Percolation Bounds for Phase Transition Temperatures of Randomized Lattice Ising Models

Isaac Liao
Machine Learning Department
Carnegie Mellon University
Pittsburgh, PA 15289
iliao@andrew.cmu.edu

Saumya Goyal
Machine Learning Department
Carnegie Mellon University
Pittsburgh, PA 15289
saumyag@andrew.cmu.edu

Abstract

In this project, we study the relationship between two seemingly unrelated phenomena: the clustering of nearby spins in samples from Ising models, and the connectivity of random graphs. It turns out that both systems exhibit a phase transition, whereby depending on the temperature and edge removal probability respectively, either the whole system agglomerates together into one spin or connected component, or the whole system fragments into locally connected pieces. Recent progress in information theory on the topic of information percolation provides an unexpected link between the two phenomena, allowing us to study one phase transition indirectly through the other. We derive an information theoretic bound between the two transition points in the two systems. We explore the properties of this bound through simulations of Ising models and percolation networks.

1 Introduction

Lattice Ising models are a type of probabilistic graphical model commonly used as a toy model for studying ferromagnetism in statistical mechanics [2]. A common and well studied kind of lattice Ising model has a binary random variable $X_{ij} \in \{-1, 1\}$ for every lattice point (i, j) in a 2 dimensional space, and has a probability distribution proportional to the exponential of a potential function $\Phi(X)$ defined as follows:

$$\Phi(X) = \sum_{i,j} J_{ij} X_{ij} + \sum_{(i_1 j_1, i_2 j_2) \in \text{adjacencies}} J_{i_1 j_1, i_2 j_2} X_{i_1 j_1} X_{i_2 j_2}, \tag{1}$$

where X is the set of all X_{ij}, and J_{ij} and $J_{i_1 j_1, i_2 j_2}$ are coefficients provided for each lattice point and adjacency, respectively. Each random variable in X represents an atom's spin within a crystal of ferromagnetic material. The coefficients, whose magnitude is inversely proportional to the temperature of crystal, can provide encouragement for spins to align with each other in large clusters to form a total spin of larger magnitude, so that the material as a whole becomes magnetic on a macroscopic scale. Ising models with constant J_{ij} and $J_{i_1 j_1, i_2 j_2}$ are well known to exhibit two distinct phases separated by a sharp transition temperature: a "paramagnetic" phase at high temperatures where clusters of aligned spins are small and numerous, and a "ferromagnetic" phase at low temperatures where the whole crystal is mostly part of one large cluster of the same spin. You can easily see this phase transition in reality, by heating a permanent magnet with a blowtorch, and observing upon reaching a very specific temperature universal to all magnets of that kind, it suddenly stops sticking to metal. The exact temperature of this phase transition is core to our understanding of magnetism, which is why condensed matter physicists are interested in finding where it is. Unfortunately though, the phase transition temperature is typically incredibly difficult to calculate/estimate for many kinds

Preprint. Under review.

of Ising models, and for decades, many a physicist have poured their blood, sweat, and tears into the process of finding it.

A separate line of work in the field of information theory has recently produced results regarding **information percolation**, which we believe can offer us help in finding the phase transition. Information percolation is the phenomenon that mutual information between two faraway nodes in a large graphical model, such as an Ising model, is upper bounded by their **percolation probability** [9]. This is the probability that the two nodes remain connected when nodes and/or edges are removed randomly [11], which can be estimated very easily in a computer via simulation.

Why do we care about the mutual information and the percolation probability? It turns out that a telltale sign of phase transition in an Ising model is when we change the temperature slightly, and faraway spins in the crystal rapidly become correlated, gaining mutual information. This is because faraway nodes will have the same spin if they are part of the same cluster, whereas with many smaller clusters they will be separated. Thus, knowing the mutual information between faraway spins tells us whether the temperature is above or below the critical temperature, which can be used to locate it.

As for the percolation probability, mathematicians have long known that percolation probabilities on lattice graphs also show a sharp phase transition, from one large connected component to many small components, dependent on the node/edge removal probability. That is, after a certain threshold value, the probability of far-off nodes being connected drops rapidly with the probability that each edge is removed, a phenomenon that reeks of similarity to the Ising model phase transition.

In light of these recent advances in information percolation, we ask in this paper whether the easily computable phase transition for percolation gives us another way to determine the transition temperatures for lattice Ising models. We estimate the percolation probability and mutual information on either side of a percolation bound, to demonstrate that the bound can give conclusions about the location of the Ising model phase transition. By measuring the gap of the bound for various grid sizes, we determine that the bound is not perfectly tight, but is still tight enough to be moderately conclusive of the transition temperature.

2 Background

This work is largely built upon the information percolation bound from Theorem 2 of [7], which we outline in this section.

2.1 Information Percolation Bound

The main information percolation bound that makes our work possible is Theorem 2 of [7], which we include in Equation 2. But first, we must explain how the theorem is set up:

We begin with an undirected graph $G = (V, E)$ where each node v has a random variable X_v from a finite alphabet $\mathcal{X}_v$, sampled independently. For each edge $e = \{u, v\}$, we define a random variable Y_e whose distribution $P(Y_e|X_u, X_v)$ is conditional only on the variables X_u and X_v at the nodes which it connects. The resulting probabilistic graphical model is then bipartite, with $X = \{X_v; v \in V\}$ on one side and $Y = \{Y_e; e \in E\}$ on the other. This lets us define the mutual information between two faraway nodes when given observations on the edges, $\mathbb{E}_{y\sim Y}[I(X_u; X_v|Y = y)]$, which appears on the left side of the bound.

Now, having already defined a conditional distribution $P(Y_e|X_u, X_v)$ for every edge e, we can write what is known as the "strong data processing inequality contraction coefficient" $\eta_{\mathrm{KL}\,e}$ of this conditional distribution. It is a scalar quantity between 0 and 1, explained in detail in [9], and is well known for many simple distributions $P(Y_e|X_u, X_v)$; it suffices to know that $\eta_{\mathrm{KL}\,e}$ is well known for the conditional distributions we will be using in this paper.

Finally, we go back to the original graph G that has only X and not Y, and delete each edge e with probability $1 - \eta_{\mathrm{KL}\,e}$. Then, we denote by $\mathrm{perc}(u \to v)$ the probability that there exists a path from node u to node v in the graph G after the edge deletion process, a.k.a. the "percolation probability", which appears on the right side of the bound.

Finally, (a weakened version of) the information percolation bound states that

$$\mathbb{E}_{y\sim Y}[I(X_u; X_v|Y = y)] \leq (\ln \max_w |\mathcal{X}_w|)\mathrm{perc}(u \to v) \tag{2}$$

This says that when faraway nodes are most likely disconnected in the percolation graph, their corresponding spins must also be mostly uncorrelated in the graphical model. The phase transition for the left side occurs when the expected mutual information suddenly jumps from zero to nonzero as the conditional distribution $P(Y_e|X_u, X_v)$ varies, so we know the phase transition is impossible wherever the right side of the bound (2) is zero, which clamps the mutual information to zero. Next, we'll aim to write the left side as the mutual information in an Ising model, and give an algorithm to estimate the right side. Then, we can determine the region of $P(Y_e|X_u, X_v)$ where the right side is computed to be nonzero, and conclude that the phase transition of the Ising model must be within that region. This will transform the bound (2) into a computational method to helps us find the phase transition in the Ising model.

3 Method

In this section, we explain how to translate Theorem (2) into a statement about the relationship between Ising models and percolation networks. Then, we outline our algorithms for simulating these percolation networks, such that we can measure the quantities described by the bound.

3.1 The Left Side of the Bound comes from an Ising Model

Here, we want to show that by picking the setup of the bound (2) cleverly, we can equate its left side with the mutual information between faraway nodes in an Ising model. We pick the setup as follows:

The graph consists of an $n \times n$ grid of random variables X_v that can take values 1 or -1 with equal probability independently. We now add edges between neighbors, and also connect opposite sides of the grid together, to form a total of $2n^2$ edges. Each edge $Y_{(u,v)}$ is a binary random variable, that takes the value $X_u X_v$ with probability p and $-X_u X_v$ with probability $1-p$. This conditional distribution $P(Y_e|X_u, X_v)$ has a known contraction coefficient of $\eta_{\mathrm{KL}\,e} = (1-2p)^2$.[1].

Now we must show that the probability distribution $P(X|Y)$ is coming from this setup is actually the same as an Ising model over variables X when Y is given and fixed:

$$
\begin{aligned}
P(X|Y) &= \frac{P(Y|X)P(X)}{P(Y)} \quad \text{(Bayes rule)} \\
&\propto P(Y|X) \quad \text{(since } X \text{ is uniform and } Y \text{ is fixed)} \\
&= \prod_{e=(u,v)\in E} P(Y_e|X_u, X_v) \quad \text{(product of likelihoods of all observations)} \\
&= \prod_{e=(u,v)\in E} \exp\left(\frac{1+Y_e X_u X_v}{2}\log p + \frac{1-Y_e X_u X_v}{2}\log 1-p\right) \quad \text{(from chosen setup)} \\
&\propto \exp\left(\sum_{e=(u,v)\in E} \frac{\log\left(\frac{p}{1-p}\right)}{2} Y_e X_u X_v\right) \quad \text{(rearranging)}
\end{aligned}
$$

So we see that $P(X|Y)$ is actually an Ising model, with $J_{ij} = 0$ and $J_{i_1 j_1, i_2 j_2} = \frac{\log\left(\frac{p}{1-p}\right)}{2} Y_e$. But as we can see, the edge coefficients $J_{i_1 j_1, i_2 j_2}$ are picked randomly since Y_e are random. The left side of the bound then quantifies the average mutual information between faraway spins *in such an Ising model*.

We can even simplify the Ising model that we are working with, using the fact that the mutual information we are looking at is invariant to certain transformations. Specifically, when we swap all the Y_e connecting to a node v, the distribution $P(X|Y)$ then becomes $P(X/X_v, -X_v|Y)$, and the mutual information does not change. Noting that the way we generated the Y_e first involved picking some X uniformly at random, this means swapping all the Y_e connecting to a node v is equivalent to picking v oppositely in the first place. So, it really doesn't matter what X we used to generate

[1]This is called the "binary symmetric channel of flip probability p", BSC(p), and the $\eta_{\mathrm{KL}\,e}$ value can be found in Equation 16 of [9]

Y_e for the expectation, and we might as well pick X to be all ones, as this will not affect the mutual information. This leads to us generating an Ising model where Y_e is i.i.d. Bernoilli with probability p of being 1 and $1-p$ of being -1. Therefore, we can also say that $P(X|Y)$ above is an Ising model with $J_{ij}=0$ and iid edge weights $J_{i_1 j_1, i_2 j_2} = \frac{\log\left(\frac{p}{1-p}\right)}{2}\text{Bernoulli}(p)$.

It turns out that this particular kind of Ising model is known in condensed matter physics as the **Edwards-Anderson model** with Bernoulli couplings, a model of a **spin glass** [4] [5]. More commonly, physicists assume Gaussian couplings, which we can recover a specific case of, by picking the setup a bit differently. Specifically, we would choose each edge $Y_{(u,v)}$ to be a normally distributed random variable with mean $J_0 X_u X_v$ and variance 1, for some J_0. The $\eta_{\text{KL}\,e}$ value associated with this choice is now equal to $2\mathbb{E}_{y\sim N(J_0,1)}[\sigma(2yJ_0)]-1$ where $\sigma(x)=1/(1+e^{-x})$, as calculated in Appendix A. This expectation can be efficiently computed via trapezoidal integration. The derivation of the resulting Ising model form is similar to before:

$$\begin{aligned} P(X|Y) &= \prod_{e=(u,v)\in E} P(Y_e|X_u, X_v) \quad \text{(as before)} \\ &\propto \prod_{e=(u,v)\in E} \exp\left(-\frac{(J_0 X_u X_v - Y_e)^2}{2}\right) \quad \text{(from chosen setup)} \\ &\propto \exp\left(\sum_{e=(u,v)\in E} (J_0 Y_e) X_u X_v\right) \quad \text{(rearranging)} \end{aligned}$$

Again, $P(X|Y)$ is an Ising model, with $J_{ij}=0$, but now the edge weights are distributed as $N(J_0^2, J_0^2)$ instead of $\frac{\log\left(\frac{p}{1-p}\right)}{2}\text{Bernoulli}(p)$, after the same simplification trick we used before to presume X is all ones. Thus we have recovered the Edwards-Anderson model of the spin glass, but this time with $N(J_0^2, J_0^2)$ couplings.

3.2 The Right Side of the Bound can be Computed Easily

Recall from the setup that the right side of the bound 2 is the probability that nodes u and v remain connected after deleting each edge with some probability $p_{drop} = 1-\eta_{\text{KL}\,e} = 1-(1-2p)^2$. So, we simulate that on a computer by creating grids that wrap around, and then creating graphs by dropping each edge with probability p_{drop}. We then breadth-first-search to determine whether the corner node $u=(0,0)$ and the middle node $v=(n/2,n/2)$ (where the grid is of size n) are connected, and average over many runs to estimate $\text{perc}(u\to v)$. Since we are working with binary variables, we have $\ln\max_w|\mathcal{X}_w| = \ln 2$, finishing off the right side of the bound.

3.3 Known Transition Temperatures from Literature

The phase transition probability of edge percolation on grids is well known to be $p_{\text{drop}}=1/2$, and can be solved using dual lattices [3]. This means when $p > 1/2 - 1/\sqrt{8}$ (which corresponds to $p_{\text{drop}}=1/2$), we must have $\text{perc}(u\to v)=0$. Due to the upper bound (2) and our Ising model equivalence that the phase transition, this means that Ising models with edge weights

$$J_{i_1 j_1, i_2 j_2} = \frac{\log(3-2\sqrt{2})}{2}\text{Bernoulli}(1/2 - 1/\sqrt{8})$$

must be in the paramagnetic, decorrelated phase. Consequently, any Ising models with weaker weights than $\frac{\log(3-2\sqrt{2})}{2}$ or weights more randomized than $\text{Bernoulli}(1/2-1/\sqrt{8})$ must also be in the paramagnetic, decorrelated phase. Therefore, the phase transition boundary for the Ising model must be at a weight strength stronger than $\frac{\log(3-2\sqrt{2})}{2}$ for randomization as $\text{Bernoulli}(1/2-1/\sqrt{8})$, or must be less random than $\text{Bernoulli}(1/2-1/\sqrt{8})$ when the weight strength is $\frac{\log(3-2\sqrt{2})}{2}$.

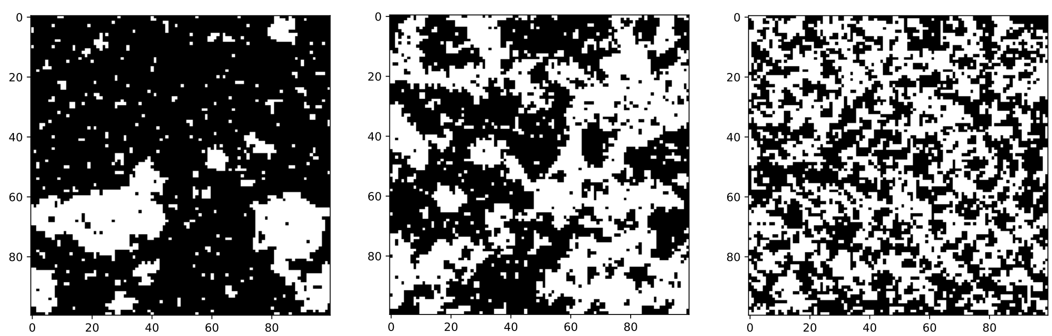
4 Preliminary Results

4.1 Baseline Method: Direct Simulation of Ising Models

In order to evaluate the efficacy of our method for bounding long range mutual information in Ising models, we need a baseline method for measuring this mutual information. As such, we perform simulations of Ising models to obtain the baseline.

More specifically, we use annealed importance sampling (AIS) to estimate the partition functions of the Ising models at various temperatures, and use these partition functions as a part of a simulated tempering routine to to sample from our Ising models at a specific temperature. We first generate a new Ising model with randomized edge weights of magnitude $\log(\frac{p}{1-p})/2$, and with positive sign with probability p, and sample 100 times to get the empirical mutual information $I(X_u; X_v|Y)$ between $u = (0, 0)$ and $v = (n/2, n/2)$. We do this with 10 Ising models using 10 samples for Y, to estimate the long range mutual information.

Some samples from our Ising models, for $p \in \{0.1, 1/2 - 1/\sqrt{8}, 0.2\}$ are shown in Figure 1.



$p = 0.1$ $\quad$ $p = 1/2 - 1/\sqrt{8}$ $\quad$ $p = 0.2$

Figure 1: Samples from Ising models with edge weights disributed as $\frac{\log\left(\frac{p}{1-p}\right)}{2}$Bernoulli$(p)$, for various values of p.

Treating our Ising model simulations as the ground truth, we now calculate our percolation bound, to plot the mutual information bound for various values of p, shown in Figure 2 in red.

5 Discussion

Our experiments show that the bound we derived holds, as indicated by the fact that the red line is above the black line by a statistically significant margin. In addition, both systems (Ising model and percolation) show a phase transition, indicated by the vertical jump in the graphs, separating regimes of low p and high p. We are pleased to observe that the red bound and the black measurements are near each other, indicating the efficacy of our method for bounding the long-range mutual information for a cheap computational cost.

For finite grid sizes, we observe in Figure (2) a soft jump in mutual information instead of a discrete jump. Nevertheless, a sharp phase transition does seem to be developing in both systems near $p = 1/2 - \sqrt{1/8} \approx 0.146$. We are unable to further resolve the phase transition for either system, due to computational constraints. The bound seems quite tight for smaller grid sizes, but unfortunately, a gap in the bound develops when the grid size increases. The well-known transition threshold for the square lattice percolation network is at $p_{\text{drop}} = 1/2$ corresponding to $p = 1/2 - \sqrt{1/8} \approx 0.146$ [3], so we expect the percolation network transition to sharpen at this location as the grid size goes to infinity.

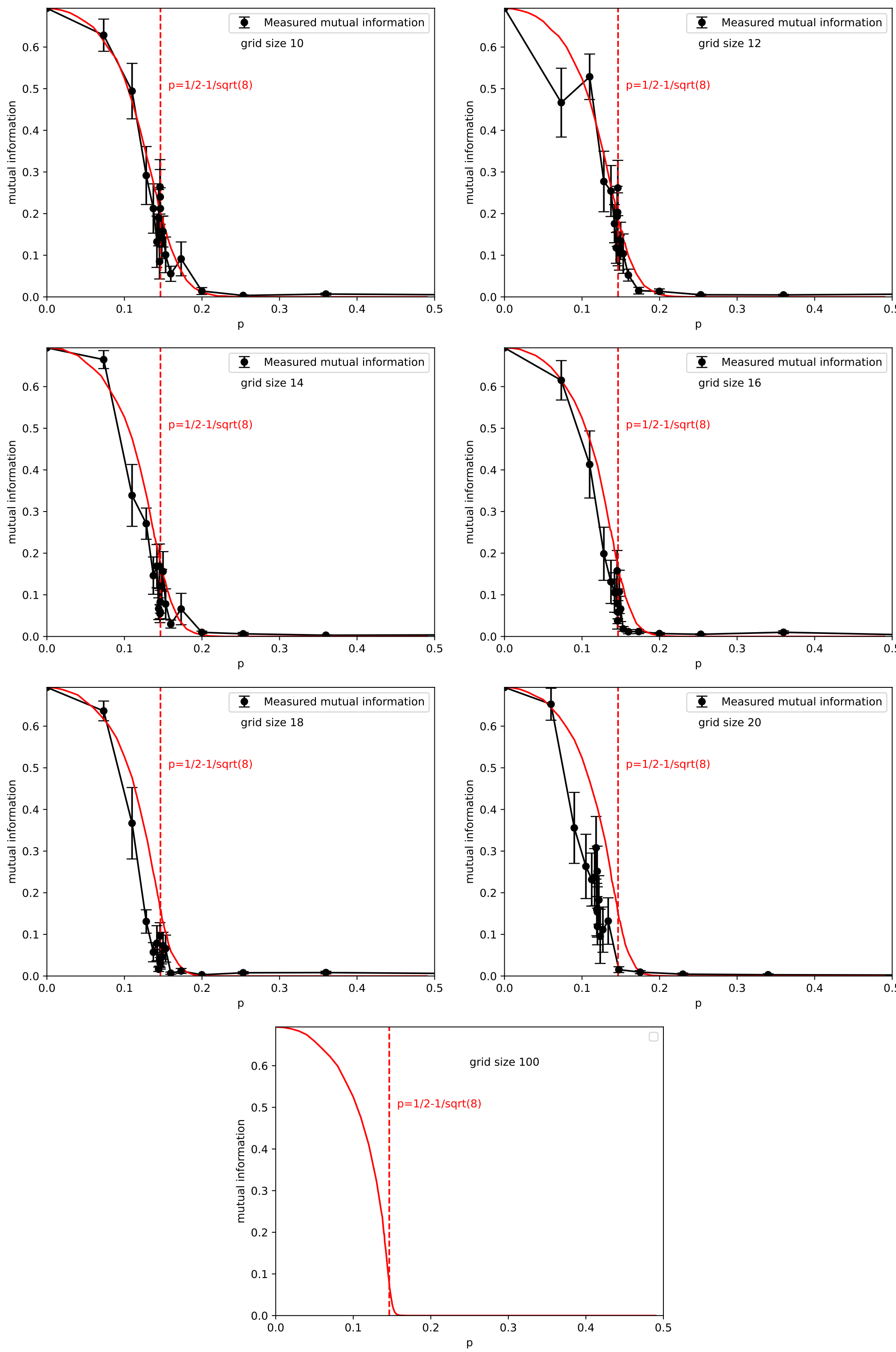


Figure 2: Left: The red line denotes our upper bound on the mutual information between distant spins in an Ising model, via measuring percolation probabilities, for various Ising model randomizations p, with various grid sizes. The actual mutual information, in black, is measured by Ising model simulation as in Section 4.1. One standard deviation uncertainties for Ising model simulations are plotted, and computed by adding in quadrature the inter-Y-sample uncertainty and the uncertainty in the mutual information estimate given by the correlation between two nodes, in each sample of Y. The red line denotes the known theoretical percolation threshold for infinite grids. [3] The bottommost plot is for grid size 100, where we were unable to directly measure the mutual information due to computational constraints; nevertheless our bound, which can still easily be computed, is shown.

6 Limitations

Our current analysis does not allow for us to vary the edge sign flip probability independently from the temperature; our results only pertain to the phases of the randomized Ising model on a one dimensional line in the two dimensional flip probability-temperature space. Similarly in the Gaussian couplings case, we were unable to find a way to allow the mean and variance of Y_e to change independently; they were also tied together. This was a major barrier which prevented us from drawing phase diagrams involving both parameters, which we would have preferred to include if possible.

Physicists have determined the critical exponents of the Ising model and percolation networks' phase transitions, and shown that the two models exhibit different limiting behavior at the boundaries of their phase transitions; ie. they belong to different universality classes. We suspect this may imply that our percolation-based bound is not saturated and can still be improved despite its success so far, though we are unsure of this conjecture. There may exist a graphical model out there whose behavior has the same critical exponents as the Ising model, and for those models, it may still be possible to saturate our bound.

7 Future Work

The framework of information percolation for Ising models which we have built, allows for a lot of modification. Variations of this work can be made for:

- Non-isotropic lattices
- Higher dimensional lattices
- Lattice graphical models with more than two possible states for a node; ie. the n-state clock model for finitely many states [12], the XY model for a continuous circle of states [10], and the classical Heisenberg model for a continuous sphere of states [6]
- Models of ferromagnetism based upon fully connected graphs, such as the Sherrington-Kirkpatrick model [1]

8 Teammates, work division

This team consists of Isaac Liao and Saumya Goyal. Isaac is primarily responsible for physical interpretations of Ising models, as well as managing the Ising model sampling code. Saumya is responsible for managing the percolation network simulations. Both are jointly responsible for coming up with the derivation showing that the left side of the bound in Theorem 2 of [7] is actually pertinent to randomized Ising models. Both are also jointly responsible for writing and generating plots. The code is publicly available on GitHub[2].

References

[1] J R L de Almeida and D J Thouless. "Stability of the Sherrington-Kirkpatrick solution of a spin glass model". In: *Journal of Physics A: Mathematical and General* 11.5 (May 1978), p. 983. DOI: 10.1088/0305-4470/11/5/028. URL: https://dx.doi.org/10.1088/0305-4470/11/5/028.

[2] Ernst Ising. "Contribution to the Theory of Ferromagnetism". In: *Z. Phys.* 31 (1925), pp. 253–258. DOI: 10.1007/BF02980577.

[3] Harry Kesten. "Analyticity properties and power law estimates of functions in percolation theory". In: *Journal of Statistical Physics* 25.4 (Aug. 1981), pp. 717–756. ISSN: 1572-9613. DOI: 10.1007/BF01022364. URL: https://doi.org/10.1007/BF01022364.

[4] John A Mydosh. *Spin glasses*. London, England: CRC Press, Apr. 2014.

[5] Hidetoshi Nishimori. *Statistical Physics of Spin Glasses and Information Processing: An Introduction*. Oxford University PressOxford, July 2001. ISBN: 9780191709081. DOI: 10.1093/acprof:oso/9780198509417.001.0001. URL: http://dx.doi.org/10.1093/acprof:oso/9780198509417.001.0001.

[2] https://github.com/saum-g/ising_percolation

[6] A.M. Polyakov. "Interaction of goldstone particles in two dimensions. Applications to ferromagnets and massive Yang-Mills fields". In: *Physics Letters B* 59.1 (Oct. 1975), pp. 79–81. ISSN: 0370-2693. DOI: 10.1016/0370-2693(75)90161-6. URL: http://dx.doi.org/10.1016/0370-2693(75)90161-6.

[7] Yury Polyanskiy and Yihong Wu. *Application of information-percolation method to reconstruction problems on graphs*. 2020. arXiv: 1806.04195 [cs.IT]. URL: https://arxiv.org/abs/1806.04195.

[8] Yury Polyanskiy and Yihong Wu. *Information Theory: From Coding to Learning*. Cambridge University Press, 2025.

[9] Yury Polyanskiy and Yihong Wu. *Strong data-processing inequalities for channels and Bayesian networks*. 2016. arXiv: 1508.06025 [cs.IT]. URL: https://arxiv.org/abs/1508.06025.

[10] H. E. Stanley. "Dependence of Critical Properties on Dimensionality of Spins". In: *Physical Review Letters* 20.12 (Mar. 1968), pp. 589–592. ISSN: 0031-9007. DOI: 10.1103/physrevlett.20.589. URL: http://dx.doi.org/10.1103/PhysRevLett.20.589.

[11] Dietrich Stauffer and Ammon Aharony. *Introduction to percolation theory*. 2nd ed. Philadelphia, PA: Taylor & Francis, Feb. 1992.

[12] F. Y. Wu. "The Potts model". In: *Reviews of Modern Physics* 54.1 (Jan. 1982), pp. 235–268. ISSN: 0034-6861. DOI: 10.1103/revmodphys.54.235. URL: http://dx.doi.org/10.1103/RevModPhys.54.235.

A Calculation of Contraction Coefficient for Gaussian Couplings

In Section 3.1, we state that for the conditional probability distribution $P(Y_{(u,v)}|X_u, X_v)$, which maps X_u, X_v to $N(J_0 X_u X_v, 1)$, the contraction coefficient is $\eta_{\text{KL}\,e} = 2\mathbb{E}_{y\sim N(J_0,1)}[\sigma(2yJ_0)] - 1$. Here, we provide the proof.

Firstly, [8] shows that for conditional probability distributions that have binary input, and are also input-symmetric, the contraction coefficient is equal to:

$$
\begin{aligned}
\eta_{\text{KL}\,e} =& I_{\chi^2}(X;Y) \quad \text{for } X \sim \text{Bernoulli}(1/2) \\
=& \mathbb{E}_{X\sim\text{Bernoulli}(1/2)}\left[\chi^2(P_{Y|X}||P_Y)\right] \\
=& \mathbb{E}_{X,Y\sim P_{XY}}\left[\frac{p(Y|X)}{p(Y)}\right] - 1 \\
=& -1 + \int \frac{\exp(-(y-J_0)^2/2)}{\sqrt{2\pi}} \frac{2\exp(-(y-J_0)^2/2)}{\exp(-(y-J_0)^2/2) + \exp(-(y+J_0)^2/2)}\, dy \\
=& -1 + \int \frac{\exp(-(y-J_0)^2/2)}{\sqrt{2\pi}} \frac{2\exp(yJ_0)}{\exp(yJ_0) + \exp(-yJ_0)}\, dy \\
=& -1 + \sqrt{\frac{2}{\pi}} \int \frac{\exp(-(y-J_0)^2/2)}{1+\exp(-2yJ_0)}\, dy \\
=& -1 + 2\frac{1}{\sqrt{2\pi}} \int \frac{\exp(-(y-J_0)^2/2)}{1+\exp(-2yJ_0)}\, dy \\
=& 2\mathbb{E}_{y\sim N(J_0,1)}[\sigma(2yJ_0)] - 1, \quad \sigma(x) = \frac{1}{1+e^{-x}}
\end{aligned}
$$

This value can then be efficiently approximated using trapezoidal integration.