t \geq \tau$, all terms with $t \geq \tau$ are zero. Our bound therefore becomes

$$\mathrm{Var} f \leq \frac{1}{2}\sum_{t=1}^{n} \mathbb{E}[|f(Y^t) - f(Y^{t-1})| \cdot \mathbf{1}_{t\leq\tau}].$$

We now add conditioning on the random vector $U_{[t-1]}$ to find

$$\mathrm{Var} f \leq \frac{1}{2}\sum_{t=1}^{n} \mathbb{E}[\mathbb{E}[|f(Y^t) - f(Y^{t-1})||U_{[t-1]}] \cdot \mathbf{1}_{t\leq\tau}].$$

Finally, we condition on the edge e_t and find inequality (†):

$$\mathrm{Var} f \leq \frac{1}{2}\sum_{e\in E}\sum_{t=1}^{n} \mathbb{E}[\mathbb{E}[|f(Y^t) - f(Y^{t-1})||U_{[t-1]}] \cdot \mathbf{1}_{t\leq\tau, e_t = e}].$$

It suffices to control the conditional expectations $\mathbb{E}[|f(Y^t) - f(Y^{t-1})||U_{[t-1]}]$ for all $1 \leq t \leq n$. In order to proceed, we need the following result.

Lemma 4.3

If g is measurable and $t \leq n$, then $\mathbb{E}[g(Y^t)|U_{[t]}] = \mu[g(\omega)]$.

Proof. Since none of the variables U_j appear in the coordinates of W^t for any $j \leq t$, it follows that conditioned on $U_{[t]}$, the coordinates of W^t are iid Unif$[0, 1]$. Therefore conditioned on $U_{[t]}$, W_i^t is independent of $U_1, \ldots, U_{i-1}$ for all $1 \leq i \leq n$. We also claim that conditioned on $U_{[t]}$,

W_i^t is independent of $e_{[i]}$ for all $1 \leq i \leq n$. It suffices to check the claim for $t \leq i \leq \tau$. By construction, the set of edges $e_{[i]}$ is determined by the variables $U_{[i-1]}$, which as shown above are conditionally independent of $W_i^t = U_i$. We may therefore apply Lemma 4.2 to the variables $e_{[n]}$ and W^t conditioned on $U_{[t]}$ to find that $Y^t = F_e(W^t)$ is distributed as μ. The result follows. $\square$

We now claim that $\mathbb{E}[|f(Y^t) - f(Y^{t-1})||U_{[t-1]}] \leq 2\mathrm{Cov}[f, \omega_e]$. To see this, we begin by rewriting the difference $|f(Y^t) - f(Y^{t-1})|$. Note that since W^t and W^{t-1} agree except at index t, the configurations Y^t and Y^{t-1} agree except possibly at edge e. (This is because we follow the fixed ordering $e_{[n]}$ in order to construct Y^t and Y^{t-1}.) If $Y_e^t = Y_e^{t-1}$, then $Y^t = Y^{t-1}$ and our difference is zero, so that the claim is trivial. Otherwise, since f is increasing, we have

$$|f(Y^t)-f(Y^{t-1})| = (f(Y^t)-f(Y^{t-1}))(Y_e^t-Y_e^{t-1}) = f(Y^{t-1})Y_e^{t-1}+f(Y^t)Y_e^t-f(Y^{t-1})Y_e^t-f(Y^t)Y_e^{t-1}.$$

We begin with the positive terms. By Lemma 4.3 applied to $g(\omega) = f(\omega)\omega_e$, we have $\mathbb{E}[f(Y^{t-1})Y_e^{t-1}|U_{[t-1]}] = \mu[f(\omega)\omega_e]$ and $\mathbb{E}[f(Y^t)Y_e^t|U_{[t-1]}] = \mathbb{E}[\mathbb{E}[f(Y^t)Y_e^t|U_{[t]}]|U_{[t-1]}] = \mu[f(\omega)\omega_e]$. Next, we handle the negative terms. Consider the conditional expectation $\mathbb{E}[f(Y^{t-1})Y_e^t|U_{[n]}]$, where we condition on the entire vector U. Given s, the configuration Y^s is increasing in V, because if $Y_{e_t}^s = 1$ for some t, then if V_t is increased, it remains the case that $Y_{e_t}^s = 1$. In particular, the entry Y_e^t is increasing. Since f is increasing, $f(Y^{t-1})$ is also increasing in V. By FKG with respect to the variables V, we have the lower bound

$$\mathbb{E}[f(Y^{t-1})Y_e^t|U_{[n]}] \geq \mathbb{E}[f(Y^{t-1})|U_{[n]}]\mathbb{E}[Y_e^t|U_{[n]}].$$

We now take conditional expectations with respect to $U_{[t-1]}$. Since Y_e^t is determined at step t, it depends only on V and $U_{[t-1]}$. Hence the conditional expectation $\mathbb{E}[Y_e^t|U_{[n]}]$ is measurable with respect to $U_{[t-1]}$. By Lemma 4.2, the bound reduces to

$$\mathbb{E}[f(Y^{t-1})Y_e^t|U_{[t-1]}] \geq \mathbb{E}[f(Y^{t-1})|U_{[t-1]}]\mathbb{E}[Y_e^t|U_{[n]}] = \mu[f(\omega)]\mu[\omega_e].$$

Similarly, the other negative term $\mathbb{E}[f(Y^{t-1})Y_e^t|U_{[t-1]}]$ is bounded below by $\mu[f(\omega)]\mu[\omega_e]$. Putting these inequalities together, we find that

$$\mathbb{E}[|f(Y^t) - f(Y^{t-1})||U_{[t-1]}] \leq 2(\mu[f(\omega)\omega_e] - \mu[f(\omega)]\mu[\omega_e]) = 2\mathrm{Cov}[f, \omega_e],$$

completing the proof of the claim.

Finally, plugging this estimate into (†), we find

$$\begin{aligned}\mathrm{Var} f &\leq \sum_{e\in E} \mathrm{Cov}[f, \omega_e] \sum_{t=1}^{n} \mu[t \leq \tau, e_t = e] \\ &= \sum_{e\in E} \delta_e(T, \mu)\mathrm{Cov}[f, \omega_e],\end{aligned}$$

where in the last step we used the fact that $\sum_{t=1}^n \mu[t \leq \tau, e_t = e]$ is the probability that T queries e by the stopping time τ. This implies the result.

We now extend the above inequality to countable sets E, following [Hut20]. Given $t \geq 1$, consider the conditional expectation $f_t = \mu[f|\mathcal{F}_t(T)]$. By definition, f_t is computed by the decision tree T by time t. Let E_t denote the set of edges that T can visit up to time t, with $E_t \nearrow E$. Since T is deterministic, E_t is finite with $|E_t| \leq 2^t$. Let T_t denote the decision tree T restricted to the set of edges E_t. Since E_t is finite, the DRT inequality implies that

$$\mathrm{Var} f_t \leq \sum_{e\in E_t} \delta_e(T_t, \mu) \cdot \mathrm{Cov}[f_t, \omega_e].$$

Since f_t forms a martingale, the martingale convergence theorem implies that $\mathrm{Var} f_t \to \mathrm{Var} f$, $\mathrm{Cov}[f_t, \omega_e] \to \mathrm{Cov}[f, \omega_e]$, and $\delta_e(T_t, \mu) \to \delta_e(T, \mu)$ as $t \to \infty$. This implies the desired inequality for f. $\square$

§4.2 Two-function DRT inequality

By following the methods of the original OSSS paper, we can provide a "two-function" version of the DRT inequality (Theorem 4.1). In this section, we follow the papers [Hut20] and [OSSS05].

Let $\mathrm{CoVr}_\mu(f,g) = (\mu \otimes \mu)[|f(\omega_1) - g(\omega_2)|] - \mu[|f(\omega_1) - g(\omega_1)|]$, where (ω_1, ω_2) is sampled from the product measure $\mu \otimes \mu$. We can think of CoVr as a generalization of the covariance of two functions. Indeed, note that if f and g take values in $\{0,1\}$, then

$$\begin{aligned}\mathrm{Cov}(f,g) &= (\mu \otimes \mu)[|f(\omega_1) - g(\omega_2)|^2] - \mu[|f(\omega_1) - g(\omega_1)|^2] \\ &= (\mu \otimes \mu)[f(\omega_1)^2 + g(\omega_2)^2 - 2f(\omega_1)g(\omega_2)] - \mu[f(\omega_1)^2 + g(\omega_1)^2 - 2f(\omega_1)g(\omega_1)] \\ &= 2\mu[f(\omega_1)g(\omega_1)] - 2\mu[f(\omega_1)]\mu[g(\omega_2)] \\ &= 2\mathrm{Cov}[f,g].\end{aligned}$$

Theorem 4.4 (Two-function DRT inequality)
Let E be a finite or countably infinite set and let μ be a monotonic measure on $\{0,1\}^E$. Then for every pair of measurable, μ-integrable functions $f, g : \{0,1\}^E \to [0,1]$ with f increasing and every decision tree T computing g, we have that

$$\frac{1}{2}|\mathrm{CoVr}_\mu(f,g)| \le \sum_{e \in E} \delta_e(T,\mu)\mathrm{Cov}_\mu[f,\omega_e].$$

Proof. We follow Section 3.3 of [OSSS05]. As it turns out, the proof is simply a reduction to the one-function proof. Let configurations X and Y be independent samples from the measure μ. As in the proof of the single-function inequality, we utilize the configurations Y^t for $0 \le t \le n$ obtained by the Lindeberg swapping method. Recall that $Y^n = Y$. Also, recall that Y^0 agrees with X on the edges visited by T up to the stopping time τ when the input configuration is X. Since T computes g, this implies that $g(Y^0) = g(X)$. With this in place, we begin by rewriting $\mathrm{CoVr}(f,g)$:

$$\begin{aligned}\mathrm{CoVr}(f,g) &= (\mu \otimes \mu)[|g(X) - f(Y)|] - \mu[|g(X) - f(X)|] \\ &= (\mu \otimes \mu)[|g(Y^0) - f(Y^n)|] - \mu[|g(Y^0) - f(Y^0)|].\end{aligned}$$

In the first term, we used the equality $g(X) = g(Y^0)$, and in the second term, we used the fact that Y^0 is distributed as μ. We now apply the triangle inequality to find

$$\begin{aligned}\mathrm{CoVr}(f,g) &\le (\mu \otimes \mu)[|g(Y^0) - f(Y^0)| + |f(Y^0) - f(Y^n)|] - \mu[|g(Y^0) - f(Y^0)|] \\ &= (\mu \otimes \mu)[|f(Y^0) - f(Y^n)|].\end{aligned}$$

The same bound holds for $-\mathrm{CoVr}(f,g)$; simply write

$$\begin{aligned}-\mathrm{CoVr}(f,g) &= (\mu \otimes \mu)[|g(Y^0) - f(Y^0)|] - (\mu \otimes \mu)[|g(Y^0) - f(Y^n)|] \\ &\le (\mu \otimes \mu)[|g(Y^0) - f(Y^n)| + |f(Y^0) - f(Y^n)|] - (\mu \otimes \mu)[|g(Y^0) - f(Y^n)|] \\ &= (\mu \otimes \mu)[|f(Y^0) - f(Y^n)|].\end{aligned}$$

It therefore suffices to bound $(\mu \otimes \mu)[|f(Y^0) - f(Y^n)|]$. But this can be done exactly as in the proof of Theorem 4.1, and the result follows. □

§4.3 Decision forest inequality

Finally, we present a generalization to so-called decision forests. In this section, we follow the paper [Hut20].

Suppose E is countable. A decision forest F is simply a countable collection of decision trees $\{T_i\}_{i=1}^{\infty}$ on $\{0,1\}^E$. Define the σ-algebra $\mathcal{F}(F) = \sigma(\{\mathcal{F}(T^i)\}_{i=1}^{\infty})$. Given a measure μ on $\{0,1\}^E$ and a measurable function $f : \{0,1\}^E \to \mathbb{R}$, the forest F computes f if f is measurable with respect to the μ-completion of $\mathcal{F}(F)$. The revealment of an edge e is defined as the probability that e is queried by some tree: $\delta_e(F,\mu) = \mu[\exists i \in \mathbb{N}, \exists t \geq 1 \text{ s.t. } T_t^i(\omega) = e]$.

Theorem 4.5 (Decision forest inequality)
Let E be a finite or countably infinite set and let μ be a monotonic measure on $\{0,1\}^E$. Then for every pair of measurable, μ-integrable functions $f, g : \{0,1\}^E \to [0,1]$ with f increasing and every decision forest F computing g, we have that

$$\frac{1}{2}|\mathrm{CoVr}_\mu(f,g)| \leq \sum_{e\in E} \delta_e(F,\mu)\mathrm{Cov}_\mu[f,\omega_e].$$

Proof. We convert the decision forest F to a decision tree T in the following manner. Suppose that $F = \{T^i\}_{i=1}^{\infty}$. Let $\{p_i\}_{i=1}^{\infty}$ denote the prime numbers in increasing order. At each time $t \in \mathbb{N}$, if $t = p_i^j$, then T queries tree T^i; else, T queries edge e_1. The edges visited by T are the same as those visited by F, so $\delta_e(T,\mu) = \delta_e(F,\mu)$ for all $e \in E$. The result follows from Theorem 4.4. ◻

§5 Applications to the random cluster model

In this section, we present three applications of the decision tree and decision forest inequalities derived above to the random cluster model. These include a differential inequality on the radius of the cluster of the origin, a differential inequality for the volume of the cluster of the origin, and a weak version of the ϕ^3 partial differential inequality.

§5.1 Application 1: differential inequality for radius

Our first application of Theorem 4.1 is a differential inequality for the radius of the cluster of the origin. As a corollary, we obtain the sharpness of the phase transition of the random cluster model. In this section, we follow the paper [DCRT18].

Given $n \geq 0$, let the box $\Lambda_n \subseteq \mathbb{Z}^d$ denote the set $[-n,n]^d \cap \mathbb{Z}^d$. Given a site x, let $\Lambda_n(x)$ denote Λ_n translated by the vector x.

Theorem 5.1 (Radius tail bound)

Fix $d \geq 2$, $n \geq 1$, $q \geq 1$, $\beta_0 \geq 0$, and $\beta < \beta_0$. For $k \geq 0$, let $\mu_k = \phi^w_{\Lambda_{2k},\beta,q}$ denote a random cluster measure on Λ_{2k} with translation invariant coupling constants, let $\theta_k(\beta) = \mu_k[0 \leftrightarrow \partial\Lambda_k]$, and let $S_n = \sum_{k=0}^{n-1} \theta_k$. Then there exists a constant $c > 0$ such that

$$\theta_n'(\beta) \geq c\frac{n}{S_n}\theta_n(\beta).$$

In order to establish the above, we use the following corollary of Theorem 4.1.

Lemma 5.2

Consider a finite connected subgraph Λ of $\mathbb{Z}^d$ containing 0. For any monotonic measure μ on $\{0,1\}^E$ and any $n \geq 1$, we have

$$\sum_{e\in E} \mathrm{Cov}_\mu[\mathbf{1}_{0\leftrightarrow\partial\Lambda_n}, \omega_e] \geq \frac{n}{4\max_{x\in\Lambda_n}\sum_{k=0}^{n-1}\mu[x \leftrightarrow \partial\Lambda_k(x)]} \cdot \mu[0 \leftrightarrow \partial\Lambda_n](1 - \mu[0 \leftrightarrow \partial\Lambda_n]).$$

Proof. The key insight is to apply Theorem 4.1 to a sequence of decision trees T_k indexed by $k \in [1, n]$, where T_k determines $\mathbf{1}_{0\leftrightarrow\partial\Lambda_n}$ starting from the set $\partial\Lambda_k$. Averaging the resulting inequalities yields a low average revealment for any fixed edge.

Given k, decision tree T_k operates as follows. At step t, let V_t denote the set of vertices that have been determined to be connected to $\partial\Lambda_k$, with $V_0 = \partial\Lambda_k$. Let F_t denote the set of discovered edges, with $F_0 = \emptyset$. Let V and F denote the set of vertices of Λ_n and edges of Λ_n, respectively.

We fix an ordering of F. At step t, we perform one of two operations.

- If there exists an edge $e = xy \in F$ with $e \notin F_t$, $x \in V_t$, and $y \notin V_t$, then we explore the smallest such edge by setting $F_{t+1} = F_t \cup \{e\}$ and

  $$V_{t+1} = \begin{cases} V_t \cup \{e\} & \omega_e = 1 \\ V_t & \text{else} \end{cases}$$

  (Note that there is no need to explore edges $e = xy$ with $x, y \in V_t$, as doing so would not change our knowledge of the connected components of $\partial\Lambda_k$.)

- If there does not exist such an e, set e_{t+1} to be the smallest edge $e \in F \setminus F_t$, set $V_{t+1} = V_t$, and set $F_{t+1} = F_t \cup \{e\}$.

  (In this case, we have completely determined the connected components of $\partial\Lambda_k$.)

Note that the stopping time τ is the last step t for which we are in the first case. Once we have determined the connected components of $\partial\Lambda_k$, we can determine whether $\{0 \leftrightarrow \partial\Lambda_n\}$ occurs; this is equivalent to the existence of a connected component of $\partial\Lambda_k$ that intersects 0 and $\partial\Lambda_n$.

Our upper bound on revealment is $\delta_e(T_k, \mu) \leq \mu[u \leftrightarrow \partial\Lambda_k] + \mu[v \leftrightarrow \partial\Lambda_k]$ for any edge $e = uv$. This is clear, because if e is discovered by time τ, then either u or v is in a connected component of $\partial\Lambda_k$, hence $u \leftrightarrow \partial\Lambda_k$ or $v \leftrightarrow \partial\Lambda_k$.

We now sum the upper bound over k. We claim that

$$\sum_{k=1}^{n} \mu[u \leftrightarrow \partial\Lambda_k] \leq \sum_{k=1}^{n} \mu[u \leftrightarrow \partial\Lambda_{|k-d(u,0)|}(u)].$$

This follows from the triangle inequality: the distance between u and a point on $\partial\Lambda_k$ is at least $|k - d(u,0)|$. Continuing, note that we may bound this sum by

$$\sum_{k=1}^{n} \mu[u \leftrightarrow \partial\Lambda_{|k-d(u,0)|}(u)] \leq 2\sum_{k=0}^{n-1} \mu[u \leftrightarrow \partial\Lambda_k(u)].$$

This is because

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

Putting the inequalities together and taking a maximum on the right-hand side, we find $(\star)$:

$$\sum_{k=1}^{n} \mu[u \leftrightarrow \partial\Lambda_k] \leq 2\max_{x\in\Lambda_n} \sum_{k=0}^{n-1} \mu[x \leftrightarrow \partial\Lambda_k(x)].$$

We now apply Theorem 4.1 to decision tree T_k for all k and average the inequalities to find

$$\frac{1}{n}\sum_{k=1}^{n} \mathrm{Var}_\mu[\mathbf{1}_{0\leftrightarrow\partial\Lambda_n}] \leq \frac{1}{n}\sum_{k=1}^{n}\sum_{e\in E} \delta_e(T_k,\mu)\mathrm{Cov}_\mu[\mathbf{1}_{0\leftrightarrow\partial\Lambda_n}, \omega_e].$$

By our bound on $\delta_e(T_k,\mu)$, we have $\sum_{k=1}^{n} \delta_e(T_k,\mu) \leq \sum_{k=1}^{n} \mu[u \leftrightarrow \partial\Lambda_k] + \sum_{k=1}^{n} \mu[v \leftrightarrow \partial\Lambda_k]$. By $(\star)$, we have

$$\sum_{k=1}^{n} \delta_e(T_k,\mu) \leq 4\max_{x\in\Lambda_n} \sum_{k=0}^{n-1} \mu[x \leftrightarrow \partial\Lambda_k(x)].$$

Plugging this bound into the right-hand side of our averaged inequality and using the fact that $\mathrm{Var}_\mu[\mathbf{1}_{0\leftrightarrow\partial\Lambda_n}] = \mu[0 \leftrightarrow \partial\Lambda_n](1 - \mu[0 \leftrightarrow \partial\Lambda_n])$, we find

$$\frac{1}{n}\mu[0 \leftrightarrow \partial\Lambda_n](1 - \mu[0 \leftrightarrow \partial\Lambda_n]) \leq 4\max_{x\in\Lambda_n} \sum_{k=0}^{n-1} \mu[x \leftrightarrow \partial\Lambda_k(x)] \cdot \sum_{e\in E} \mathrm{Cov}_\mu[\mathbf{1}_{0\leftrightarrow\partial\Lambda_n}, \omega_e].$$

Rearranging yields the desired inequality. □

We are ready to prove Theorem 5.1.

Proof. We apply Russo's formula (Proposition 2.6) and manipulate to arrive at the differential inequality. We study the connection event $\{0 \leftrightarrow \partial\Lambda_n\}$ in the finite box Λ_{2n}.

Consider Lemma 5.2 where Λ_{2n} is equipped with the measure μ_n. We begin by bounding the maximum on the right-hand side by a constant multiple of S_n. We claim that for $x \in \Lambda_n$, we have

$$\sum_{k=0}^{n-1} \mu_n[x \leftrightarrow \partial\Lambda_k(x)] \leq 2 \sum_{k \leq n/2} \mu_n[x \leftrightarrow \partial\Lambda_k(x)]$$
$$\leq 2 \sum_{k \leq n/2} \mu_k[0 \leftrightarrow \partial\Lambda_k]$$
$$\leq 2S_n$$

To see the second step, consider Figure 4. Note that $\Lambda_{2k}(x) \subseteq \Lambda_{2n}$. Therefore for any given configuration $\hat{\omega}$, the Domain Markov Property in Proposition 2.5 implies that

$$\mu_n[x \leftrightarrow \partial\Lambda_k(x) | \omega_{|\Lambda_{2n}\setminus\Lambda_{2k}(x)} = \hat{\omega}_{|\Lambda_{2n}\setminus\Lambda_{2k}(x)}] = \phi^{\xi}_{\Lambda_{2k}(x),\beta,q}[x \leftrightarrow \partial\Lambda_k(x)],$$

where ξ is the boundary condition induced by $\hat{\omega}$. By comparison of boundary conditions in Proposition 2.5, this is bounded by the probability of the event with respect to the wired random cluster measure on $\Lambda_k(x)$, namely $\phi^{w}_{\Lambda_{2k}(x),\beta,q}[x \leftrightarrow \partial\Lambda_k(x)]$. Removing the conditioning, we find an upper bound of $\phi^{w}_{\Lambda_{2k}(x),\beta,q}[x \leftrightarrow \partial\Lambda_k(x)]$. It is now clear that $\phi^{w}_{\Lambda_{2k}(x),\beta,q}[x \leftrightarrow \partial\Lambda_k(x)] = \mu_k[0 \leftrightarrow \partial\Lambda_k]$, since the wired random cluster models on $\Lambda_{2k}(x)$ and Λ_{2k} are identical up to a translation by x. The claim follows.

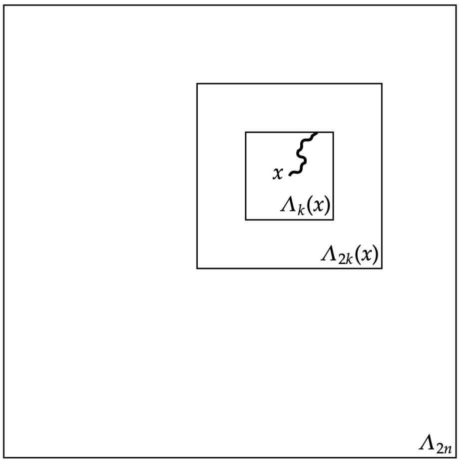

Figure 4: The boxes used in the proof of Theorem 5.1; adapted from [DCRT18].

Therefore Lemma 5.2 implies

$$\sum_{e \in E} \mathrm{Cov}_{\mu_n}[\mathbf{1}_{0 \leftrightarrow \partial\Lambda_n}, \omega_e] \geq \frac{n}{8S_n} \cdot \theta_n(\beta)(1 - \theta_n(\beta)).$$

By Russo's formula,

$$\theta_n'(\beta) = \sum_{e \in E} \frac{J_e}{e^{\beta J_e} - 1} \mathrm{Cov}_{\mu_n}[\mathbf{1}_{0 \leftrightarrow \partial\Lambda_n}, \omega_e],$$

which, since $\beta_0 > \beta$, is bounded below by

$$\min_{e \in E} \left[\frac{J_e}{e^{\beta_0 J_e} - 1} \right] \cdot \sum_{e \in E} \mathrm{Cov}_{\mu_n}[\mathbf{1}_{0 \leftrightarrow \partial\Lambda_n}, \omega_e].$$

Note that since $\theta_n(\beta_0) \leq \theta_1(\beta_0)$ and $\theta_n(\beta) \leq \theta_n(\beta_0)$, we have $1 - \theta_n(\beta) \geq 1 - \theta_1(\beta_0)$. Hence

$$\theta_n'(\beta) \geq c \frac{n}{S_n} \theta_n(\beta),$$

where $c = \frac{1-\theta_1(\beta_0)}{8} \min_{e\in E}\left[\frac{J_e}{e^{\beta_0 J_e}-1}\right]$ is a positive constant. $\square$

As stated at the beginning of this section, a corollary of Theorem 5.1 is the sharpness of the phase transition.

Theorem 5.3 (Sharpness of random cluster model)
Given the random cluster model on $\mathbb{Z}^d$ with $d \geq 2$, wired boundary condition, and $q \geq 1$:

- There exists $c > 0$ such that $\theta(\beta) \geq \beta - \beta_c$ for $\beta \geq \beta_c$ sufficiently close to β_c;
- For any $\beta < \beta_c$, there exists $c_\beta > 0$ such that $\phi^w_{\Lambda_n,\beta,q}[0 \leftrightarrow \partial\Lambda_n] \leq \exp(-c_\beta n)$.

Proof. By the purely analytic Lemma 3.1 of [DCRT18], Theorem 5.1 implies the result. See [DCRT18] for details. $\square$

Note that although Theorem 5.3 is stated for the wired boundary condition, the coincidence of the critical points of the wired and free boundary conditions implies sharpness for the free boundary condition.

§5.2 Application 2: differential inequality for volume

We now use Theorem 4.5 to derive a differential inequality for the volume $|C_0|$ of the cluster of the origin. In this section, we follow the paper [Hut20].

Theorem 5.4 (Volume tail bound)
Given $q \geq 1$ and boundary condition $\xi \in \{f, w\}$, we have

$$\max_{e\in E}\left[\frac{e^{\beta J_e}-1}{J_e}\right]\left(\frac{d}{d\beta}\right)_+ \log \phi^\xi_{\beta,q}[|C_0| \geq n] \geq \frac{1}{2}\left[\frac{n(1-e^{-\lambda})}{\lambda\sum_{m=1}^{\lceil n/\lambda\rceil}\phi^\xi_{\beta,q}[|C_0|\geq m]} - 1\right]$$

for all $\beta \geq 0$, $\lambda > 0$, and $n \geq 1$.

As in the derivation of the differential inequality for the radius, we begin with an intermediate bound.

Proposition 5.5
Let $G = (V, E)$ be a countable graph. Let μ be a monotonic measure on $\{0,1\}^E$. Fix $v \in V$, $n \geq 1$, and $\lambda > 0$. Then

$$\sum_{e\in E}\mathrm{Cov}_\mu[\mathbf{1}_{|C_v|\geq n}, \omega_e] \geq \left[\frac{(1-e^{-\lambda}) - \mu[1-e^{-\lambda|C_v|/n}]}{2\sup_{u\in V}\mu[1-e^{-\lambda|C_u|/n}]}\right]\mu[|C_v| \geq n].$$

Proof. Surprisingly, we reintroduce a technique from [AB87]: sprinkling green sites among the vertices! Let $\omega \in \{0,1\}^E$ denote the configuration of edges. We define the ghost field η to be a random variable independent of ω that independently assigns each $u \in V$ the value $\eta_u = 1$ with probability $h = 1 - e^{-\lambda/n}$ and the value $\eta_u = 0$ with probability $1 - h$. We call vertices u such that $\eta_u = 1$ the green sites, and let $\mathcal{G}$ denote the random set of green sites. Let $\mathbb{P}$ denote the joint distribution of (ω, η) and let $\mathbb{E}$ be the associated expectation.

Consider the functions $f(\omega, \eta) = \mathbf{1}_{|C_v| \geq n}$ and $g(\omega, \eta) = \mathbf{1}_{v \leftrightarrow \mathcal{G}}$. Here, $v \leftrightarrow \mathcal{G}$ is the event that v is connected to some green site by an open path of edges. We will apply Theorem 4.5 to the pair (f, g) for a certain decision forest computing g. The intuition for the choice of (f, g) is as follows: in order to understand the function f, we construct an approximation g which affords a convenient decision forest. When $f(\omega, \eta) = 1$, the expected number of green sites is at least $\frac{\lambda}{n}|C_v| \geq \lambda$, and so it is likely that $g(\omega, \eta) = 1$. On the other hand, if the cluster volume is smaller than n, we expect to see very few green sites in C_v and $g(\omega, \eta)$ is likely 0.

With this intuition established, we describe our decision forest. Fix an ordering of the edges in E. The forest F consists of trees $\{T_u : u \in V\}$. The decision tree T_u operates in the following way. First, T_u queries vertex u. If u is not green, then the process terminates. Else if u is green, then T_u explores the cluster of u by processing edges in increasing order. The exploration of each cluster is identical to the exploration of the tree T_k in the proof of Lemma 5.2. It is clear that F computes g, since $v \leftrightarrow \mathcal{G}$ if and only if v is discovered in the cluster of some green site u.

It is important to note that unlike in the previous example of a decision tree, the tree T_u takes as input a vertex u and a sequence of edges. As a result, the vertex u is associated to a revealment $\delta_u(F, \mathbb{P})$ which must be accounted for in our decision forest inequality.

Theorem 4.5 therefore reads

$$\mathrm{Cov}[f, g] \leq \sum_{u \in V} \delta_u(F, \mathbb{P}) \mathrm{Cov}[f, \eta_u] + \sum_{e \in E} \delta_e(F, \mathbb{P}) \mathrm{Cov}[f, \omega_e],$$

where covariance is taken with respect to $\mathbb{P}$. Since the random variables η and ω are independent and since f is independent of η, we have $\mathrm{Cov}[f, \eta_u] = 0$, hence the first sum of the right-hand side vanishes.

The revealment $\delta_e(F, \mathbb{P})$ can be bounded as follows. If $e = xy$ for vertices x and y, then if e is visited by the forest, e must be connected to $\mathcal{G}$. In other words, either $x \leftrightarrow \mathcal{G}$ or $y \leftrightarrow \mathcal{G}$. The probability that there is no green site in the cluster of x is exactly $\mu[(1-h)^{|C_x|}] = \mu[e^{-\lambda|C_x|/n}]$, hence $x \leftrightarrow \mathcal{G}$ occurs with probability at most $\sup_{u \in V} \mu[1 - e^{-\lambda|C_u|/n}]$. The same holds for y, so by a union bound, we have

$$\delta_e(F, \mathbb{P}) \leq 2 \sup_{u \in V} \mu[1 - e^{-\lambda|C_u|/n}].$$

Finally, we massage the left-hand side of the inequality. By definition of covariance, we have

$$\begin{aligned}
\mathrm{Cov}[f, g] &= \mathbb{E}[\mathbf{1}_{|C_v| \geq n} \mathbf{1}_{v \leftrightarrow \mathcal{G}}] - \mathbb{E}[\mathbf{1}_{|C_v| \geq n}] \mathbb{E}[\mathbf{1}_{v \leftrightarrow \mathcal{G}}] \\
&= (\mathbb{P}[v \leftrightarrow \mathcal{G} | |C_v| \geq n] - \mathbb{P}[v \leftrightarrow \mathcal{G}]) \mu[|C_v| \geq n] \\
&= (\mathbb{E}[(1 - e^{-\lambda|C_v|/n}) | |C_v| \geq n] - \mathbb{P}[v \leftrightarrow \mathcal{G}]) \mu[|C_v| \geq n] \\
&\geq ((1 - e^{-\lambda}) - \mu[1 - e^{-\lambda|C_v|/n}]) \mu[|C_v| \geq n].
\end{aligned}$$

Plugging in this bound and our bound on $\delta_e(F, \mathbb{P})$, we find

$$((1 - e^{-\lambda}) - \mu[1 - e^{-\lambda|C_v|/n}]) \mu[|C_v| \geq n] \leq 2 \sup_{u \in V} \mu[1 - e^{-\lambda|C_u|/n}] \cdot \sum_{e \in E} \mathrm{Cov}[f, \omega_e],$$

which implies the result. □

We are ready to prove Theorem 5.4.

Proof. Consider Proposition 5.5 where $G = \mathbb{Z}^d$, $\mu = \phi^\xi_{\beta,q}$, and $v = 0$. Note that since $1 - e^{-x} \leq x$ for all x, we have the bound $1 - e^{-\lambda|C_u|/n} \leq 1 \wedge \frac{\lambda|C_u|}{n}$. Also, note that any nonnegative integer-valued random variable X satisfies $\mathbb{E}[X] = \sum_{m\geq 1} \mathbb{P}[X \geq m]$. It follows that

$$\begin{aligned}
\phi^\xi_{\beta,q}[1 - e^{-\lambda|C_u|/n}] &\leq \frac{\lambda}{n}\phi^\xi_{\beta,q}\left[\frac{n}{\lambda} \wedge |C_u|\right] \\
&= \frac{\lambda}{n}\sum_{m\geq 1}\phi^\xi_{\beta,q}\left[\frac{n}{\lambda} \wedge |C_u| \geq m\right] \\
&= \frac{\lambda}{n}\sum_{m=1}^{\lceil n/\lambda\rceil}\phi^\xi_{\beta,q}[|C_u| \geq m],
\end{aligned}$$

where in the last step we used the fact that $\frac{n}{\lambda} \wedge |C_u|$ is at most $\lceil n/\lambda\rceil$. By translation invariance in infinite volume, we in fact have that

$$\sup_{u\in V}\phi^\xi_{\beta,q}[1 - e^{-\lambda|C_u|/n}] \leq \frac{\lambda}{n}\sum_{m=1}^{\lceil n/\lambda\rceil}\phi^\xi_{\beta,q}[|C_0| \geq m].$$

Therefore Proposition 5.5 implies that

$$\sum_{e\in E}\mathrm{Cov}[\mathbf{1}_{|C_0|\geq n}, \omega_e] \geq \left[\frac{1 - e^{-\lambda}}{2\frac{\lambda}{n}\sum_{m=1}^{\lceil n/\lambda\rceil}\phi^\xi_{\beta,q}[|C_0| \geq m]} - \frac{\phi^\xi_{\beta,q}[1 - e^{-\lambda|C_0|/n}]}{2\sup_{u\in V}\phi^\xi_{\beta,q}[1 - e^{-\lambda|C_u|/n}]}\right]\cdot\phi^\xi_{\beta,q}[|C_0| \geq n],$$

where covariance is taken with respect to the measure $\phi^\xi_{\beta,q}$. Upon bounding the subtracted fraction by $1/2$ and rearranging, we find

$$\sum_{e\in E}\mathrm{Cov}[\mathbf{1}_{|C_0|\geq n}, \omega_e] \geq \frac{1}{2}\left[\frac{n(1 - e^{-\lambda})}{\lambda\sum_{m=1}^{\lceil n/\lambda\rceil}\phi^\xi_{\beta,q}[|C_0| \geq m]} - 1\right]\cdot\phi^\xi_{\beta,q}[|C_0| \geq n].$$

We now apply Russo's formula in infinite volume (Proposition 2.7) to find

$$\left(\frac{d}{d\beta}\right)_+\phi^\xi_{\beta,q}[|C_0| \geq n] \geq \sum_{e\in E}\frac{J_e}{e^{\beta J_e} - 1}\mathrm{Cov}[\mathbf{1}_{|C_0|\geq n}, \omega_e].$$

Uniformly bounding $\frac{J_e}{e^{\beta J_e}-1}$, we find

$$\max_{e\in E}\left[\frac{e^{\beta J_e} - 1}{J_e}\right]\left(\frac{d}{d\beta}\right)_+\phi^\xi_{\beta,q}[|C_0| \geq n] \geq \sum_{e\in E}\mathrm{Cov}[\mathbf{1}_{|C_0|\geq n}, \omega_e].$$

Combining this with the lower bound above and applying the logarithmic derivative formula for Dini derivatives yields the result. ☐

§5.3 Application 3: weak ϕ^3 inequality

For our last application, we return to Conjecture 3.2. The following are recent developments of Aizenman and El Bahri, by personal communication. The ϕ^3 inequality for percolation on $\mathbb{Z}^d$ with $d \geq 2$ states that

$$M \leq h\frac{\partial M}{\partial h} + \beta M\frac{\partial M}{\partial \beta} + M^2,$$

where $M = \mathbb{E}[1-e^{-h|C_0|}]$. As we have seen, the proof for percolation does not directly generalize to the random cluster model due to its dependence on the BK inequality. We present a unified proof of a weak version of this inequality for all $q \geq 1$ and boundary conditions $\xi \in \{f, w\}$.

To state our inequality, we define a ghost field η on sites of $\mathbb{Z}^d$ such that $\eta_x = 1$ with probability $1-e^{-h}$ and $\eta_x = 0$ otherwise, with η independent of the configuration ω. We write μ for the distribution of η. We define the magnetization of the random cluster model as $M = \mathbb{E}[1-e^{-h|C_0|}]$, where expectation is taken with respect to the joint distribution $\mathbb{P}$ of the ghost field η and the configuration $\omega \sim \phi^{\xi}_{\beta,q}$.

Theorem 5.6 (Weak ϕ^3 inequality)
Given $d \geq 2$, $q \geq 1$, and boundary condition $\xi \in \{f, w\}$, we have

$$M \leq h\left(\frac{\partial}{\partial h}\right)_+ M + \beta M\left(\frac{\partial}{\partial \beta}\right)_+ M + M^2$$

for all $\beta \geq 0$ and $h \geq 0$.

In our proof, we apply infinite volume Russo's formula (Proposition 2.7) to the product measure $\mathbb{P}$. Although the configuration ω and ghost field η individually satisfy Russo's formula, the product space is not naturally a random cluster model. For this reason we explicitly justify the use of Russo's formula below.

Proof. We apply Theorem 4.5 to the single function $f(\omega, \eta) = \mathbf{1}_{0 \leftrightarrow \mathcal{G}}$ with the decision forest utilized in the proof of Proposition 5.5. As mentioned above, we choose the ghost field such that a given site u is green with probability $1-e^{-h}$. Theorem 4.5 reads

$$\mathrm{Var} f \leq \sum_{u \in V} \delta_u(F, \mathbb{P}) \mathrm{Cov}[f, \eta_u] + \sum_{e \in E} \delta_e(F, \mathbb{P}) \mathrm{Cov}[f, \omega_e],$$

where covariance is taken with respect to $\mathbb{P}$. As a first simplification, note that since f takes values in $\{0, 1\}$, we have $\mathrm{Var} f = \mathbb{E}[f] - \mathbb{E}[f]^2 = M - M^2$.

We begin by bounding the second sum on the right-hand side. Note that an edge $e = xy$ is discovered by our forest if and only if e is open and x lies in the cluster of some green site. We have performed this computation before: the probability that x lies in the cluster of some green site equals $\mathbb{E}[1-e^{-h|C_x|}]$, which by translation invariance is precisely M. Hence $\delta_e(F, \mathbb{P}) \leq M$ for all $e \in E$, so that the second sum is bounded by $M \sum_{e \in E} \mathrm{Cov}[f, \omega_e]$.

We claim that the following inequality $(**)$ holds:

$$\left(\frac{\partial}{\partial \beta}\right)_+ M \geq \sum_{e \in E} \frac{J_e}{e^{\beta J_e} - 1} \mathrm{Cov}[f, \omega_e].$$

To see this, first note that

$$\begin{aligned}
\left(\frac{\partial}{\partial \beta}\right)_+ M &= \left(\frac{\partial}{\partial \beta}\right)_+ \sum_{\hat{\eta} \in \{0,1\}^{\mathbb{Z}^d}} \phi^{\xi}_{\beta,q}[f(\omega, \hat{\eta})] \cdot \mu[\hat{\eta}] \\
&\geq \sum_{\hat{\eta} \in \{0,1\}^{\mathbb{Z}^d}} \left(\frac{\partial}{\partial \beta}\right)_+ \phi^{\xi}_{\beta,q}[f(\omega, \hat{\eta})] \cdot \mu[\hat{\eta}],
\end{aligned}$$

where in the first step we conditioned on $\eta = \hat{\eta}$, and where in the second step we used the superaddivity of the Dini derivative. We may now apply Proposition 2.7 to the function $f(\omega, \hat{\eta})$ to find the lower bound

$$\sum_{\hat{\eta}\in\{0,1\}^{\mathbb{Z}^d}} \sum_{e\in E} \frac{J_e}{e^{\beta J_e}-1} \mathrm{Cov}_{\phi^{\xi}_{\beta,q}}[f(\omega,\hat{\eta}),\omega_e] \cdot \mu[\hat{\eta}] = \sum_{e\in E} \frac{J_e}{e^{\beta J_e}-1} \sum_{\hat{\eta}\in\{0,1\}^{\mathbb{Z}^d}} \mathrm{Cov}_{\phi^{\xi}_{\beta,q}}[f(\omega,\hat{\eta}),\omega_e] \cdot \mu[\hat{\eta}]$$
$$= \sum_{e\in E} \frac{J_e}{e^{\beta J_e}-1} \mathrm{Cov}[f,\omega_e],$$

where the last covariance is with respect to $\mathbb{P}$. The inequality $(**)$ follows.

Since $e^x - 1 \geq x$ for all x, the fraction $\frac{J_e}{e^{\beta J_e}-1}$ is at most β. Multiplying both sides of $(**)$ by β, we see that

$$\beta\left(\frac{\partial}{\partial\beta}\right)_+ M \geq \sum_{e\in E} \mathrm{Cov}[f,\omega_e].$$

Therefore the second sum is bounded by $\beta M\left(\frac{\partial}{\partial\beta}\right)_+ M$.

Next, we bound the first sum on the right-hand side. Apply the trivial bound $\delta_u(F,\mathbb{P}) \leq 1$ for all u, so that we have the bound $\sum_{u\in\mathbb{Z}^d} \mathrm{Cov}[f,\eta_u]$. We claim that the inequality $(*)$ holds:

$$\left(\frac{\partial}{\partial h}\right)_+ M \geq \frac{1}{1-e^{-h}} \sum_{u\in\mathbb{Z}^d} \mathrm{Cov}[f,\eta_u].$$

We proceed as in the previous case. Note that Proposition 2.7 applies to η and yields the inequality

$$\left(\frac{\partial}{\partial h}\right)_+ g(\eta) \geq \frac{\partial(1-e^{-h})}{\partial h} \cdot \frac{1}{e^{-h}(1-e^{-h})} \sum_{u\in V} \mathrm{Cov}_{\mu}[g,\eta_u] = \frac{1}{1-e^{-h}} \sum_{u\in V} \mathrm{Cov}_{\mu}[g,\eta_u]$$

for any increasing function $g : \{0,1\}^{\mathbb{Z}^d} \to \mathbb{R}$. Following the proof of $(**)$, we write

$$\left(\frac{\partial}{\partial h}\right)_+ M \geq \sum_{\hat{\omega}\in\{0,1\}^E} \left(\frac{\partial}{\partial h}\right)_+ \mu[f(\hat{\omega},\eta)] \cdot \phi^{\xi}_{\beta,q}[\hat{\omega}]$$
$$\geq \sum_{\hat{\omega}\in\{0,1\}^E} \sum_{u\in\mathbb{Z}^d} \frac{1}{1-e^{-h}} \mathrm{Cov}_{\mu}[f(\hat{\omega},\eta),\eta_u] \cdot \phi^{\xi}_{\beta,q}[\hat{\omega}]$$
$$= \frac{1}{1-e^{-h}} \sum_{u\in\mathbb{Z}^d} \mathrm{Cov}[f,\eta_u],$$

where the last covariance is with respect to $\mathbb{P}$. The inequality $(*)$ follows.

Rearranging $(*)$, we find

$$(1-e^{-h})\left(\frac{\partial}{\partial h}\right)_+ M \geq \sum_{u\in\mathbb{Z}^d} \mathrm{Cov}[f,\eta_u].$$

Since $1 - e^{-h} \leq h$ for all $h > 0$, it follows that the first sum is bounded by $h\left(\frac{\partial}{\partial h}\right)_+ M$.

Putting the bounds on the first and second sums together and rearranging, we arrive at the desired inequality. □

§6 Discussion

In this paper, we presented differential inequalities for the random cluster model, including a radius tail bound, a volume tail bound, and a weak version of a ϕ^3 inequality. We saw that classical techniques from Bernoulli percolation do not directly generalize, and instead new tools from the theory of decision trees and decision forests are useful. Although we did not present the analysis of the differential inequalities in this report, a number of the results imply so-called critical exponent bounds. Both Theorem 5.1 and Theorem 5.4 imply that $\theta(\beta)$ grows at least linearly for $\beta \geq \beta_c$ with β sufficiently close to β_c, and Theorem 5.4 implies other critical exponent bounds on the tail of the volume of C_0 [DCRT18] [Hut20]. If Conjecture 3.2 is true, then a number of critical exponent bounds on the magnetization M from [AB87] would carry over to the random cluster model for $q \in [1, 2]$. At present, we do not have a method of attacking the Burgers' inequality, but given the simplicity of its proof for percolation it is likely that the result is within reach of current tools. Regardless, the development of new differential inequalities to probe critical phenomena will remain an exciting part of statistical mechanics in the years to come.

§7 Acknowledgements and Honor Pledge

I would like to thank Professor Michael Aizenman for serving as my Junior Paper advisor and for guiding me through the field of percolation theory.

"I pledge my honour that this paper represents my own work in accordance with University regulations."

– Sunay Joshi

References

[AB87] Michael Aizenman and David J Barsky. Sharpness of the phase transition in percolation models. *Communications in Mathematical Physics*, 108(3):489–526, 1987.

[ABF87] Michael Aizenman, David J Barsky, and Roberto Fernandez. The phase transition in a general class of ising-type models is sharp. *Journal of Statistical Physics*, 47(3/4), 1987.

[DC17] Hugo Duminil-Copin. Lectures on the ising and potts models on the hypercubic lattice, 2017.

[DC18] Hugo Duminil-Copin. Introduction to bernoulli percolation. 2018.

[DCRT18] Hugo Duminil-Copin, Aran Raoufi, and Vincent Tassion. Sharp phase transition for the random-cluster and potts models via decision trees, 2018.

[Dur19] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.

[Gri06] Geoffrey R Grimmett. *The Random-Cluster Model*, volume 333. Springer Science & Business Media, 2006.

[Hut20] Tom Hutchcroft. New critical exponent inequalities for percolation and the random cluster model. *Probability and Mathematical Physics*, 1(1):147–165, nov 2020.

[Lan11] Mikel Landajuela. Burgers equation. *BCAM Internship report: Basque Center for Applied Mathematics*, 2011.

[Men86] Mikhail V Menshikov. Coincidence of critical points in percolation problems. In *Soviet Mathematics Doklady*, volume 33, pages 856–859, 1986.

[OSSS05] Ryan O'Donnell, Michael Saks, Oded Schramm, and Rocco A Servedio. Every decision tree has an influential variable. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 31–39. IEEE, 2005.