Cutoff for the Ising Model on Finite Regular Graphs via Information Percolation

Nishant Lamboria

under the guidance of

Dr. Anirban Basak and Dr. Riddhipratim Basu

Project Report

S.N. Bhatt Memorial Excellence Fellowship Program

International Centre for Theoretical Sciences (ICTS)
Bengaluru, India

July 2025

Abstract

We study how Glauber dynamics for the Ising model mixes on finite transitive graphs in the high-temperature regime. In particular, we focus on the cutoff phenomenon, where the system undergoes a sharp transition from being far from stationarity to being close to it. To analyze this behavior, we use the information percolation framework developed by Lubetzky and Sly, which traces the influence of the initial state backward in time. We develop the necessary theory from the ground up, assuming only basic familiarity with Markov chains, and then present an outline of the proof of the cutoff phenomenon.

1 Introduction

The Ising model is a fundamental object in statistical physics and probability, used to describe systems of interacting spins on a graph. Glauber dynamics provides a natural Markov chain for sampling from the Ising distribution by updating one spin at a time according to a local rule. A central question in this setting is: how quickly does the chain converge to its stationary distribution?

Of particular interest is the *cutoff phenomenon*, where convergence to stationarity occurs abruptly over a short time window. Understanding when and why cutoff occurs remains a key challenge in the study of Markov chains.

In this report, we examine the cutoff behavior of Glauber dynamics on finite transitive graphs in the high-temperature regime. Our analysis is based on the *information percolation* framework developed by Lubetzky and Sly, which traces the influence of the initial configuration backward in time. This framework organizes update histories into RED, GREEN, and BLUE clusters [4.3], each representing different modes of information flow. Using this classification, we outline how one can rigorously derive both upper and lower bounds on the mixing time and establish the existence of a cutoff.

2 Preliminaries

Let $(X_t)_{t\geq 0}$ be a time-homogeneous Markov chain on a finite state space Ω with transition matrix P. If the chain is *irreducible* and *aperiodic*, then it admits a unique stationary distribution π satisfying $\pi P = \pi$. Moreover, for any initial distribution μ, the distribution μP^t converges to π as $t \to \infty$. To measure how close the distribution of X_t is to stationarity, we use several standard ways of comparing probability distributions μ and ν on Ω.

Definition 2.1 (Total Variation Distance)**.** *The total variation distance between two probability measures μ and ν on a finite space Ω is defined as:*

$$\|\mu - \nu\|_{\mathrm{TV}} := \max_{A\subset\Omega} |\mu(A) - \nu(A)| = \frac{1}{2}\sum_{x\in\Omega} |\mu(x) - \nu(x)|.$$

Definition 2.2 (L^p Distance)**.** *For $p \geq 1$, if $\nu(x) > 0$ for all $x \in \Omega$, the $L^p(\nu)$ distance between μ and ν is defined as:*

$$\|\mu - \nu\|_{L^p(\nu)} := \left(\sum_{x \in \Omega} \left(\frac{\mu(x)}{\nu(x)} - 1 \right)^p \nu(x) \right)^{1/p}.$$

Let X_t be a Markov chain with stationary distribution π. The worst-case total variation distance at time t is defined as:

$$d_{\mathrm{TV}}(t) := \max_{x_0 \in \Omega} \|\mathbb{P}^t(x_0, \cdot) - \pi(\cdot)\|_{\mathrm{TV}}.$$

This quantity is non-increasing in nature, which allows us to define the **mixing time** for a given $\varepsilon > 0$ as:

$$t_{\text{MIX}}(\varepsilon) := \inf \{ t \geq 0 : d_{\mathrm{TV}}(t) \leq \varepsilon \}.$$

Similarly, the worst-case L^p distance ($p \geq 1$) is defined as:

$$d^{(p)}(t) := \max_{x_0 \in \Omega} \|\mathbb{P}^t(x_0, \cdot) - \pi(\cdot)\|_{L^p(\pi)}.$$

Theorem 2.3 ([4], Section 4.7)**.** *The Total Variation distance and the L^p distances satisfy:*

$$d_{\mathrm{TV}}(t) = \frac{1}{2} d^{(1)}(t) \quad \textit{and} \quad d^{(p)}(t) \leq d^{(q)}(t) \ \forall\ 1 \leq p \leq q.$$

To capture the abrupt drop in these distances during convergence, we now define the notion of a cutoff. A sequence of finite Markov chains indexed by n is said to exhibit a *cutoff* with window size w_n if there is a sequence of times $T_n \to \infty$ such that for every fixed $\varepsilon \in (0, 1)$,

$$t_{\text{MIX}}^{(n)}(1 - \varepsilon) \geq T_n - \Theta(w_n) \quad \text{and} \quad t_{\text{MIX}}^{(n)}(\varepsilon) \leq T_n + \Theta(w_n),$$

where $t_{\text{MIX}}^{(n)}$ is the mixing time for the n^{th} chain. In other words, the transition from *far-from-stationarity* to *close-to-stationarity* happens sharply, over a small window of time of width $w_n = o(T_n)$.

This phenomenon has been shown to occur in many natural examples, including random walks on groups, card shuffling processes, and certain spin systems. In this report, we focus on proving cutoff for the Glauber dynamics of the Ising model.

3 The Ising Model and Glauber Dynamics

3.1 The Ising Model

Let $G = (V, E)$ be a finite undirected graph. The Ising model at inverse temperature $\beta \geq 0$ is a probability distribution on spin configurations $\sigma \in \Omega := \{-1, +1\}^V$, where each vertex holds a positive or negative spin. The probability of a configuration σ is given as follows:

$$\pi(\sigma) = \frac{1}{Z(\beta)} \exp\left(\beta \sum_{\{u,v\} \in E} \sigma(u)\sigma(v)\right),$$

where $Z(\beta)$ is the normalizing constant (partition function), ensuring that π is a valid probability distribution.

3.2 Glauber Dynamics

Let $G = (V, E)$ be a finite graph and $\Omega \subset S^V$ a space of configurations over some finite set S. Suppose π is a probability distribution on Ω.

To simulate the distribution π, we use **Glauber dynamics**, a Markov chain on Ω specifically designed so that π is its stationary distribution. The dynamics are defined as follows: given the current configuration σ, a vertex $v \in V$ is chosen uniformly at random. Let $\Omega(\sigma, v) := \{\tau \in \Omega : \tau(w) = \sigma(w) \ \forall \ w \neq v\}$. The transition probability from σ to τ is given by:

$$P(\sigma, \tau) = \frac{1}{|V|} \cdot \frac{\pi(\tau)}{\pi(\Omega(\sigma, v))} \cdot \mathbb{1}\{\tau \in \Omega(\sigma, v)\}.$$

This definition ensures that the Markov chain is reversible with respect to π, i.e., it satisfies the detailed balance equations:

$$\pi(\sigma)P(\sigma, \tau) = \pi(\tau)P(\tau, \sigma) \quad \forall \ \sigma, \tau \in \Omega.$$

Therefore, π is the stationary distribution of the chain.

In the case of the Ising model, $S = \{-1, +1\}$, and the configuration space is $\Omega = \{-1, +1\}^V$. For a configuration $\sigma \in \Omega$ and a vertex $v \in V$, let $N(v)$ denotes the neighbors of v in G, define the *local field*:

$$S(\sigma, v) := \sum_{u \in N(v)} \sigma(u).$$

The Glauber dynamics proceeds as follows: choose a vertex $v \in V$ uniformly at random. Replace the spin at v with a new spin drawn from the marginal of π conditioned on the spins of its neighbors. Specifically, update $\sigma(v)$ to $+1$ with probability:

$$p(\sigma, v) = \frac{e^{\beta S(\sigma,v)}}{e^{\beta S(\sigma,v)} + e^{-\beta S(\sigma,v)}} = \frac{1 + \tanh(\beta S(\sigma, v))}{2},$$

and to -1 with the remaining probability. The dynamics is irreducible and aperiodic on any connected graph, and its stationary distribution is precisely the Ising measure π.

3.3 Continuous-Time Glauber Dynamics

We will work with the continuous-time version of Glauber dynamics. In this setup, each vertex $v \in V$ has its own independent Poisson clock that rings at rate 1. Whenever the clock at a vertex rings, we update the spin at that vertex using the update rule of Glauber dynamics. This defines a continuous-time Markov chain $(X_t)_{t\geq 0}$ on the space $\{-1,+1\}^V$, which has the Ising distribution π as its unique stationary distribution.

We are interested in understanding how quickly this chain converges to stationarity, and whether the convergence exhibits a sharp cutoff. The following theorem answers this question in the high-temperature regime for finite transitive graphs:

Theorem 3.1 ([1], Theorem 1)**.** *For any $d \geq 2$, there exists $\beta_0 = \beta_0(d) > 0$ such that the following holds. Let G be a d-regular transitive graph on n vertices. For any fixed $0 < \varepsilon < 1$, continuous-time Glauber dynamics for the Ising model on G at inverse-temperature $0 \leq \beta \leq \beta_0$ satisfies*

$$t_{MIX}(\varepsilon) = T_n \pm O_\varepsilon(1).$$

In particular, the dynamics on a sequence of such graphs has cutoff with an $O(1)$-window around $T_n := \inf\{t > 0 : \mathbb{E}[X_t^+] \leq 1/\sqrt{n}\}$, where $(X_t^+)_{t\geq 0}$ is the dynamics started from the all-plus *state.*

4 Update Sequence and Oblivious Updates

To understand how the Glauber dynamics evolves over time, we represent the process as a deterministic function of the initial configuration and a random update sequence.

In the continuous-time Glauber dynamics, each vertex has an independent exponential clock of rate 1. When a clock rings, the corresponding vertex updates its spin. We can represent the entire update process using a sequence of tuples $(J_1, U_1, t_1), (J_2, U_2, t_2), \ldots$, where $t_1 < t_2 < t_3 < \ldots$ are the times when the clocks ring, $J_i \in V$ is the vertex whose clock rings at time t_i, and $U_i \sim \mathrm{Unif}[0,1]$ are i.i.d. random variables, used to decide the new spin.

Given a starting state X_0, let $X_t = X_{t_{i-1}}$ for all $t \in (t_{i-1}, t_i)$ for all $i \geq 1$. To determine X_{t_i}, let $S_i := \sum_{u\sim J_i} X_{t_{i-1}}(u)$ be the sum of spins at the neighbors of J_i at time t_{i-1}. Update the spin at J_i according to:

$$X_{t_i}(J_i) = \begin{cases} +1 & \text{if } U_i < \dfrac{1+\tanh(\beta S_i)}{2}, \\ -1 & \text{otherwise.} \end{cases}$$

4.1 Backward Search and Oblivious Updates

To study how the configuration $X_{t_\star}$ depends on the initial state X_0, we take a backward view of the Glauber dynamics. That is, instead of simulating the process forward from time 0 to $t_\star$, we look at the sequence of updates that occurred between time 0 and $t_\star$, and trace which parts of the system influence the spin at a given vertex at time $t_\star$.

Formally, let $t_\star = t_0 > t_1 > t_2 > \ldots$ denote the update times in decreasing order over the interval $[0, t_\star]$. At each step, we examine the update that occurred at time t_i, and determine how it affects the current spin configuration.

Now, recall that in Glauber dynamics, the probability that a vertex takes the value $+1$ at update time t_i depends on the sum S_i of spins in its neighborhood. This probability is given by:

$$\mathbb{P}[X_{t_i}(J_i) = +1] = \frac{1 + \tanh(\beta S_i)}{2}.$$

However, for all possible values of S_i, this probability always stays within a fixed range away from 0 and 1. Specifically, using that $|S_i| \leq d$, we define $\theta := 1 - \tanh(\beta d)$. This gives us symmetric lower and upper bounds on $p(\sigma, v)$ as:

$$\frac{1 + \tanh(\beta S_i)}{2} \in \left[\frac{\theta}{2}, 1 - \frac{\theta}{2}\right].$$

In other words, even if all neighboring spins are aligned, there is always a positive probability that the vertex takes the opposite spin.

This observation motivates a reformulation of the update rule that separates part of the randomness into a uniform choice independent of the local configuration. For any update time t_i, let S_i be the sum of spins in the neighborhood of the vertex being updated. The original update rule sets the spin to $+1$ if $U_i < \frac{1}{2}(1 + \tanh(\beta S_i))$.

We now write this update rule in the following equivalent form: with probability θ, the update is called **oblivious**, and the spin is resampled uniformly from $\{+1, -1\}$. With the remaining probability $1 - \theta$, we expose the local field S_i, and update the spin to $+1$ if

$$U_i < \frac{1 + \tanh(\beta S_i)}{2} + \frac{\theta}{2},$$

and to -1 otherwise.

This reformulation preserves the original update probabilities exactly, since:

$$\mathbb{P}[X_{t_i}(J_i) = +1] = (\theta) \cdot \left(\frac{1}{2}\right) + (1 - \theta) \cdot \left(\frac{1 + \tanh(\beta S_i)}{2} + \frac{\theta}{2} - \theta\right) \cdot \left(\frac{1}{1 - \theta}\right) = \frac{1 + \tanh(\beta S_i)}{2}.$$

This reformulation can also be understood visually. Originally, the probability of updating to $+1$ depended on the local field through a biased subinterval of $[0, 1]$. By introducing oblivious updates, we effectively reshuffle and divide the interval into two parts:

Figure 1: The left plot shows the original threshold depending on the local field, while the right plot splits off an oblivious region for uniform resampling.

Note that whenever a vertex encounters an oblivious update, its new state is independent of the current local field and the past. This will allow us to identify points in the update history where the influence from the initial condition is completely erased.

4.2 Clusters in the Space-Time Slab

Given the update sequence and the reformulated dynamics, we can now trace how the spin at a vertex v at time $t_\star$ depends on earlier parts of the system. Starting from v, we first locate its most recent update. If this update is oblivious, the tracing process stops. Otherwise, if the update is non-oblivious, we branch out to its neighbors and recursively trace their update histories backward from that point.

This naturally gives rise to a branching process in the space-time slab $V \times [0, t_\star]$, with temporal edges connecting successive updates at each vertex and potential spatial edges linking neighboring vertices during non-oblivious updates. The process evolves recursively, branching whenever a non-oblivious update occurs, and continues until it either encounters an oblivious update—cutting off further influence along that branch—or reaches time zero.

The collection of all space-time points that influence the value of $X_{t_\star}(v)$ is called the *history* of vertex v, denoted by $\mathcal{H}_v$. For a subset of vertices $A \subset V$, we define the combined history as

$$\mathcal{H}_A := \bigcup_{v \in A} \mathcal{H}_v.$$

For each time $t \in [0, t_\star]$, let $\mathcal{H}_A(t) := \mathcal{H}_A \cap (V \times \{t\})$. This defines a subset of the space-time slab $V \times [0, t_\star]$ that encodes all the paths of influence contributing to the spins at A at time $t_\star$.

4.3 Classification of Clusters

We now group together vertices whose histories are connected in the space-time slab. These connected components, called **information percolation clusters**, describe how information propagates through the system.

Each cluster represents a region of space and time that jointly influences some part of the configuration at time $t_\star$. We classify these clusters into three types:

- **Red Clusters:** A cluster is called RED if it survives all the way back to time $t = 0$, i.e., $\mathcal{H}_A(0) \neq \emptyset$. In this case, the configuration at time $t_\star$ depends nontrivially on the initial state.

- **Blue Clusters:** A cluster is BLUE if it dies out before time 0 and consists of exactly one vertex. Since it doesn't depend on the initial condition, its spin at time $t_\star$ is essentially uniform and independent.

- **Green Clusters:** All other clusters are called GREEN. These are nontrivial clusters that do not survive to time 0 but do contain more than one vertex. Their behavior is random, but not completely independent.

This classification allows us to partition the vertex set into three subsets V_{RED}, V_{BLUE}, and V_{GREEN}, according to the cluster each vertex belongs to. The corresponding histories are denoted by $\mathcal{H}_{\text{RED}}$, $\mathcal{H}_{\text{BLUE}}$, and $\mathcal{H}_{\text{GREEN}}$, respectively. Analyzing the behavior of the dynamics conditioned on the green clusters is plays a crucial part in the proof of the cutoff phenomenon.

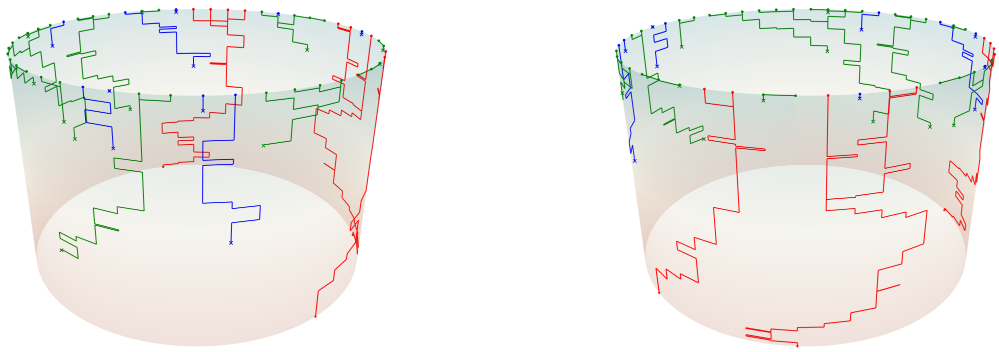

Figure 2: Simulation of clusters on an n-cycle. Here, *copying one of the neighbors* at a non-oblivious update is an equivalent rule to the Glauber dynamics ([1], Example 1).

5 Proof of the Main Theorem

The key idea is to condition on the history of GREEN clusters and analyze the impact of RED clusters. For which, we define the following key quantity:

Definition 5.1. *Let $\mathcal{H}_A^-$ denote the history of all vertices not in A up to time $t_\star$, i.e.,*

$$\mathcal{H}_A := \bigcup_{v \notin A} \mathcal{H}_v.$$

Then, for any $A \subset V$, we define:

$$\Psi_A := \sup_{\mathcal{X}} \mathbb{P}\left(A \in \mathit{RED} \,\middle|\, \mathcal{H}_A^- = \mathcal{X},\ \{A \in \mathit{RED}\} \cup \{A \subset V_{\mathit{BLUE}}\}\right).$$

This quantity captures the worst-case probability (over compatible green histories) that a given subset of vertices becomes a red cluster. To control this quantity, we apply the following lemma from Lubetzky and Sly:

Lemma 5.2 ([1], Lemma 2.1)**.** *For any $d \geq 2$ and $\lambda > 0$, there exist constants $\beta_0, C_0 > 0$ such that for all $0 < \beta < \beta_0$, and any subset $A \subset V$, the red cluster probability satisfies:*

$$\Psi_A \leq C_0 m_t e^{-\lambda \mathfrak{W}(A)},$$

where $\mathfrak{W}(A)$ is the size of the smallest connected subgraph of G containing A, and $m_t := \mathbb{E}[X_t^+]$ denotes the average magnetization at time t for the chain started from all-plus.

5.1 Upper Bound on Mixing Time

Let X_t and X_t' denote two copies of the Glauber dynamics starting from initial configurations x_0 and π, respectively. Then by Jensen's inequality:

$$\begin{aligned} d_{\mathrm{TV}}(t) &= \max_{x_0} \left\| \mathbb{P}_{x_0}(X_t \in \cdot) - \mathbb{P}_{X_0'}(X_t \in \cdot) \right\|_{\mathrm{TV}} \\ &\leq \max_{x_0} \mathbb{E}\left[\left\| \mathbb{P}_{x_0}(X_t \in \cdot \mid \mathcal{H}_{\textsc{Green}}) - \mathbb{P}_{X_0'}(X_t \in \cdot \mid \mathcal{H}_{\textsc{Green}}) \right\|_{\mathrm{TV}} \right]. \end{aligned}$$

By taking supremum over $\mathcal{H}_{\textsc{Green}}$, we obtain:

$$d_{\mathrm{TV}}(t) \leq \sup_{\mathcal{X}} \max_{x_0} \left\| \mathbb{P}_{x_0}(X_t \in \cdot \mid \mathcal{H}_{\textsc{Green}} = \mathcal{X}) - \mathbb{P}_{X_0'}(X_t \in \cdot \mid \mathcal{H}_{\textsc{Green}} = \mathcal{X}) \right\|_{\mathrm{TV}}.$$

Now, since the spin values on $V_{\textsc{Green}}$ are conditionally independent of the initial configuration, we may project onto $V \setminus V_{\textsc{Green}}$ to get:

$$\begin{aligned} d_{\mathrm{TV}}(t) \leq \sup_{\mathcal{X}} \max_{x_0} &\| \mathbb{P}_{x_0}(X_t(V \setminus V_{\textsc{Green}}) \in \cdot \mid \mathcal{H}_{\textsc{Green}} = \mathcal{X}) \\ &- \mathbb{P}_{X_0'}(X_t(V \setminus V_{\textsc{Green}}) \in \cdot \mid \mathcal{H}_{\textsc{Green}} = \mathcal{X}) \big\|_{\mathrm{TV}}. \end{aligned}$$

Let ν_A be the uniform measure on $A \subset V$, then by triangle inequality:

$$d_{\mathrm{TV}}(t) \leq 2 \sup_{\mathcal{X}} \max_{x_0} \left\| \mathbb{P}_{x_0}(X_t(V \setminus V_{\textsc{Green}}) \in \cdot \mid \mathcal{H}_{\textsc{Green}} = \mathcal{X}) - \nu_{V \setminus V_{\textsc{Green}}} \right\|_{\mathrm{TV}}$$

We now invoke the following lemma due to Miller and Peres:

Lemma 5.3 ([3], Proposition 3.2)**.** *Let μ be a measure obtained by randomly sampling a subset $R \subset V$, assigning spins on R according to some distribution φ_R, and assigning spins on $V \setminus R$ uniformly. Let ν be the uniform distribution on $\{\pm 1\}^V$. Then:*

$$\|\mu - \nu\|_{L^2(\nu)}^2 \leq \mathbb{E}\left[2^{|R \cap R'|} \right] - 1,$$

where R, R' are independent samples of the random subset.

Applying this in our setting with $R = (V_{\text{RED}} | \mathcal{H}_{\text{GREEN}} = \mathcal{X})$, and using the relation between $d_{\text{TV}}(t)$ and $d^{(2)}(t)$ we get:

$$d_{\text{TV}}(t) \leq \left(\sup_{\mathcal{X}} \mathbb{E}\left[2^{\left|V_{\text{RED}} \cap V'_{\text{RED}}\right|} \mid \mathcal{H}_{\text{GREEN}} = \mathcal{X} \right] - 1 \right)^{1/2}.$$

To bound this expectation, we use a coupling lemma from Lubetzky and Sly:

Lemma 5.4 ([1], Corollary 2.4)**.** *There exists a family of independent indicator variables $Y_{A,A'}$ satisfying $\mathbb{P}(Y_{A,A'} = 1) = \Psi_A \Psi_{A'}$, such that:*

$$(|V_{RED} \cap V'_{RED}| \mid \mathcal{H}_{GREEN} = \mathcal{X}) \leq \sum_{A \cap A' \neq \emptyset} |A \cup A'| \cdot Y_{A,A'}.$$

This implies:

$$\sup_{\mathcal{X}} \mathbb{E}\left[2^{\left|V_{\text{RED}} \cap V'_{\text{RED}}\right|} \mid \mathcal{H}_{\text{GREEN}} = \mathcal{X} \right] \leq \prod_{A \cap A' \neq \emptyset} \mathbb{E}\left[2^{(|A| + |A'|) Y_{A,A'}} \right] \leq \exp\left(n \left(\sum_{A \ni v} 2^{|A|} \Psi_A \right)^2 \right)$$

for any fixed vertex $v \in V$ as the graph is transitive. Now, using $e^x - 1 \leq 2x \ \forall\, x \in [0,1]$,

$$d_{\text{TV}}(t) \leq \left(e^{n\left(\sum_{A \ni v} 2^{|A|} \Psi_A\right)^2} - 1 \right)^{1/2} \leq \sqrt{2n} \sum_{A \ni v} 2^{|A|} \Psi_A.$$

Using the bound from Lemma 5.2, we get:

$$\sum_{A \ni v} 2^{|A|} \Psi_A \leq C_0 m_t \sum_k \sum_{\substack{A \ni v \\ \mathfrak{W}(A) = k}} 2^k e^{-\lambda k}.$$

Now Section 3 in [5] bounds the size of $\{A \ni v \mid \mathfrak{W}(A) \leq k\}$ by $(2ed)^k$. So, choosing $\lambda = \log(8ed)$, we obtain:

$$\sum_{A \ni v} 2^{|A|} \Psi_A \leq C_0 m_t \sum_k (4ede^{-\lambda})^k \leq 2 C_0 m_t,$$

$$\implies d_{\text{TV}}(t) \leq 2\sqrt{2} C_0 m_t \sqrt{n}.$$

We now conclude the upper bound by using the decay of magnetization over time.

Lemma 5.5 ([2], Claim 3.3)**.** *For any $t_0, t \geq 0$, the expected magnetization satisfies:*

$$e^{-t} m_{t_0} \leq m_{t_0 + t} \leq e^{-(1-\beta d)t} m_{t_0}.$$

Letting $t_\star = T_n + s_\star$ and applying the upper bound from the lemma:

$$m_{t_\star} = m_{T_n+s_\star} \le e^{-(1-\beta d)s_\star} m_{T_n} \le \frac{e^{-(1-\beta d)s_\star}}{\sqrt{n}},$$

we conclude:

$$d_{\mathrm{TV}}(t_\star) \le 2\sqrt{2} C_0 e^{-(1-\beta d)s_\star}.$$

Choosing $s_\star = s_\star(\varepsilon)$ large enough so that the right-hand side is at most ε, we obtain:

$$d_{\mathrm{TV}}(T_n + s_\star) \le \varepsilon.$$

This completes the upper bound.

5.2 Lower Bound on Mixing Time

To establish the lower bound, we consider the function:

$$f(\sigma) := \sum_{v \in V} \sigma(v),$$

which represents the total magnetization of a configuration σ.

We analyze the behavior of this function under two independent copies of the Glauber dynamics, $(X_t^+)_{t\ge 0}$, started from the all-plus configuration, and $(X_t)_{t\ge 0}$, started from the stationary distribution.

Since the stationary distribution is symmetric, we have $\mathbb{E}[f(X_t)] = 0$ for all $t \ge 0$, while $\mathbb{E}[f(X_t^+)] = nm_t$, where m_t is the average magnetization at time t, by transitivity.

Let $t_\star^- = T_n - s_\star^-$. Using the upper bound on magnetization decay, we get:

$$\mathbb{E}[f(X_{t_\star^-}^+)] = nm_{t_\star^-} \ge e^{(1-\beta d)s_\star^-}\sqrt{n} \ge e^{s_\star^-/2}\sqrt{n} \quad \forall\, \beta \le 1/2d.$$

To control the variance of $f(X_{t_\star^-})$ and $f(X_{t_\star^-}^+)$, we use the following lemma:

Lemma 5.6 ([2], Claim 3.4)**.** *For some constants $\beta_0 = \beta_0(d) > 0$ and $\gamma = \gamma(d) > 0$, if $\beta < \beta_0$, then for any $t > 0$, we have:*

$$\sum_u \mathrm{Cov}(X_t(u), X_t(v)) \le \gamma \quad \textit{and} \quad \sum_u \mathrm{Cov}(X_t^+(u), X_t^+(v)) \le \gamma \quad \textit{for all } v \in V.$$

This immediately gives:

$$\max\{\mathrm{Var}(f(X_{t_\star^-})), \mathrm{Var}(f(X_{t_\star^-}^+))\} \le \gamma n.$$

We now apply the following lemma to use f as a distinguishing statistic:

Lemma 5.7 ([4], Proposition 7.9)**.** *Let μ and ν be two distributions on a finite set Ω, and let $f : \Omega \to \mathbb{R}$ be a function. Let $\sigma^2 := \max\{\mathrm{Var}_\mu(f), \mathrm{Var}_\nu(f)\}$. If $|\mathbb{E}_\mu(f) - \mathbb{E}_\nu(f)| \geq r\sigma$, then:*

$$\|\mu - \nu\|_{\mathrm{TV}} \geq 1 - \frac{8}{r^2}.$$

Proof. Assume without loss of generality that $\mathbb{E}_\mu(f) \leq \mathbb{E}_\nu(f)$, and define the interval:

$$A = \left(\mathbb{E}_\mu(f) + \frac{r\sigma}{2}, \infty\right).$$

By Chebyshev's inequality:

$$\mu(f^{-1}(A)) \leq \mathbb{P}_\mu\left(|f(X) - \mathbb{E}_\mu(f)| > \frac{r\sigma}{2}\right) \leq \frac{4}{r^2},$$

$$\nu(f^{-1}(A)) \geq \mathbb{P}_\nu\left(f(X) > \mathbb{E}_\nu(f) - \frac{r\sigma}{2}\right) \geq 1 - \mathbb{P}_\nu\left(|f(X) - \mathbb{E}_\nu(f)| \geq \frac{r\sigma}{2}\right) \geq 1 - \frac{4}{r^2}.$$

Hence,

$$\|\mu - \nu\|_{\mathrm{TV}} \geq |\mu(f^{-1}(A)) - \nu(f^{-1}(A))| \geq 1 - \frac{8}{r^2}.$$

□

Applying this lemma to our context,

$$|\mathbb{E}[f(X^+_{t_\star^-})] - \mathbb{E}[f(X_{t_\star^-})]| \geq \frac{e^{s_\star^-/2}}{\sqrt{\gamma}} \cdot \sqrt{\gamma n} = r\sigma,$$

we obtain:

$$d_{\mathrm{TV}}(t_\star^-) \geq 1 - \frac{8\gamma}{e^{s_\star^-}} \geq 1 - \varepsilon,$$

for sufficiently large $s_\star^- = s_\star^-(\varepsilon)$. Hence proved.

Since both $s_\star$ and $s_\star^-$ depend only on ε, we have established the existence of a cutoff with a constant-size window for sufficiently small β.

6 Conclusion

In this project, we studied the cutoff phenomenon in the high-temperature Ising model on finite transitive graphs, where the symmetry of the graph allows for a relatively clean analysis using information percolation. However, a much deeper and more general result was established by Lubetzky and Sly, who proved that cutoff occurs throughout the entire subcritical regime ($\beta < \beta_c$) on arbitrary finite graphs [6].

Acknowledgements

This report marks the end of an engaging journey through the study of Markov chains and mixing times. I am grateful to my professors for suggesting this topic and for their unmatched support and patient guidance throughout the project.

I would especially like to thank Naman and Aishee for their helpful suggestions and for always being around to discuss ideas, clarify doubts, and share the occasional frustration. This project wouldn't have been the same without your company.

Here's to mixing well — in both probability and friendships!

References

[1] E. Lubetzky and A. Sly. An exposition to information percolation for the Ising model. *Annales de la Faculté des Sciences de Toulouse 24*, 2015.

[2] E. Lubetzky and A. Sly. Universality of cutoff for the Ising model. *Annals of Probability 45* , 2017.

[3] J. Miller and Y. Peres. Uniformity of the uncovered set of random walk and cutoff for lamplighter chains. *The Annals of Probability*, 2012.

[4] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov Chains and Mixing Times.* American Mathematical Society, 2009.

[5] R. Uehara. The Number of Connected Components in Graphs and Its Applications *IEICE Technical Report, COMP99-10*, 1999.

[6] E. Lubetzky and A. Sly, Information percolation and cutoff for the stochastic Ising model. *Journal of the American Mathematical Society (JAMS) 29*, 2016.