Glauber Dynamics for Ising Model on Convergent Dense Graph Sequences*

Rupam Acharyya[1] and Daniel Štefankovič[2]

1 Department of Computer Science, University of Rochester, Rochester, NY, USA
racharyy@cs.rochester.edu
2 Department of Computer Science, University of Rochester, Rochester, NY, USA
stefanko@cs.rochester.edu

Abstract

We study the Glauber dynamics for Ising model on (sequences of) dense graphs. We view the dense graphs through the lens of graphons [19]. For the ferromagnetic Ising model with inverse temperature β on a convergent sequence of graphs $\{G_n\}$ with limit graphon W we show fast mixing of the Glauber dynamics if $\beta\lambda_1(W) < 1$ and slow (torpid) mixing if $\beta\lambda_1(W) > 1$ (where $\lambda_1(W)$ is the largest eigenvalue of the graphon). We also show that in the case $\beta\lambda_1(W) = 1$ there is insufficient information to determine the mixing time (it can be either fast or slow).

1998 ACM Subject Classification G.3 Probability and Statistics

Keywords and phrases Spin systems, Glauber dynamics, Ising model, graphons

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2017.23

1 Introduction

Spin systems have been extensively studied in physics [11], mathematics [25], and machine learning [24]. An important and challenging computational question is efficiently sampling configurations from the distribution of a model (spin system). A popular sampling method (and the focus of our paper) is *Glauber dynamics* [11]. One of the most studied spin models is *Ising model* [14, 12]. Even though there is a polynomial-time algorithm to sample from the distribution of the ferromagnetic Ising model [13] it is still useful (for reasons of simplicity, generality, and speed) to study the Glauber dynamics for the model [15, 21]. A basic question is: what properties of the underlying graph and the temperature make the Glauber dynamics fast (or slow)? In the case of sparse graphs the dynamics was studied for, for example, $\mathbb{Z}^2$ (see, e.g., [20]), general bounded degree graphs [22], and graphs with bounded connective constant [27, 26].

In the case of dense graphs the dynamics was studied for the complete graph [15] (for more general models on the complete graph, see [6, 2]). Our goal is to understand the impact of the structural properties (analaogously to the connective constant) of the dense graphs and the speed of Glauber dynamics. We will view dense graphs through the lens of *graphons* [19] and use the notions of free energy of a spin system on a graphon [4]. We give a threshold for the inverse temperature below which Glauber dynamics is rapidly mixing

* This material is based upon work supported by the National Science Foundation under NSF grant CCF-1563757.

and above which the mixing is slow. This generalizes [15] from complete graphs to general dense graph sequences. We also show that at the critical point it is not possible to draw a conclusion about the mixing time for a convergent sequence of graphs just by looking at the limit graphon.

We obtain our lower bound results by studying the typical configurations of the model [23, 15]. A phase of a spin configuration denotes what fraction of vertices get what spin. The most probable (dominant) phases play an important role in influencing the speed of the Glauber dynamics. Intuitively, a unique dominant phase (at a high temperature) corresponds to fast mixing of Glauber dynamics, whereas multiple dominant phases (at a low temperature) correspond to slow mixing of Glauber (and other) dynamics (moving between phases requires the chain to move through a high energy barrier). The typical phases were previously studied, for example, to show slow mixing of Glauber dynamics [23, 15, 8] and to prove hardness results for sampling [10, 9, 28].

2 Background

2.1 Homogeneous Ferromagnetic Ising Model (with no external field)

Ising Model was introduced in 1920's by Lenz [14] and Ising [12]. Let $G = (V(G), E(G))$ be a finite graph. In a configuration of the model, each vertex is assigned a spin from the set $\{+1, -1\}$. The energy of a configuration σ, is specified by the Hamiltonian of the configuration

$$H(\sigma) = -\sum_{v\sim w} J(v,w)\sigma(v)\sigma(w),$$

where $v \sim w$ denotes v is a neighbor of w in G and $J(v, w)$ denotes the interaction strength between vertices v and w.

We study homogeneous ferromagnetic Ising Model, that is, we assume $J(v, w) = 1$ for all $v, w \in V$. The probability measure μ, on the set of configurations $\Omega = \{+1, -1\}^{|V(G)|}$, for this model is given by,

$$\mu(\sigma) = \frac{e^{-\beta H(\sigma)}}{Z(\beta)},$$

where $\beta > 0$ is called the inverse temperature and $Z(\beta)$, the normalization factor is called the *partition function*.

This work focuses on dense graphs. We follow [15] in re-parameterizing the inverse temperature β as β/n, where $n = |V(G)|$. So the probability measure for dense graphs can be rewritten as,

$$\mu(\sigma) = \frac{e^{(\beta/n)\cdot S(\sigma)}}{Z(\beta)},$$

where $S(\sigma) = \sum_{v,w\in V, v\sim w} \sigma(v)\sigma(w)$.

2.2 Glauber Dynamics

In this paper we analyze Glauber Dynamics to sample from the distribution of the model. The (single site) Glauber Dynamics for the probability measure μ is defined by the following transition rule.

1. Pick a vertex v (also called site) uniformly at random from $V(G)$.
2. Change the spin of v with respect to the spins of its neighbors, i.e., in the new configuration, spin of v will be $+1$ with a probability of $p(\sigma, v)$, where

$$p(\sigma, v) := \frac{\exp(\frac{\beta}{n} S_v(\sigma))}{\exp(\frac{\beta}{n} S_v(\sigma)) + \exp(-\frac{\beta}{n} S_v(\sigma))},$$

and $S_v(\sigma) = \sum_{w \in V, v \sim w} \sigma(w)$.

We study the following (standard) notion of mixing time. The mixing time $\tau_{mix}(\varepsilon)$ of a Markov chain with state space Ω, transition matrix P and stationary distribution π is

$$\tau_{mix}(\varepsilon) = \max_{X_0 \in \Omega} \min\{t : d_{TV}(P^t(X_0, \cdot), \pi) \leq \varepsilon\}.$$

Usually $\varepsilon = \frac{1}{4}$ or $\varepsilon = \frac{1}{2e}$ is used.

2.3 Convergent Sequence of Dense Graphs

We study sequences of dense graphs using notions of convergence defined in [3, 4].

Let G be a weighted graph with non-negative vertex weights α_v that sum to 1 and edge weights $\beta_{uv} \in [0, 1]$. Let G' be another weighted graph with non-negative vertex weights α'_i that sum to 1 and edge weights $\beta'_{ij} \in [0, 1]$. Let $\chi(G, G')$ be the set of fractional overlays between G and G', where a fractional overlay (between G and G') is $X \in \mathbb{R}_{\geq 0}^{V(G) \times V(G')}$ such that $\sum_i X_{vi} = \alpha_v(G)$ and $\sum_v X_{vi} = \alpha'_i(G')$. The *cut distance* between G and G' is (see [19])

$$\delta_\square(G, G') = \min_{X \in \chi(G, G')} d_\square(G, G', X), \tag{1}$$

where

$$d_\square(G_n, G, X) = \max_{Q, R \subset V(G) \times V(G')} \left| \sum_{(v,i) \in Q, (u,j) \in R} X_{vi} X_{uj} (\beta_{uv} - \beta'_{ij}) \right|. \tag{2}$$

The *free energy* of an Ising model with parameter β/n for a dense graph G_n is defined as follows (see [17]).

$$\hat{\mathcal{F}}(G_n, \beta) = -\frac{1}{|V(G_n)|} \ln Z(G_n, \beta),$$

where $Z(G_n, \beta) = \sum_{\sigma : V(G_n) \to \{+1, -1\}} \exp(\frac{1}{n} \sum_{(u,v) \in E(G_n)} \beta \sigma(u) \sigma(v))$.

Microcanonical free energy is a more detailed version of free energy – we compute the free energy for each phase (by phase we mean the fraction of the vertices with positive spin), formally defined for $a \in [0, 1]$ as follows (see [17]):

$$\hat{\mathcal{F}}_a(G, \beta) = -\frac{1}{n} \ln Z_a(G, \beta),$$

where $Z_a(G, \beta) = \sum_{\sigma \in \Omega_a(G)} \exp(\frac{\beta}{n} \sum_{(u,v) \in E(G)} \sigma(u) \sigma(v))$ and

$$\Omega_a(G) = \{\sigma : V(G) \to \{-1, +1\} \,\big|\, \big||\sigma^{-1}(\{+1\})| - a|V(G)|\big| \leq 1\}.$$

In [4] it has been shown that convergence w.r.t. cut metric implies convergence w.r.t. microcanonical free energy and free energy (they also show converse if one has convergence w.r.t. microcanonical free energies for all spin models).

2.4 Limit Object of Convergence: Graphon

The limits of the convergence w.r.t. the cut norm are graphons [19].

▶ **Definition 1** (Graphon, [19]). A graphon W is a symmetric measurable function $W : [0,1]^2 \to [0,1]$. (The symmetry means $W(x,y) = W(y,x)$ for all $x, y \in [0,1]$.)

The simplest graphons correspond to step functions with finitely many steps.

▶ **Definition 2** (Step Graphon, [19]). Let $S_1, \dots, S_k$ be a disjoint decomposition of $[0,1]$ into intervals for some finite k and let P be a symmetric $k \times k$ matrix with entries from $[0,1]$. A function $U : [0,1]^2 \to [0,1]$ is a step graphon with value matrix P if $\forall i, j$ and $\forall (x,y) \in S_i \times S_j$ $U(x,y) = P_{ij}$. We call $\alpha_1, \dots, \alpha_k$ the step sizes of the step graphon, where $\alpha_i = |S_i|$ for $i \in \{1, \dots, k\}$.

Given a weighted graph H (with $|V(H)| = n$) a step graphon W_H can be naturally constructed as follows. Let $S_1, \dots, S_n$ be disjoint sub-intervals of $[0,1]$ such that S_i is of size α_i, where α_i is the weight of the vertex $i \in V(G)$. For $x \in S_i$ and $y \in S_j$ we let $W_H(x,y) = \beta_{ij}$, where β_{ij} is the weight of the edge between vertices i and j (if there is no edge between i and j we let $\beta_{ij} = 0$).

▶ **Definition 3** (Eigenvalue of a Graphon). [17] Given a graphon W, consider the following operator $T_W : L_2[0,1] \to L_1[0,1]$:

$$(T_W f)(x) = \int_{[0,1]} W(x,y) f(y)\, dy.$$

The operator T_W has discrete spectrum, i.e., a multi-set of real nonzero eigenvalues $\lambda_1, \lambda_2, \dots$ (sorted in the non-increasing order by their absolute value), such that $\lambda_n \to 0$. We call these the eigenvalues of the graphon W. The eigenvalue with highest absolute value is denoted $\lambda_1(W)$.

The notions of cut distance, free energy, and micro-canonical free energy extend from graphs to graphons (see [17]).

The *cut distance* between two graphons is:

$$\delta_\square(W,U) := \inf_\phi \|W^\phi - U\|_\square := \inf_\phi \sup_{S,T} \left| \int_{S\times T} W^\phi(x,y) - U(x,y)\, dx\, dy \right|,$$

where $\phi : [0,1] \to [0,1]$ is a measure preserving function and $W^\phi(x,y) = W(\phi(x), \phi(y))$. The cut distance between a graph G and a step graphon W is denoted by $\delta_\square(G,W) = \delta_\square(W_G, W)$.

The *free energy* of a graphon is defined as

$$\mathcal{F}(W,\beta) = \inf_{m:[0,1]\to[-1,1]} \mathcal{E}(W,\beta,m), \tag{3}$$

where

$$\mathcal{E}(W,\beta,m) = -\frac{\beta}{2}\langle m, T_W m\rangle - \mathrm{Ent}(m), \tag{4}$$

and

$$\mathrm{Ent}(m) = -\int_0^1 \frac{1}{2}(1-m(x)) \log(\frac{1}{2}(1-m(x)))\, dx - \int_0^1 \frac{1}{2}(1+m(x)) \log(\frac{1}{2}(1+m(x)))\, dx,$$

and

$$\langle m, T_W m\rangle = \int_{[0,1]^2} W(x,y)m(x)m(y)\,dx\,dy.$$

The *microcanonical free energy* of a graphon with phase $a \in [0,1]$ is

$$\mathcal{F}_a(W,\beta) = \inf_{m:[0,1]\to[-1,1] \text{ and } \int_{[0,1]} m(x)\,dx = a} \mathcal{E}(W,\beta,m), \tag{5}$$

where $\mathcal{E}(W,\beta,m)$ is defined as in (4).

A sequence of dense graphs $\{G_n\}$ is said to be convergent to a graphon W if $\delta_\square(G_n, W) \to 0$. For a sequence of dense graphs $\{G_n\}$ converging to a graphon W it has been shown [3, 4] that the free energy and microcanonical free energy of the dense graphs converge to the free energy and microcanonical free energy of the graphon.

▶ **Proposition 4** ([4]). *Suppose $\{G_n\}$ be a sequence of dense graphs convergent to a graphon $W : [0,1]^2 \to [0,1]$. Then*
1. $\hat{\mathcal{F}}(G_n,\beta) \to \mathcal{F}(W,\beta)$.
2. $\forall a \in [0,1]$, $\hat{\mathcal{F}}_a(G_n,\beta) \to \mathcal{F}_a(W,\beta)$.

3 Main Results and Related Works

3.1 Results for Mixing Time

The Glauber dynamics for Ising model has been extensively studied in [13, 15, 22]. The dynamics is well understood when the graph has bounded degree [13, 21, 1]. In the dense scenario the dynamics has been analyzed for the complete graph [15] (so-called mean field model). The complete graph corresponds to graphon with $W(x,y) = 1$ (for all $x, y \in [0,1]$). Our goal here is to extend this work to general graphons (we aim to understand the connection between the mixing time, the inverse temperature, and the structure of the graphon).

For the Ising model on the complete graph [15] show that the mixing of Glauber Dynamics is fast when $\beta < 1$ and it is exponentially slow when $\beta > 1$. This threshold behavior extends to convergent dense graph sequences – we provide a threshold for the parameter β, such that, below the threshold mixing of Glauber dynamics is fast and above the threshold mixing is slow (our result matches the threshold for the complete graph – the threshold is $\beta = 1/\lambda_1(W)$ and for complete graph $\lambda(W) = 1$). Formally we have the following results.

▶ **Theorem 5.** *Consider a homogeneous ferromagnetic Ising model (with no external field) with inverse temperature β and a graphon W. If $\{G_n\} \to W$, then the mixing time of the Glauber Dynamics for Ising model on G_n satisfies the following:*
1. *If $\lambda_1(W) \cdot \beta < 1$ then $\tau_{mix}(G_n) = O(n \log(n))$.*
2. *If $\lambda_1(W) \cdot \beta > 1$ then $\tau_{mix}(G_n) = e^{\Omega(n)}$.*

Remark (mixing in critical case): In the above theorem we haven't stated any result for the critical temperature, i.e., when $\lambda_1(W)\beta = 1$. This is because, at the critical temperature one cannot draw conclusion about the mixing time for a convergent sequence of graphs just by looking at the limit graphon. We show examples of two different graph sequences which converge to the same graphon, even though at critical temperature mixing is fast for one sequence and slow for the other. These examples are discussed in Section 9.

3.2 Results for Phase Diagram

A phase α of the Ising model is the set of configurations which has αn fraction of vertices with $+1$ spin. The weight of a phase is the value of the partition function when restricted to the configurations with the given phase signature. The phase which has maximum weight is called the *dominant phase*. It has been seen earlier that when the model is studied on a graph, the phase diagram of the model changes with different values of the parameter β. For example, when Ising model is studied on complete graphs the model exhibits an unique dominant phase if $\beta < 1$ and it has multiple dominant phases when $\beta > 1$. It has been shown in [15] that coexistence of multiple dominant phases implies slow mixing, because to get from one phase to another it requires to pass through a high free energy barrier. Hence studying phase diagram for spin models has been focus of numerous previous studies [23, 10, 9]. The goal of these studies was to understand the speed of the dynamics. As we know from Section 2.4 that the free energy is defined as the negative of the logarithm of the partition function, to find the dominant phase we need to find the phase which minimizes the free energy. In this paper our interest is to study the behavior of the phase transition on a sequence of graphs. For this purpose we study the behavior of the free energy on the limit graphon, i.e., we try to find for what values of β there is an unique minimizer (equivalently unique dominant phase) in the expression for the free energy. Formally we have the following theorem.

▶ **Theorem 6.** *Consider a graphon W and the free energy function for the graphon W with respect to the inverse temperature parameter β is defined as in* (4)*.*

1. *If $\lambda_1(W) \cdot \beta < 1$ then the function $\mathcal{E}(W, \beta, m)$ has unique*[1] *local minimum.*
2. *If $\lambda_1(W) \cdot \beta > 1$ then the function $\mathcal{E}(W, \beta, m)$ has multiple*[2] *global minima.*

4 Organization

In Section 5 we prove for a convergent graph sequence than one can align the graphs in the sequence with a step graphon (that is close to the limit graphon) in such a way that most vertices have same neighborhood statistics as the step graphon. This property will later be used to prove the upper bound result of Theorem 5. Next in Section 6 we establish the phase digram for different values of β (Theorem 6). The result about phase diagram is an important tool to prove the lower bound of mixing time of Theorem 5. Finally in Section 7 we prove the upper bound result at high temperature and in Section 8 we prove that the mixing is slow on the graphs in the sequence at low temperature. All the remaining proofs can be found in the Appendix.

5 Labeling Graphs in a Convergent Graph Sequence

In this section we will deduce some properties of convergent graph sequences which will be used to prove the upper bound result of Theorem 5.

▶ **Definition 7** (GOOD and BAD vertices)**.** Let U be a step graphon with k steps, value matrix P, and step sizes $\alpha_1, \ldots, \alpha_k$. Let G be a graph and let $\phi : V(G) \to \{1, \ldots, k\}$ be a

[1] By unique we mean unique up to measurability, i.e., m_1 and m_2 are two solutions then the set where they differ has measure zero

[2] By multiple we mean there exists at least two functions m_1 and m_2 such that the set where they differ has measure greater than zero

labeling. Let v be a vertex of G and let $i = \phi(v)$. We call the vertex v to be GOOD with ε tolerance if for all $j \in \{1, \ldots, k\}$,

$$|\{w \mid w \sim v; \phi(w) = j\}| \leq (P_{ij}\alpha_j + \varepsilon)n.$$

Otherwise we call the vertex to be BAD w.r.t. ε tolerance.

▶ **Definition 8** (Proper Labeling). Let G be a graph and U be a step graphon. A labeling $\phi : V(G) \to \{1, \ldots, k\}$ is said to be proper up to ε tolerance w.r.t. U if there are at most εn many BAD vertices w.r.t. ε tolerance.

With the above definitions we can now state the following lemma.

▶ **Lemma 9.** *Let $\{G_n\}$ be a sequence of graphs such that $G_n \to W$ for some graphon W. Then for any $\varepsilon > 0$ there exists $k = k(\varepsilon), n_0 = n_0(\varepsilon)$ and a step graphon U with k steps such that $\delta_\square(W, U) \leq \varepsilon$ and such that $\forall n \geq n_0$ we have that G_n has a proper labeling up to ε tolerance w.r.t. U.*

To prove the Lemma 9 we will first prove an easier version of the lemma when the limit graphon is a step graphon.

▶ **Lemma 10.** *Let $\{G_n\}$ be a sequence of graphs such that $G_n \to U$ for some step graphon U. Then for any $\varepsilon > 0$ there exists $n_0 = n_0(\varepsilon)$ such that $\forall n \geq n_0$ we have that G_n has a proper labeling up to ε tolerance w.r.t. U.*

Proof of Lemma 10. We know that $G_n \to U$ implies that for given $\varepsilon > 0$ there exists n_0 such that $\forall n \geq n_0$,

$$\delta_\square(G_n, U) \leq \varepsilon^2/2. \tag{6}$$

Since every step graphon can be viewed as arising from a weighted graph G by the construction shown in in Section 2.3, we will, w.l.o.g., assume $U = W_G$. Hence $\delta_\square(G_n, U) = \delta_\square(G_n, W_G) = \delta_\square(G_n, G)$. Now for two weighted graphs we have

$$\delta_\square(G_n, G) = \min_{X \in \mathcal{X}(G_n, G)} d_\square(G_n, G, X), \tag{7}$$

where X is a fractional overlay, i.e., $\sum_i X_{vi} = \frac{1}{n}$ and $\sum_v X_{vi} = \alpha_i(G)$ and

$$d_\square(G_n, G, X) = \max_{Q, R \subset V(G_n) \times V(G)} \left| \sum_{(v,i) \in Q, (u,j) \in R} X_{vi} X_{uj} (1 - P_{ij}) \right|. \tag{8}$$

Note that we give weight $\frac{1}{n}$ to each vertex $v \in V(G_n)$ (as G_n is originally unweighted). The 1 in (8) is the weight of the edge $(u, v) \in E(G_n)$. Similarly $\alpha_i(G)$ is the weight of the vertex $i \in V(G)$ and P_{ij} is the weight of the edge $(i, j) \in E(G)$. Now let X be the fractional overlay which minimizes the cut distance. We assign the label ϕ of a vertex $v \in V(G_n)$ from the distribution $\{nX_{vi}\}_i$, i.e., $\phi(v) = i$ with probability nX_{vi}. Note that for any vertex,

$$\mathbb{E}\Big[\big|\{w | w \sim v; \phi(w) = j\}\big|\Big] = n \sum_{w | w \sim v} X_{wj}.$$

Now we call a vertex v to be *dangerous* for (i, j) if $\phi(v) = i$ and

$$\sum_{w | w \sim v} X_{wj} \geq \alpha_j P_{ij} + \frac{\varepsilon}{2}. \tag{9}$$

Now we will show that there are not too many such dangerous vertices.

Bound on number of Dangerous Vertices: First we fix i and j. Let Q be the set of all dangerous vertices for (i, j) and R be the set of all vertices $w \in V(G)$ with label j . Then from (9), (6) and (8) we have:

$$\sum_{v\in Q} X_{vi}(\alpha_j P_{ij} + \frac{\varepsilon}{2}) \leq \sum_{v\in Q} X_{vi} \sum_{w\sim v} X_{wj} \leq \sum_{v\in Q} X_{vi} \sum_{w\in R} X_{wj}P_{ij} + \frac{\varepsilon^2}{2} = \sum_{v\in Q} X_{vi}\alpha_j P_{ij} + \frac{\varepsilon^2}{2}. \tag{10}$$

Hence from (10) we have $\sum_{v\in Q} X_{vi} \leq \varepsilon$. So from Chernoff Bound w.h.p. the number of dangerous vertices are at most εn. Next we look at the vertices which are not dangerous for any (i, j), i.e., if $\phi(v) = i$, then for all j we have

$$\sum_{w|w\sim v} X_{wj} \leq \alpha_j P_{ij} + \frac{\varepsilon}{2}. \tag{11}$$

We now move on to prove that the probability there there exists too many BAD vertices is very low. We now use Y_v as an indicator variable to denote whether the vertex v is BAD or not. Hence it is enough to bound $Pr[\sum_v Y_v \geq \varepsilon n]$. Now from Markov's inequality we have:

$$Pr[\sum_v Y_v \geq \varepsilon n] \leq \frac{\sum_v \mathbb{E}[Y_v]}{\varepsilon n}. \tag{12}$$

Hence we now need to bound $\mathbb{E}[Y_v]$. Again using Markov's inequality we have

$$\begin{aligned}\mathbb{E}[Y_v] &= Pr[v \text{ is BAD}]\\ &= \sum_i Pr[v \text{ gets label } i] \cdot Pr[\exists j \ni |\{w|w \sim v; \phi(w) = j\}| \geq (P_{ij}\alpha_j + \varepsilon)n \Big| \phi(v) = i]\\ &\leq \sum_i nX_{vi} \sum_j Pr[|\{w|w \sim v; \phi(w) = j\}| \geq (P_{ij}\alpha_j + \varepsilon)n]. \end{aligned} \tag{13}$$

Using Chernoff-Hoeffding bound for any non-dangerous vertex v we have,

$$Pr[|\{w|w \sim v; \phi(w) = j\}| \geq (P_{ij}\alpha_j + \varepsilon)n] \leq \exp(-n\varepsilon^2/4). \tag{14}$$

Now from (12) and (13) we have:

$$Pr[\sum_v Y_v \geq \varepsilon n] \leq \frac{kn\exp(-n\varepsilon^2/4)}{\varepsilon n} = \frac{k}{\varepsilon}\exp(-n\varepsilon^2/4).$$

Hence we have the lemma. ◀

Proof of Lemma 9. As $G_n \to W$ we have for given $\varepsilon > 0$ there exists n_0 such that

$$\delta_\square(G_n, W) \leq \frac{\varepsilon^2}{4}. \tag{15}$$

Also from [17] we have for any graphon W we have that $\exists$ a step function U with k steps (where k is sufficiently large) such that

$$\delta_\square(U, W) \leq \sqrt{\frac{2}{\log_2 k}}\|U\|_2 \leq \frac{\varepsilon^2}{4}. \tag{16}$$

Hence from (15) and (16) we have the following analog of (6)

$$\delta_\square(U, G_n) = \delta_\square(G_U, G_n) \leq \frac{\varepsilon^2}{2}, \tag{17}$$

where G_U is a graph on k vertices. Now the remainder of the proof of the lemma is identical to the proof of Lemma 10. ◀

6 Phase Diagram

In this section we will prove Theorem 6. As free energy of the model is the infimum over the set of all measurable functions from $[0,1]$ to $[-1,1]$ (defined in Section 2.4) we first need to prove that there exists some such function at which the infimum is achieved. Then we will analyze its properties.

▶ **Lemma 11.** *Let $\mathcal{E}(W,\beta,m)$ be the function as defined by (4). Then the following infimum is attained for some measurable function m:*

$$\inf_{m:[0,1]\to[-1,1]} \mathcal{E}(W,\beta,m). \tag{18}$$

Proof of Lemma 11 has been deferred to Appendix. Assuming the existence we now move on to prove Theorem 6.

Proof of Theorem 6.

Case I: $\boldsymbol{\lambda_1(W)\beta < 1}$. In this case we will prove that the functional $m \mapsto \mathcal{E}(W,\beta,m)$ is strictly convex. Then there will be an unique minimum up to measurability (strict convexity implies unique minimum because if there were two minima available then by strict convexity there average will have a strictly lesser functional value which is a contradiction). Formally we prove the following lemma.

▶ **Lemma 12.** *$\mathcal{E}(W,\beta,m)$ is defined as in (3). Then for all $0 \le \alpha \le 1$ and for all measurable functions m,p from $[0,1]$ to $[-1,1]$ we have :*

$$\mathcal{E}(W,\beta,(1-\alpha)m+\alpha p) < (1-\alpha)\mathcal{E}(W,\beta,m) + \alpha\mathcal{E}(W,\beta,p).$$

whenever $\lambda_1(W)\beta < 1$.

We prove the above lemma in the Appendix.

Case II: $\boldsymbol{\lambda_1(W)\beta > 1}$. For the purpose of the proof we slightly re-parameterize the functions. In particular we define $\rho(x) := \frac{1}{2}(m(x)+1)$. Hence the optimization problem can be written as:

$$\inf_{\rho:[0,1]\to[0,1]} \mathcal{E}(W,\beta,2\rho-1). \tag{19}$$

If two measurable functions $f,g:[0,1]\to[0,1]$ differ on a set of measure zero we write $f \underset{m}{\approx} g$. Now we define a new set $S = \{\rho : [0,1]\to[0,1] \mid \int_{[0,1]} \rho(x)\,dx = \frac{1}{2}\}$. We will show that the minimum doesn't lie in the set S. For the function $\rho(x) = 1/2$ everywhere we argue that it cannot be the minimum by a local perturbation argument. For all the other functions $\rho \in S$ we use the following transformation to produce a function with a smaller value.

▶ **Definition 13.** Given a function $\rho \in S$ we define another measurable function $\hat{\rho} : [0,1] \to [0,1]$ as follows:

$$\hat{\rho}(x) = \begin{cases} \rho(x) & \textit{if } \rho(x) \ge \frac{1}{2}, \\ 1-\rho(x) & \textit{otherwise.} \end{cases}$$

Now we have the following lemma for $\hat{\rho}$ the proof of which has been deferred to Appendix.

▶ **Lemma 14.** *If* $\rho \in S = \{\rho : [0,1] \to [0,1] | \int_{[0,1]} \rho(x)\,dx = \frac{1}{2}\} \setminus \{\rho : [0,1] \to [0,1] | f \underset{m}{\approx} \rho$ *and* $\rho(x) = \frac{1}{2} \forall x \in [0,1]\}$ *and* $\hat{\rho}$ *is defined as in Definition 13, then*

$$\mathcal{E}(W, \beta, 2\hat{\rho} - 1) < \mathcal{E}(W, \beta, 2\rho - 1). \tag{20}$$

It remains to rule out the function $\rho(x) = \frac{1}{2}$ (for $x \in [0,1]$), that is, to show that this function is also not an minimum point for $\mathcal{E}(W, \beta, 2\rho - 1)$. More formally we have the following lemma.

▶ **Lemma 15.** *Consider the following minimization problem from* (18)

$$\inf_{\rho:[0,1]\to[0,1]} \mathcal{E}(W, \beta, 2\rho - 1).$$

If $\lambda_1(W)\beta > 1$ *then* $\rho : [0,1] \to [0,1] | \rho(x) = \frac{1}{2} \forall x$ *is not a minimizer of the optimization problem.*

Hence from the Lemma 14 and 15 we have that the minimizers of $\mathcal{E}(W, \beta, 2\rho - 1)$ is not in the set $S = \{\rho : [0,1] \to [0,1] | \int_{[0,1]} \rho(x)\,dx = \frac{1}{2}\}$. Note that $\mathcal{E}(W, \beta, 2\rho - 1) = \mathcal{E}(W, \beta, 2(1-\rho)-1)$. Hence if ρ_{opt} is a minimizer of $\mathcal{E}(W, \beta, 2\rho - 1)$ so is $1 - \rho_{opt}$. Hence the optimization problem has multiple minima. ◀

7 Upper Bound for the Mixing Time

We will now prove the upper bound result stated in the Theorem 5 using path coupling, a well known proof technique for bounding mixing time. We state a lemma from [5] which will be used for the proof.

▶ **Lemma 16.** *[5] Let* $\mathcal{X}$ *be a Markov chain. Let* $G_{\mathcal{X}}$ *be the graph of the Markov chain. Let* ℓ *be a length function on the edges of* $G_{\mathcal{X}}$ *such that* $\ell(x, y) \geq 1$ *for each edge* $\{x, y\} \in E(G_{\mathcal{X}})$*. This then naturally extends to a metric (which we also denote by* ℓ*), where* $\ell(x, y)$ *is the length of the shortest path from* x *to* y*. Suppose that for each edge* $(x, y) \in G_{\mathcal{X}}$ *there exists a coupling* (X_1, Y_1) *of* $P(x, \cdot)$ *and* $P(y, \cdot)$ *such that the following holds:*

$$\mathbb{E}_{x,y}[\ell(X_1, Y_1)] \leq \ell(x, y)e^{-\alpha}.$$

Then

$$t_{mix}(\eta) \leq \left\lceil \frac{-\log(\eta) + \log(\operatorname{diam}(\mathcal{X}))}{\alpha} \right\rceil,$$

where $\operatorname{diam}(\mathcal{X}) = \max_{x,y \in G_{\mathcal{X}}} \ell(x, y)$ *is the diameter of* $G_{\mathcal{X}}$

Now we prove the main theorem about fast mixing in high temperature.

Proof of Theorem 5.1. From Lemma 9 we know that for any $\varepsilon > 0$ for any sufficiently large n the graph G_n can be properly labeled up to ε tolerance (call) w.r.t. some step graphon U such that U is ε-close to the limit graphon W, i.e., $\delta_{\square}(U, W)\varepsilon$. Let's call the labeling as ϕ. Let k be the number of steps in U and $\alpha_1, \ldots, \alpha_k$ be the step sizes. Now we define the length function ℓ to be used in the path coupling argument.

Defining the Distance: For a vertex $v \in V(G)$ we define the following quantity,

$$\hat{d}_v = \begin{cases} d_{\phi(v)} & \text{if } v \text{ is GOOD w.r.t. } \phi, \\ \frac{1}{\lambda_1(U)} \sum_j d_j & \text{if } v \text{ is BAD w.r.t. } \phi, \end{cases}$$

where $(d_1, \ldots, d_k)$ is the eigenvector corresponding to the largest eigenvalue $(\lambda_1(U))$ of the step graphon U, where the eigenvector is scaled so that $d_i \geq 1$. Note that if for all i we have $d_i \geq 1$ then $\frac{1}{\lambda_1(U)} \sum_j d_j \geq 1$. Now the distance between any two arbitrary configurations σ' and τ' is defined as:

$$\ell(\sigma', \tau') = \sum_{v \in V(G) \text{ and } \sigma'(v) \neq \tau'(v)} \hat{d}_v.$$

Choice of ε: Now let $\varepsilon_0 > 0$ be such that $(\lambda_1(W) + \varepsilon_0)\beta = 1$. We will choose U such that $|\lambda_1(W) - \lambda_1(U)| \leq \varepsilon_0/4$ and take $\varepsilon > 0$ such that $\varepsilon d_{\text{bad}}(1 + \lambda_1(U)) = (\min_j d_j)\frac{\varepsilon_0}{4}$, where $d_{\text{bad}} = \frac{1}{\lambda_1(U)} \sum_j d_j$.

Defining the Path Coupling: Let σ, τ be two configurations such that the two configurations differ only at v and $\sigma(v) = -1$ and $\tau(v) = +1$. Now we describe a coupling (X, Y) such that X starts with σ and Y starts with τ.

- Pick one vertex w u.a.r from V.
- If $w \notin \mathcal{N}(v)$ then update the spin of w in both X and Y with transition probability specified by the dynamics [in Section 2.2].
- If $w \in \mathcal{N}(v)$ then pick a number $Z \in [0, 1]$ and set

  $$X_1(w) = \begin{cases} +1, & \text{if } Z \leq p(\sigma, v), \\ -1, & \text{otherwise,} \end{cases}$$

  and

  $$Y_1(w) = \begin{cases} +1, & \text{if } Z \leq p(\tau, v) \\ -1, & \text{otherwise,} \end{cases}$$

  where

  $$p(\sigma, v) = \frac{e^{\beta S_v(\sigma)}}{e^{\beta S_v(\sigma)} + e^{-\beta S_v(\sigma)}}, \tag{21}$$

  and $S_v(\sigma) = \sum_{v \sim w} \sigma(w)$.

From the definition of the coupling we can see that the disagreement of the two configurations spreads further with probability $p(\tau, v) - p(\sigma, v)$. We have the following upper bound on the probability of spreading disagreement.

▶ **Claim 17** (see, e.g., [16]). *Consider Ising model on a dense graph G with inverse temperature β and let σ, τ be two configurations such that the two configurations differ only at v and $\sigma(v) = -1$ and $\tau(v) = +1$. Also $p(\sigma, v)$ is defined as in (21). Then we have*

$$p(\tau, v) - p(\sigma, v) \leq \tanh(\frac{\beta}{n}).$$

Now we analyze the expected decrease of the coupling distance in two cases to satisfy the hypothesis of the Lemma 16.

Case I: v is GOOD: As we can see from Lemma 9 if v is a GOOD vertex then we have number of neighboring vertices of v with label j is $\leq (P_{ij}\alpha_j + \varepsilon)n$. As we have seen in the coupling we choose a vertex w u.a.r., i.e., w.p. $\frac{1}{n}$. Now we have the following cases:

- If $w = v$, then $d(X_1, Y_1) = 0$.
- If $w \notin \mathcal{N}(v) \cup \{v\}$, then $d(X_1, Y_1) = d_i$.
- If $w \in \mathcal{N}(v)$ and w gets label j by the labeling, then w.p. $p(\tau, v) - p(\sigma, v)$,
  - $\ell(X_1, Y_1) = d_i + d_j$, if w is GOOD,
  - $\ell(X_1, Y_1) = d_i + d_{\text{bad}}$, if w is BAD.

where $d_{\text{bad}} = \frac{1}{\lambda_1(U)} \sum_j d_j$. Also from Lemma 9 there are at most εn many BAD vertices. So from Claim 17 and the above discussion we have

$$\begin{aligned}
\mathbb{E}[\ell(X_1, Y_1)] &\leq d_i(1 - \frac{1}{n}) + \frac{1}{n} \cdot \tanh(\beta/n)\Big[\sum_j (P_{ij}\alpha_j + \varepsilon)n \cdot d_j + \varepsilon n \cdot d_{\text{bad}}\Big] \\
&\leq d_i(1 - \frac{1}{n}) + \frac{1}{n} \cdot \beta\Big[\sum_j (P_{ij}\alpha_j + \varepsilon) \cdot d_j + \varepsilon \cdot d_{\text{bad}}\Big] \\
&= d_i(1 - \frac{1}{n}) + \frac{1}{n} \cdot \beta\Big[\sum_j \big(P_{ij}\alpha_j d_j + \varepsilon d_{\text{bad}}(1 + \lambda(U)\big)\Big] \\
&= d_i(1 - \frac{1}{n}) + \frac{1}{n} \cdot \beta\Big[\lambda_1(U) d_i + \varepsilon d_{\text{bad}}(1 + \lambda(U))\Big] \\
&\leq d_i \exp\Big(- \frac{1}{n}\Big(1 - \beta\big(\lambda_1(U) + \varepsilon \frac{d_{\text{bad}}}{d_i}(1 + \lambda_1(U))\big)\Big)\Big).
\end{aligned} \tag{22}$$

By the choice of ε we then have

$$\beta\big(\lambda_1(U) + \varepsilon \frac{d_{\text{bad}}}{d_i}(1 + \lambda_1(U)\big) \leq \beta\big(\lambda_1(U) + \frac{\varepsilon_0}{4}\big) \leq (\lambda_1(W) + \frac{\varepsilon_0}{2})\beta < 1. \tag{23}$$

Hence using (23) in (22) we have

$$\mathbb{E}[\ell(X_1, Y_1)] \leq d_i \exp(-\frac{1}{n}c),$$

where $c = 1 - \beta\big(\lambda_1(U) + \varepsilon \frac{d_{\text{bad}}}{d_i}(1 + \lambda_1(U))\big) > 0$ and so from Lemma 16 we have the theorem.

Case II: v is BAD: In this case we will consider that v is BAD w.r.t. the labeling and so it can be connected to all the vertices in the worst case. Using similar discussion for case I we have:

- If $w = v$, then $d(X_1, Y_1) = 0$.
- If $w \notin \mathcal{N}(v) \cup \{v\}$, then $d(X_1, Y_1) = d_{\text{bad}}$.
- If $w \in \mathcal{N}(v)$ and w gets label j by the labeling, then w.p. $p(\tau, v) - p(\sigma, v)$,
  - $\ell(X_1, Y_1) = d_{\text{bad}} + d_j$, if w is GOOD.
  - $\ell(X_1, Y_1) = d_{\text{bad}} + d_{\text{bad}}$, if w is BAD.

Similarly we have,

$$\begin{aligned}
\mathbb{E}[\ell(X_1, Y_1)] &\leq d_{\text{bad}}(1 - \frac{1}{n}) + \frac{1}{n} \cdot \tanh(\beta/n)\Big[\sum_j n d_j + \varepsilon n \cdot d_{\text{bad}}\Big] \\
&\leq d_{\text{bad}}(1 - \frac{1}{n}) + \frac{1}{n} \cdot \beta\Big[\sum_j d_j + \varepsilon \cdot d_{\text{bad}}\Big] \\
&\leq d_{\text{bad}}(1 - \frac{1}{n}) + \frac{1}{n} \cdot \beta \cdot d_{\text{bad}}\Big[\lambda_1(U) + \varepsilon\Big] \\
&\leq d_{\text{bad}} \exp\Big(- \frac{1}{n}\Big(1 - \beta\big(\lambda_1(U) + \varepsilon\big)\Big)\Big).
\end{aligned}$$

By the choice of ε we then have

$$\beta(\lambda_1(U)+\varepsilon) \leq \beta(\lambda_1(W)+\frac{\varepsilon_0}{2}) < 1.$$

Hence we will have the theorem from Lemma 16. ◀

8 Lower Bound for Mixing Time

Here we will prove the result about slow mixing of Theorem 5 using the well known conductance bound technique [7].

▶ **Lemma 18.** *[7] Let $\mathcal{M}$ be a Markov chain with state space Ω, transition matrix P, and stationary distribution μ. Let $A \subset \Omega$ such that $\mu(A) \leq \frac{1}{2}$, and $B \subset \Omega$ that forms a barrier in the sense $P_{ij} = 0$ for $i \in A \setminus B$ and $j \in A^{\mathsf{c}} \setminus B$. Then the mixing time of $\mathcal{M}$ is at least $\frac{\mu(A)}{8\mu(B)}$.*

To find such sets we look at the sets with given signature or phase. Formally we define

$$A_\alpha := \{\sigma \big| |\{v \in V(G) | \sigma(v) = +\}| = \alpha n\}. \tag{24}$$

Now let Z_α denotes the partition function with signature α. To apply the Lemma 18 we consider $A = A_{<\frac{1}{2}} = \bigcup\limits_{\alpha<\frac{1}{2}} A_\alpha$ and $B = A_{\frac{1}{2}}$. Trivially B is barrier between A and $A^{\complement}$. Now to show lower bound of $\frac{\mu(A)}{8\mu(B)}$ we give a lower bound on $\mu(A)$ and an upper bound on $\mu(B)$.

Lower bound on $\mu(A)$. Assume $\{G_n\}$ be a convergence sequence of dense graphs which converges to a graphon W, then the graphs also converge w.r.t. the microcanonical free energy, where microcanonical energy $\mathcal{F}_{\mathbf{a}}(W,\beta)$ is defined as

$$\mathcal{F}_{\mathbf{a}}(W,\beta) := \inf_{\rho:\alpha(\rho)=\mathbf{a}} \mathcal{E}(W,\beta,2\rho-1).$$

Now let's look at the free energy from (3):

$$\mathcal{F}(W,\beta) = \inf_{\rho:[0,1]\to[0,1]} \mathcal{E}(W,\beta,2\rho-1) = \mathcal{E}(W,\beta,2\rho^{opt}-1).$$

Now let's say we have $\int_{[0,1]} \rho^{opt}(x)\,dx = \alpha_c$ for some constant α_c (w.l.o.g., we can assume $\alpha_c < 1/2$). We denote $Z'_\alpha = Z(\beta)|_{\Omega_\alpha}$ and $Z_\alpha = Z(\beta)|_{A_\alpha}$, where Ω_α is defined in Section 2.3. Then from Proposition 4 we have :

$$\begin{aligned}
&\Big|\frac{1}{n}\log(Z'_{\alpha_c}) - \sup_{\int_{[0,1]}\rho(x)\,dx=\alpha_c} \big(-\mathcal{F}_{\alpha_c}(W,\beta)\big)\Big| < \varepsilon \\
&\Rightarrow \Big|\frac{1}{n}\log(Z'_{\alpha_c}) + \mathcal{E}(W,\beta,2\rho^{opt}-1)\Big| < \varepsilon \\
&\Rightarrow \frac{1}{n}\log(Z'_{\alpha_c}) > \mathcal{E}(W,\beta,2\rho^{opt}-1) - \varepsilon \\
&\Rightarrow \frac{1}{n}\log(Z_{<\frac{1}{2}}) > \frac{1}{n}\log(Z'_{\alpha_c}) > -\mathcal{E}(W,\beta,2\rho^{opt}-1) - \varepsilon \\
&\Rightarrow Z_{<\frac{1}{2}} > \exp(n(-\mathcal{E}(W,\beta,2\rho^{opt}-1)-\varepsilon)).
\end{aligned} \tag{25}$$

where we denote $Z_{<\frac{1}{2}} = \cup_{\alpha<\frac{1}{2}} Z_\alpha$.

Upper bound on $\mu(B)$: Suppose $\sup_{\int_{[0,1]}\rho(x)=\frac{1}{2}} \mathcal{E}_{\frac{1}{2}}(W,\beta,2\rho-1) = \mathcal{E}(W,\beta,2\rho^*-1)$, for some ρ^*.
Now from Proposition 4 we have that,

$$\begin{aligned}
&\left|\frac{1}{n}\log(Z'_{\frac{1}{2}}) - \sup_{\int_{[0,1]}\rho(x)=\frac{1}{2}}\left(-\mathcal{F}_{\frac{1}{2}}(W,\beta)\right)\right| < \varepsilon \\
\Rightarrow &\left|\frac{1}{n}\log(Z'_{\frac{1}{2}}) + \mathcal{E}(W,\beta,2\rho^*-1)\right| < \varepsilon \\
\Rightarrow &\frac{1}{n}\log(Z'_{\frac{1}{2}}) < -\mathcal{E}(W,\beta,2\rho^*-1) + \varepsilon \\
\Rightarrow &\frac{1}{n}\log(Z_{\frac{1}{2}}) < \frac{1}{n}\log(Z'_{\frac{1}{2}}) < -\mathcal{E}(W,\beta,2\rho^*-1) + \varepsilon \\
\Rightarrow &Z_{\frac{1}{2}} < \exp(n(-\mathcal{E}(W,\beta,2\rho^*-1)+\varepsilon)).
\end{aligned} \tag{26}$$

Proof of Theorem 5.2. From (25) and (26) we have that

$$\begin{aligned}
\frac{\mu(A)}{8\mu(B)} &\geq \frac{\exp(n(-\mathcal{E}(W,\beta,2\rho^{opt}-1)-\varepsilon))}{\exp(n(-\mathcal{E}(W,\beta,2\rho^*-1)+\varepsilon))} \\
&= \exp(n(-\mathcal{E}(W,\beta,2\rho^{opt}-1)+\mathcal{E}(W,\beta,2\rho^*-1)-2\varepsilon)) \\
&= \exp(n(c-2\varepsilon)).
\end{aligned} \tag{27}$$

where $c = -\mathcal{E}(W,\beta,2\rho^{opt}-1) + \mathcal{E}(W,\beta,2\rho^*-1)$. As in this case we have $\beta\lambda_1(W) > 1$ and so from Theorem 6.2 we have $c > 0$. Hence choosing ε sufficiently small we obtain the theorem. ◀

9 Counterexample at Critical Temperature

9.1 Example of Fast Mixing at Critical Temperature

In this section we show a sequence of graphs $\{G_n\}$ such that $\{G_n\} \to W$ and we assume $\lambda_1(W)\beta = 1$. But the mixing time of Glauber dynamics on G_n is $O(n\log n)$. To show this we consider the graphs sampled from the model $\mathcal{G}(n, \frac{1}{2} - \frac{1}{\log n})$. Note that, if G_n is sampled from the model $\mathcal{G}(n, \frac{1}{2} - \frac{1}{\log n})$ then $\{G_n\} \to W$, where W is constant function such that $W(x,y) = \frac{1}{2}$ for all x, y. So we assume $\beta = 2$. By Chernoff bound it can be shown that w.h.p. for each vertex we have the number of neighbors of v is $\leq \frac{n}{2}$. Hence following the same path coupling defined in Section 7, we get fast mixing.

9.2 Example of Slow Mixing at Critical Temperature

In this section we show a sequence of graphs $\{G_n\}$ such that $\{G_n\} \to W$ and we assume $\lambda_1(W)\beta = 1$. But the mixing time of Glauber dynamics on G_n is super-polynomial (more precisely, $\exp(\Omega(\sqrt{n}))$). To show this we consider the graphs sampled from the model $\mathcal{G}(n, \frac{1}{2} + \frac{1}{\log n})$. Note that, if G_n is sampled from the model $\mathcal{G}(n, \frac{1}{2} + \frac{1}{\log n})$ then $\{G_n\} \to W$, where W is constant function such that $W(x,y) = \frac{1}{2}$ for all x, y. So we assume $\beta = 2$.

9.2.1 Properties of Random Graph

▶ **Lemma 19.** *Given a graph $G(= (V,E)) \sim \mathcal{G}(n, \frac{1}{2} + \frac{1}{\log n})$. Assume $S \subset V$ such that $|S| = \frac{n}{2} + k$, for some $k \geq 0$. Then we have w.h.p.:*

1. $E(S,S^c) = (\frac{1}{2} + \frac{1}{\log n})(\frac{n^2}{4} - k^2)[1 \pm \frac{c}{\sqrt{n}}]$.
2. $E(S,S) = (\frac{1}{2} + \frac{1}{\log n})(\binom{n/2+k}{2})[1 \pm \frac{c}{\sqrt{n}}]$.
3. $E(S^c,S^c) = (\frac{1}{2} + \frac{1}{\log n})(\binom{n/2-k}{2})[1 \pm \frac{c}{\sqrt{n}}]$.

Proof. The lemma follows from Chernoff bound. ◀

Upper Bound on Balanced Configuration: From Lemma 19 we have w.h.p. for balanced configurations we have

$$\begin{aligned}
\mu(B) &\le \binom{n}{n/2} \exp(2\frac{2}{n}(\frac{1}{2} + \frac{1}{\log n})\binom{n/2}{2}[1 + \frac{c}{\sqrt{n}}]) \exp(-\frac{2}{n}(\frac{1}{2} + \frac{1}{\log n})\frac{n^2}{4}[1 - \frac{c}{\sqrt{n}}]) \\
&= \binom{n}{n/2} \exp(\frac{2}{n}(\frac{1}{2} + \frac{1}{\log n})[2\binom{n/2}{2}(1 + \frac{c}{\sqrt{n}}) - \frac{n^2}{4}(1 - \frac{c}{\sqrt{n}})]) \\
&\le 2^n \exp(\frac{2}{n}(\frac{1}{2} + \frac{1}{\log n})[\frac{n^2}{4}(1 + \frac{c}{\sqrt{n}} - 1 + \frac{c}{\sqrt{n}})]) \\
&= 2^n \exp(c\sqrt{n}[\frac{1}{2} + \frac{1}{\log n}]). \qquad (28)
\end{aligned}$$

Lower Bound on Unbalanced Configuration: Similarly from Lemma 19 we have w.h.p. for configurations with $(\frac{n}{2} + k)$ pluses and $(\frac{n}{2} - k)$ minuses we have

$$\begin{aligned}
\mu(A) &\ge \binom{n}{n/2+k} \exp(\frac{2}{n}(\frac{1}{2} + \frac{1}{\log n})[\binom{n/2+k}{2} + \binom{n/2-k}{2}] \cdot [1 - \frac{c}{\sqrt{n}}]) \\
&\cdot \exp(-\frac{2}{n}(\frac{1}{2} + \frac{1}{\log n})[\frac{n^2}{4} - k^2][1 + \frac{c}{\sqrt{n}}]) \\
&= \binom{n}{n/2+k} \exp(\frac{2}{n}(\frac{1}{2} + \frac{1}{\log n})[\frac{n^2}{4} + k^2] \cdot [1 - \frac{c}{\sqrt{n}}]) \\
&\cdot \exp(-\frac{2}{n}(\frac{1}{2} + \frac{1}{\log n})[\frac{n^2}{4} - k^2][1 + \frac{c}{\sqrt{n}}]) \\
&= \binom{n}{n/2+k} \exp(\frac{2}{n}(\frac{1}{2} + \frac{1}{\log n})[(\frac{n^2}{4} + k^2) \cdot (1 - \frac{c}{\sqrt{n}}) - (\frac{n^2}{4} - k^2)(1 + \frac{c}{\sqrt{n}})]) \\
&= \binom{n}{n/2+k} \exp(\frac{2}{n}(\frac{1}{2} + \frac{1}{\log n})[2(-\frac{n^{3/2}c}{4} + k^2)]) \\
&\ge \frac{1}{\sqrt{\pi n/2}} 2^n \exp(-\frac{2k^2}{n} - \frac{4k^3}{n^2}) \exp(\frac{2}{n}(\frac{1}{2} + \frac{1}{\log n})[2(-\frac{n^{3/2}c}{4} + k^2)]). \qquad (29)
\end{aligned}$$

Here we use the fact that,

$$\binom{n}{n/2+k} \ge \frac{1}{\sqrt{\pi n/2}} 2^n \exp(-\frac{2k^2}{n} - \frac{4k^3}{n^2}).$$

Now taking $k = c_1 n^{7/8}$ we have from (28) and (29) we have:

$$\begin{aligned}
\frac{\mu(A)}{\mu(B)} &\ge \frac{1}{\sqrt{\pi n/2}} \exp(-\frac{2k^2}{n} - \frac{4k^3}{n^2}) \cdot \exp([\frac{1}{2} + \frac{1}{\log n}](-2c\sqrt{n} + \frac{4k^2}{n})) \\
&= \frac{1}{\sqrt{\pi n/2}} \exp(-\frac{2k^2}{n} - \frac{4k^3}{n^2} - c\sqrt{n} - \frac{2c\sqrt{n}}{\log n} + \frac{2k^2}{n} + \frac{4k^2}{n \log n})
\end{aligned}$$

$$= \frac{1}{\sqrt{\pi n/2}} \exp(\frac{4k^2}{n}[\frac{1}{\log n} - \frac{c_1}{n^{1/8}}] - c\sqrt{n} - \frac{2c\sqrt{n}}{\log n})$$
$$= \frac{1}{\sqrt{\pi n/2}} \exp(\Omega(\sqrt{n})).$$

Hence we have the lower bound.

Acknowledgements. The authors thank the anonymous referees for the helpful suggestions.

References

1 Noam Berger, Claire Kenyon, Elchanan Mossel, and Yuval Peres. Glauber dynamics on trees and hyperbolic graphs. *Probability Theory and Related Fields*, 131(3):311–340, 2005.

2 Antonio Blanca and Alistair Sinclair. Dynamics for the Mean-field Random-cluster Model. In Naveen Garg, Klaus Jansen, Anup Rao, and José D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2015)*, volume 40 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 528–543, Dagstuhl, Germany, 2015. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/LIPIcs.APPROX-RANDOM.2015.528`.

3 Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi. Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219(6):1801–1851, 2008.

4 Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós, and Katalin Vesztergombi. Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Annals of Mathematics*, 176(1):151–219, 2012.

5 Russ Bubley and Martin E. Dyer. Path coupling: A technique for proving rapid mixing in Markov chains. In *38th Annual Symposium on Foundations of Computer Science, FOCS'97, Miami Beach, Florida, USA, October 19-22, 1997*, pages 223–231, 1997. `doi:10.1109/SFCS.1997.646111`.

6 Paul Cuff, Jian Ding, Oren Louidor, Eyal Lubetzky, Yuval Peres, and Allan Sly. Glauber dynamics for the mean-field Potts model. *Journal of Statistical Physics*, 149(3):432–477, 2012.

7 Martin Dyer, Alan Frieze, and Mark Jerrum. On counting independent sets in sparse graphs. *SIAM Journal on Computing*, 31(5):1527–1541, 2002.

8 Andreas Galanis, Daniel Štefankovic, and Eric Vigoda. Swendsen-Wang Algorithm on the Mean-Field Potts Model. In Naveen Garg, Klaus Jansen, Anup Rao, and José D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2015)*, volume 40 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 815–828, Dagstuhl, Germany, 2015. Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik. `doi:10.4230/LIPIcs.APPROX-RANDOM.2015.815`.

9 Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. *Combinatorics, Probability and Computing*, 25(04):500–559, 2016.

10 Andreas Galanis, Daniel Štefankovič, Eric Vigoda, and Linji Yang. Ferromagnetic Potts model: Refined #BIS-hardness and related results. *SIAM Journal on Computing*, 45(6):2004–2065, 2016.

11 Roy J. Glauber. Time-dependent statistics of the Ising model. *Journal of mathematical physics*, 4(2):294–307, 1963.

12 Ernst Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, 1925.

13 Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM Journal on computing*, 22(5):1087–1116, 1993.
14 Wilhelm Lenz. Beitrag zum Verständnis der magnetischen Erscheinungen in festen Körpern. *Z. Phys.*, 21:613–615, 1920.
15 David A. Levin, Malwina J. Luczak, and Yuval Peres. Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability. *Probability Theory and Related Fields*, 146(1):223–265, 2010.
16 David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times*. American Mathematical Soc., 2009.
17 László Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.
18 László Lovász and Vera T. Sós. Generalized quasirandom graphs. *Journal of Combinatorial Theory, Series B*, 98(1):146–163, 2008.
19 László Lovász and Balázs Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933–957, 2006.
20 Fabio Martinelli. Lectures on Glauber dynamics for discrete spin models. *Lectures on probability theory and statistics*, pages 93–191, 2004.
21 Fabio Martinelli, Alistair Sinclair, and Dror Weitz. Glauber dynamics on trees: boundary conditions and mixing time. *Communications in Mathematical Physics*, 250(2):301–334, 2004.
22 Elchanan Mossel and Allan Sly. Exact thresholds for Ising–Gibbs samplers on general graphs. *The Annals of Probability*, 41(1):294–328, 2013.
23 Elchanan Mossel, Dror Weitz, and Nicholas Wormald. On the hardness of sampling independent sets beyond the tree threshold. *Probability Theory and Related Fields*, 143(3-4):401–439, 2009.
24 Kevin P. Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.
25 Wolfgang Paul and Jörg Baschnagel. *Stochastic processes*. Springer, 1999.
26 Alistair Sinclair, Piyush Srivastava, Daniel Štefankovič, and Yitong Yin. Spatial mixing and the connective constant: Optimal bounds. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1549–1563. Society for Industrial and Applied Mathematics, 2015.
27 Alistair Sinclair, Piyush Srivastava, and Yitong Yin. Spatial mixing and approximation algorithms for graphs with bounded connective constant. In *54th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 300–309, 2013.
28 Allan Sly and Nike Sun. The computational hardness of counting in two-spin models on d-regular graphs. In *53rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 361–369. IEEE, 2012.

10 Appendix

10.1 Proof of Theorem 6

Here we will prove the remaining proofs of Theorem 6. Before moving on to the proofs we need the following standard definitions.

10.1.1 Preliminaries

We will use the following with X being the space of measurable functions $[0, 1] \to [0, 1]$ and X^* will be the dual space.

▶ **Definition 20** (Weak Convergence). Let X be a normed space. Then a sequence $\{f_n\}$ in X is said to be weak-convergent to $f \in X$ if

$$\forall L \in X^* \text{ we have } L(f_n) \to L(f) \text{ as } n \to \infty,$$

and we denote this by $f_n \underset{w}{\to} f$.

▶ **Definition 21** (Weak Compactness). Let $(X, \|\cdot\|)$ be a normed space with dual space X^*. Then a set $M \subset X$ is called weak compact, if every sequence in M has a weak convergent subsequence with limit in M.

Next we will state two facts and a lemma which we will use in the proof of main theorem.

▶ **Fact 22.** *Let $\{f_n\}$ be a sequence in X such that $f_n \underset{w}{\to} f$ then $\langle f_n, h\rangle \to \langle f, h\rangle$ for all $h \in X$.*

▶ **Fact 23.** *The set of measurable function from $[0,1]$ to $[0,1]$ are weak-compact.*

▶ **Lemma 24.** *Let $f_n \underset{w}{\to} f$ then we have*

$$\limsup_{n\to\infty} H(f_n) \le H(f).$$

▶ **Definition 25** (Smoothed Function). Let U be a step graphon with steps $S_1, \dots, S_k$ and f be a measurable function such that $f : [0,1] \to [0,1]$. Then smoothed version of f w.r.t. U is defined by the step function g with the step $S_1, \dots, S_k$ as follows:

$$g(x) = c_i \text{ if } x \in S_i,$$

where $c_i = \sqrt{\frac{\langle f_n, T_{W_n} f_n\rangle}{\int_{S_i \times S_i} W(x,y)\,dx\,dy}}$.

▶ **Fact 26.** *Let f be a measurable function from $[0,1]$ to $[0,1]$ and g_n be the smoothed version of f w.r.t. the step graphon W_n. Then*

1. $\langle f, T_{W_n} f\rangle = \langle g_n, T_{W_n} g_n\rangle$,
2. $\mathrm{Ent}(f_n) \le \mathrm{Ent}(g_n)$.

10.1.2 Proof of Lemma 11

Proof of Lemma 11. Let's denote the value of the objective function at infimum by D, i.e.,

$$D = \inf_{m:[0,1]\to[-1,1]} \mathcal{E}(W, \beta, m) = \inf_{m:[0,1]\to[-1,1]} \Big(-\frac{\beta}{2}\langle m, T_W m\rangle - \mathrm{Ent}(m)\Big),$$

where

$$\begin{aligned}\mathrm{Ent}(m) = &-\int_0^1 \frac{1}{2}(1 - m(x)) \log(\frac{1}{2}(1 - m(x)))\,dx \\ &-\int_0^1 \frac{1}{2}(1 + m(x)) \log(\frac{1}{2}(1 + m(x)))\,dx.\end{aligned}$$

Let $\{m'_n\}$ be the sequence of functions such that $-\frac{\beta}{2}\langle m'_n, T_W m'_n\rangle - \mathrm{Ent}(m'_n) \to D$. By weak compactness of the set of measurable functions we know that there exists a subsequence $\{m_n\}$ of $\{m'_n\}$ and a measurable function m such that $-\frac{\beta}{2}\langle m_n, T_W m_n\rangle \to -\frac{\beta}{2}\langle m, T_W m\rangle$. From [18] we also have for any graphon W there exist a sequence of step graphons $\{W_t\}$'s such that $W_t \to W$ in L_1 distance. Hence we can also write $-\frac{\beta}{2}\langle m, T_{W_t} m\rangle \to -\frac{\beta}{2}\langle m, T_W m\rangle$. Let g_t be the smoothed version of m w.r.t. the step graphon W_t as defined in Definition 25. Using Fact 26 and the weak compactness of the set of measurable functions there exists a function g such that

$$\langle m, T_{W_t} m\rangle = \langle g_t, T_{W_t} g_t\rangle \to \langle g, T_W g\rangle. \tag{30}$$

Now let's assume another function $\rho : [0,1] \to [0,1]$ such that $\rho(x) = \frac{1}{2}(1 - m(x))\ \forall x \in [0,1]$. Also we define the functional for any measurable $\rho : [0,1] \to [0,1]$ as $H(\rho) = -\int_{[0,1]} \rho(x)\log(\rho(x))\,dx$. Hence $\mathrm{Ent}(m) = H(\rho) + H(1-\rho)$. Now Now from Lemma 24 and Fact 26 we have

$$\mathrm{Ent}(g) \ge \limsup_{n\to\infty} \mathrm{Ent}(g_t) \ge \limsup_{t\to\infty} \mathrm{Ent}(m). \tag{31}$$

Hence from (30) and (31) we have

$$\begin{aligned}\liminf_{t\to\infty}\Big(-\frac{\beta}{2}\langle m, T_{W_t} m\rangle - \mathrm{Ent}(m)\Big) &\ge -\frac{\beta}{2}\langle g, T_W g\rangle - \mathrm{Ent}(g)\\ D &\ge -\frac{c}{2}\langle g, T_W g\rangle - \mathrm{Ent}(g). \end{aligned}\tag{32}$$

Hence the optimum is achieved for g which is a measurable function from $[0,1]$ to $[0,1]$ by weak*-compactness of the set. Hence the infimum is achieved. ◀

10.1.3 Remaining Proofs for Theorem 6.1

We need to prove the strict convexity of the functional defined in Lemma 12. Taking $\rho(x) := \frac{1}{2}(m(x)+1)$ in (3) we have

$$\begin{aligned}\mathcal{E}(W,\beta,m) &= -\frac{\beta}{2}\langle m, T_W m\rangle - \mathrm{Ent}(m)\\ &= -\frac{\beta}{2}\langle (2\rho-1), T_W(2\rho-1)\rangle - \mathrm{Ent}(2\rho-1)\\ &= \underbrace{-\frac{\beta}{2}\int_{[0,1]^2}(2\rho(x)-1)(2\rho(y)-1)W(x,y)\,dx\,dy}_{I(\rho)} - \mathrm{Ent}(2\rho-1).\end{aligned}\tag{33}$$

We use this re-parameterization of the function as the function ρ is an eigenvector for the operator T_W and we will use the property of the eigenvector in the proof.

Proof of Lemma 14. We assume $\rho(x) := \frac{1}{2}(m(x)+1)$ and $s(x) := \frac{1}{2}(p(x)+1)$. Now from (33) for any $\alpha \in [0,1]$ we have the following lemma about the functional I defined in (33).

▶ **Lemma 27.** *For any $\rho, s : [0,1] \to [0,1]$ ($\rho \ne s$ up to measurability) and any $\alpha \in [0,1]$ we have*

$$I((1-\alpha)\rho + \alpha s) - (1-\alpha)I(\rho) - \alpha I(s) < 2\alpha(1-\alpha)||\rho - s||_2^2.$$

Also for the Ent functional we have the following lower bound.

▶ **Lemma 28.** *For any $\rho, s : [0,1] \to [0,1]$ and any $\alpha \in [0,1]$ we have*

$$\mathrm{Ent}((1-\alpha)(2\rho-1)+\alpha(2s-1)) - (1-\alpha)\mathrm{Ent}(2\rho-1) - \alpha\mathrm{Ent}(2s-1) \ge 2\alpha(1-\alpha)||\rho-s||_2^2.$$

From the statement of Lemma 27 and 28, Lemma 12 directly follows. ◀

Now we finish the remaining proofs.

Proof of Lemma 27. From (33) we have

$$\begin{aligned}
&I((1-\alpha)\rho+\alpha s)-(1-\alpha)I(\rho)-\alpha I(s)\\
&=-\frac{\beta}{2}\Big[\int_{[0,1]^2}(2((1-\alpha)\rho(x)+\alpha s(x))-1)(2((1-\alpha)\rho(y)+\alpha s(y))-1)W(x,y)\,dx\,dy\\
&-\int_{[0,1]^2}\Big((1-\alpha)(2\rho(x)-1)(2\rho(y)-1)-\alpha(2s(x)-1)(2s(y)-1)\Big)W(x,y)\,dx\,dy\Big]\\
&=\frac{\beta}{2}\int_{[0,1]^2}\Big[4\alpha(1-\alpha)[\rho(x)\rho(y)+s(x)s(y)-2\rho(x)s(y)]\Big]W(x,y)\,dx\,dy\\
&=2\beta\alpha(1-\alpha)\int_{[0,1]^2}\Big[(\rho(x)-s(x))(\rho(y)-s(y))\Big]W(x,y)\,dx\,dy. \qquad (34)
\end{aligned}$$

Now in (34) we use the fact that $\lambda_1(W)$ is the largest eigenvalue of the graphon W as defined in Definition 3 and also $\lambda_1(W)\beta<1$. So we can rewrite (34) as,

$$\begin{aligned}
&I((1-\alpha)\rho+\alpha s)-(1-\alpha)I(\rho)-\alpha I(s)\\
&=2\alpha(1-\alpha)\beta\int_{[0,1]}(\rho(y)-s(y))\Big[\int_{[0,1]}(\rho(x)-s(x))W(x,y)\,dx\Big]\,dy\\
&\le 2\alpha(1-\alpha)(\lambda_1(W)\beta)\int_{[0,1]}(\rho(y)-s(y))^2\,dy<2\alpha(1-\alpha)||\rho-s||_2^2. \qquad (35)
\end{aligned}$$

This completes the proof. ◀

Proof of Lemma 28. To prove the lemma we will use the following lemma as the main tool.

▶ **Lemma 29.** *Let $\alpha\in[0,1]$ and $R,S\in(0,1)$. Then we have*

$$\begin{aligned}
-(1-\alpha)R\ln(1+\frac{\alpha}{R}(S-R))-\alpha S\ln(1+\frac{1-\alpha}{S}(R-S))\\
-(1-\alpha)(1-R)\ln(1+\frac{\alpha}{1-R}(R-S))-\alpha(1-S)\ln(1+\frac{1-\alpha}{1-S}(S-R))\\
\ge 2\alpha(1-\alpha)(S-R)^2.
\end{aligned}$$

Now we apply Lemma 29 for each point of the integral, i.e., we set $R=\rho(x)$ and $S=s(x)$ and taking integral over $[0,1]$ we have

$$\begin{aligned}
&\mathrm{Ent}((1-\alpha)(2\rho-1)+\alpha(2s-1))-(1-\alpha)\mathrm{Ent}(2\rho-1)-\alpha\mathrm{Ent}(2s-1)\\
&=-(1-\alpha)\int_{[0,1]}\Big[\rho(x)\ln(1+\frac{\alpha}{\rho(x)}(s(x)-\rho(x)))-\alpha s(x)\ln(1+\frac{1-\alpha}{s(x)}(\rho(x)-s(x)))\\
&-(1-\alpha)(1-\rho(x))\ln(1+\frac{\alpha}{1-\rho(x)}(\rho(x)-s(x)))\\
&-\alpha(1-s(x))\ln(1+\frac{1-\alpha}{1-s(x)}(s(x)-\rho(x)))\Big]\,dx\\
&\ge 2\alpha(1-\alpha)\int_{[0,1]}(\rho(x)-s(x))^2\,dx=2\alpha(1-\alpha)||\rho-s||_2^2.
\end{aligned}$$

This completes the proof of Lemma 28. ◀

Now we state another lemma which is used to prove Lemma 29.

▶ **Lemma 30.** *Let $R \in (0,1)$ and $x \in (-R, 1-R)$. Then*

$$F(R,x) := -R\ln(1+\frac{x}{R}) - (1-R)\ln(1-\frac{x}{1-R}) - 2x^2 \geq 0. \tag{36}$$

Proof. We have $F(R,x) = F(1-R,-x)$ and hence it is enough to show (36) for $x \geq 0$. Note that

$$F(R,0) = 0 \quad \text{and} \quad \lim_{x\to(1-R)^-} F(R,x) = \infty. \tag{37}$$

We have that $F(R,x)$ is differentiable on $(-R, 1-R)$ with

$$\frac{\partial}{\partial x}F(R,x) = \frac{-x(2x+2R-1)^2}{(x+R)(x-(1-R))}.$$

If $R \geq 1/2$ there are no critical points of $F(R,x)$ on $(0, 1-R)$ and from (37) we get $F(R,x) \geq 0$ for $x \in (0, 1-R)$. Now assume $R < 1/2$. The only critical point of $F(R,x)$ on $(0, 1-R)$ is $x = 1/2 - R$. We only need to prove that for all $R \in (0, 1/2)$

$$F(R, 1/2 - R) \geq 0.$$

It will be convenient to parameterize $R = 1/2 - T$. We have

$$F(1/2-T, T) = (1/2-T)\ln(1-2T) + (1/2+T)\ln(1+2T) - 2T^2 =: G(T).$$

Note that $G(0) = 0$ and

$$G'(T) = -\ln(1-2T) + \ln(1+2T) - 4T.$$

We will show $G'(T) \geq 0$ for $T \in [0, 1/2)$. Note that $G'(0) = 0$ and for $T \in [0, 1/2)$ we have

$$G''(T) = \frac{16T^2}{1-4T^2} \geq 0,$$

and hence $G'(T) \geq 0$ for $T \in [0, 1/2)$. ◀

Proof of Lemma 29. From Lemma 30 we have

$$(1-\alpha)F(R, \alpha(S-R)) + \alpha F(S, (1-\alpha(R-S)),$$

which is equivalent to the inequality we are proving. ◀

10.1.4 Remaining Proofs for Theorem 6.2

Recall that $S = \{\rho : [0,1] \to [0,1] | \int_{[0,1]} \rho(x)\,dx = \frac{1}{2}\}$. Also assume $A^l_\rho = \{x \in [0,1] | \rho(x) < \frac{1}{2}\}$ and $A^g_\rho = \{x \in [0,1] | \rho(x) > \frac{1}{2}\}$. Then note that A^l_ρ has positive measure if and only if A^g_ρ has positive measure. Also denote $A^{geq}_\rho = A^g_\rho \cup A^{eq}_\rho$, where $A^{eq}_\rho = \{x \in [0,1] | \rho(x) = \frac{1}{2}\}$.

▶ **Definition 31.** Given a function $\rho \in S$ we define another measurable function $\hat{\rho} : [0,1] \to [0,1]$ as follows:

$$\hat{\rho}(x) = \begin{cases} \rho(x) & \textit{if } x \in A^g_\rho, \\ 1-\rho(x) & \textit{otherwise.} \end{cases}$$

We have the following important property of $\hat{\rho}$:

▶ **Claim 32.** *If $x \in A^l_\rho$ then $\hat{\rho}(x) > \rho(x)$.*

▶ **Claim 33.** *Assume $\rho : [0,1] \to [0,1]$ is a measurable function and $\hat{\rho}$ as defined in definition 31. Then*
1. *$\hat{\rho}$ is also a measurable function.*
2. $\mathrm{Ent}(2\hat{\rho} - 1) = \mathrm{Ent}(2\rho - 1)$.

Proof of Claim 33.
1. Follows from the properties of measurability.
2. This follows from the symmetry of Ent function. ◀

Proof of Lemma 33. From Claim 33 we know that $\mathrm{Ent}(2\hat{\rho} - 1) = \mathrm{Ent}(2\rho - 1)$. Hence to prove (20) it is enough to prove that if $\rho \in S = \{\rho : [0,1] \to [0,1] | \int_{[0,1]} \rho(x)\, dx = \frac{1}{2}\} \setminus \{\rho : [0,1] \to [0,1] | f \underset{m}{\approx} \rho \text{ and } \rho(x) = \frac{1}{2} \forall x\}$, then

$$I(\hat{\rho}) < I(\rho).$$

where $I(\rho)$ is defined as in (33). This follows because $\rho(x) \leq \hat{\rho(x)}$ for all x and in particular $\rho(x) < \hat{\rho(x)}$, when $x \in A^l_\rho$ and also $W(x,y)$ is positive everywhere. ◀

Proof of Lemma 15. Let's consider the following function $\rho^b : [0,1] \to [0,1]$:

$$\rho^b(x) = \frac{1}{2}(1 + \varepsilon e_1(x)),$$

for all $x \in [0,1]$, where e_1 is the eigenfunction w.r.t. the largest eigenvalue of W, i.e., $\int_{[0,1]} W(x,y) e_1(y) dy = \lambda_1(W) e_1(x)$ and $\varepsilon > 0$ is some parameter. Now we have

$$\begin{aligned} I(\rho^b) = I(\frac{1}{2}(1 + \varepsilon e_(x)) &= -\frac{\beta}{2} \int_{[0,1]^2} (\varepsilon e_1(x))(\varepsilon e_1(y)) W(x,y)\, dx\, dy \\ &= -\varepsilon^2 \frac{\beta}{2} \int_{[0,1]^2} e_1(x) e_1(y) W(x,y)\, dx\, dy \\ &= -\varepsilon^2 \frac{\beta}{2} \lambda_1(W) \|e_1\|_2^2 < -\frac{\varepsilon^2}{2} \|e_1\|_2^2. \end{aligned} \tag{38}$$

Similarly for entropy we have:

$$\begin{aligned} \mathrm{Ent}(2\rho^b - 1) &= \mathrm{Ent}(\varepsilon \cdot e) \\ &= -\int_{[0,1]} \frac{1}{2}(1 + \varepsilon e_1(x)) \log(\frac{1}{2}(1 + \varepsilon e_1(x))) - \int_{[0,1]} \frac{1}{2}(1 - \varepsilon e_1(x)) \log(\frac{1}{2}(1 - \varepsilon e_1(x))) \\ &= -\log\frac{1}{2} - \int_{[0,1]} \frac{1}{2}(1 + \varepsilon e_1(x)) \log(1 + \varepsilon e_1(x)) - \int_{[0,1]} \frac{1}{2}(1 - \varepsilon e_1(x)) \log(1 - \varepsilon e_1(x)) \\ &= -\log\frac{1}{2} - \int_{[0,1]} \frac{1}{2}(1 + \varepsilon e_1(x))[\varepsilon e_1(x) - \varepsilon^2 e_1^2(x) + \varepsilon^3 e_1^3(x) - \cdots] \\ &\qquad - \int_{[0,1]} \frac{1}{2}(1 + \varepsilon e_1(x))[-\varepsilon e_1(x) - \varepsilon^2 e_1^2(x) - \varepsilon^3 e_1^3(x) - \cdots] \\ &\approx -\log\frac{1}{2} - \frac{\varepsilon^2}{2}\|e_1\|_2^2. \end{aligned} \tag{39}$$

Hence from (38) and (39) $\mathcal{E}(W, \beta, 2\rho^b - 1) \approx \mathcal{E}(W, \beta, 2\rho^{\frac{1}{2}} - 1) - c_\varepsilon \cdot \varepsilon^2$ which implies that $\mathcal{E}(W, \beta, 2\rho - 1)$ is decreasing in the given direction and we have the lemma. ◀