# GLAUBER DYNAMICS FOR THE MEAN-FIELD ISING MODEL: CUT-OFF, CRITICAL POWER LAW, AND METASTABILITY

DAVID A. LEVIN, MALWINA J. LUCZAK, AND YUVAL PERES

Abstract. We study the Glauber dynamics for the Ising model on the complete graph, also known as the Curie-Weiss Model. For $\beta < 1$, we prove that the dynamics exhibits a cut-off: the distance to stationarity drops from near 1 to near 0 in a window of order $n$ centered at $[2(1-\beta)]^{-1} n \log n$. For $\beta = 1$, we prove that the mixing time is of order $n^{3/2}$. For $\beta > 1$, we study metastability. In particular, we show that the Glauber dynamics restricted to states of non-negative magnetization has mixing time $O(n \log n)$.

## 1. Introduction

1.1. **Ising model and Glauber dynamics.** Let $G = (V, \mathcal{E})$ be a finite graph. Elements of the state space $\Omega := \{-1, 1\}^V$ will be called *configurations*, and for $\sigma \in \Omega$, the value $\sigma(v)$ will be called the *spin* at $v$. The *nearest-neighbor energy* $H(\sigma)$ of a configuration $\sigma \in \{-1, 1\}^V$ is defined by

$$H(\sigma) := -\sum_{\substack{v,w \in V,\\ v \sim w}} J(v, w)\sigma(v)\sigma(w), \tag{1.1}$$

where $w \sim v$ means that $\{w, v\} \in \mathcal{E}$. The parameters $J(v, w)$ measure the interaction strength between vertices; we will always take $J(v, w) \equiv J$, where $J$ is a positive constant.

For $\beta \geq 0$, the *Ising model* on the graph $G$ with parameter $\beta$ is the probability measure $\mu$ on $\Omega$ given by

$$\mu(\sigma) = \frac{e^{-\beta H(\sigma)}}{Z(\beta)}, \tag{1.2}$$

where $Z(\beta) = \sum_{\sigma \in \Omega} e^{-\beta H(\sigma)}$ is a normalizing constant.

---

*Date*: February 5, 2008.
2000 *Mathematics Subject Classification.* 60J10, 60K35, 82C20.
*Key words and phrases.* Markov chains, Ising model, Curie-Weiss model, mixing time, cut-off, coupling, Glauber dynamics, metastability, heat-bath dynamics, mean-field model.
Research of Y. Peres was supported in part by NSF grant DMS-0605166.

The parameter $\beta$ is interpreted physically as the inverse of temperature, and measures the influence of the energy function $H$ on the probability distribution. At *infinite temperature*, corresponding to $\beta = 0$, the measure $\mu$ is uniform over $\Omega$ and the random variables $\{\sigma(v)\}_{v\in V}$ are independent.

The (single-site) *Glauber dynamics* for $\mu$ is the Markov chain on $\Omega$ with transitions as follows: When at $\sigma$, a vertex $v$ is chosen uniformly at random from $V$, and a new configuration is generated from $\mu$ conditioned on the set

$$\{\eta \in \Omega \,:\, \eta(w) = \sigma(v),\ w \neq v\}.$$

In other words, if vertex $v$ is selected, the new configuration will agree with $\sigma$ everywhere except possibly at $v$, and at $v$ the spin is $+1$ with probability

$$p(\sigma; v) := \frac{e^{\beta S^v(\sigma)}}{e^{\beta S^v(\sigma)} + e^{-\beta S^v(\sigma)}}, \tag{1.3}$$

where $S^v(\sigma) := J \sum_{w\,:\,w\sim v} \sigma(w)$. Evidently, the distribution of the new spin at $v$ depends only on the current spins at the neighbors of $v$. It is easily seen that $(X_t)$ is reversible with respect to the measure $\mu$ in (1.2).

In what follows, the Glauber dynamics will be denoted by $(X_t)_{t=0}^{\infty}$. We use $\mathbf{P}_\sigma$ and $\mathbf{E}_\sigma$ respectively to denote the underlying probability measure and associated expectation operator when $X_0 = \sigma$.

A *coupling* of the Glauber dynamics with starting states $\sigma$ and $\tilde{\sigma}$ is a process $(X_t, \tilde{X}_t)_{t\geq 0}$ such that $(X_t)$ is a version of the Glauber dynamics with starting state $\sigma$ and $(\tilde{X}_t)$ is a version of the Glauber dynamics with starting state $\tilde{\sigma}$. If a coupling $(X_t, \tilde{X}_t)$ is a Markov chain, we call it a *Markovian coupling*. We write $\mathbf{P}_{\sigma,\tilde{\sigma}}$ and $\mathbf{E}_{\sigma,\tilde{\sigma}}$ for the probability measure and associated expectation respectively corresponding to a coupling with initial states $\sigma$ and $\tilde{\sigma}$.

1.2. **Order $n \log n$ mixing and cut-off.** Given a sequence $G_n = (V_n, E_n)$ of graphs, we write $\mu_n$ for the Ising measure and $(X_t^n)$ for the Glauber dynamics on $G_n$. The worst-case distance to stationarity of the Glauber dynamics chain after $t$ steps is

$$d_n(t) := \max_{\sigma\in\Omega_n} \|\mathbf{P}_\sigma(X_t^n \in \cdot) - \mu_n\|_{\mathrm{TV}}, \tag{1.4}$$

where $\|\mu - \nu\|_{\mathrm{TV}}$ denotes the total variation distance between the probability measures $\mu$ and $\nu$. The *mixing time* $t_{\mathrm{mix}}(n)$ is defined as

$$t_{\mathrm{mix}}(n) := \min\{t \,:\, d_n(t) \leq 1/4\}. \tag{1.5}$$

Note that $t_{\mathrm{mix}}(n)$ is finite for each fixed $n$ since, by the convergence theorem for ergodic Markov chains, $d_n(t) \to 0$ as $t \to \infty$. Nevertheless, $t_{\mathrm{mix}}(n)$ will in general tend to infinity with $n$. Our concern here is with the growth rate of the sequence $t_{\mathrm{mix}}(n)$.

The Glauber dynamics is said to exhibit a *cut-off* at $\{t_n\}$ with *window* $\{w_n\}$ if $w_n = o(t_n)$ and

$$\lim_{\gamma\to\infty} \liminf_{n\to\infty} d_n(t_n - \gamma w_n) = 1,$$
$$\lim_{\gamma\to\infty} \limsup_{n\to\infty} d_n(t_n + \gamma w_n) = 0.$$

The first part of this paper is motivated by the following conjecture, due to the third author:

**Conjecture 1.** *Let $(G_n)$ be a sequence of transitive graphs. If the Glauber dynamics on $G_n$ has $t_{\mathrm{mix}}(n) = O(n \log n)$, then there is a cut-off.*

We establish this conjecture in the special case when $G_n$ is the complete graph on $n$ vertices and $\beta < 1$ (the "high temperature" regime), where the Glauber dynamics has $O(n \log n)$ mixing time.

1.3. **Results.** Here we take $G_n$ to be $K_n$, the complete graph on $n$ vertices. That is, the vertex set is $V_n = \{1, 2, \dots, n\}$, and the edge set $\mathcal{E}_n$ contains all $\binom{n}{2}$ pairs $\{i, j\}$ for $1 \le i < j \le n$. We take the interaction parameter $J$ to be $1/n$; in this case, the Ising measure $\mu$ on $\{-1, 1\}^n$ is given by

$$\mu(\sigma) = \mu_n(\sigma) = \frac{1}{Z(\beta)} \exp\left( \frac{\beta}{n} \sum_{1\le i<j\le n} \sigma(i)\sigma(j) \right). \tag{1.6}$$

In the physics literature, this is usually referred to as the *Curie-Weiss* model. For the remainder of this paper, *Ising model* will always refer to the measure $\mu$ in (1.6), and *Glauber dynamics* will always refer to the one corresponding to this measure. We will often omit the explicit dependence on $n$ in our notation.

It is a consequence of the Dobrushin-Shlosman uniqueness criterion that $t_{\mathrm{mix}}(n) = O(n \log n)$ when $\beta < 1$ (Aizenman and Holley, 1987). See also Bubley and Dyer (1997). Our first result is that there is a cut-off phenomenon in this regime:

**Theorem 1.** *Suppose that $\beta < 1$. The Glauber dynamics for the Ising model on $K_n$ has a cut-off at $t_n = [2(1-\beta)]^{-1} n \log n$ with window size $n$.*

*Remark* 1. Most examples of Markov chains for which the cut-off phenomenon has been proved tend to have ample symmetry, for example, random walks on groups. Part of the interest in Theorem 1 is that the chain studied here is not of this type, and our methods are strictly probabilistic – in particular, based on coupling. Recently, Diaconis and Saloff-Coste (2006) gave a sharp criterion for cut-off (for separation distance) for birth-and-death chains.

In the critical case $\beta = 1$, we prove that the mixing time of the Glauber dynamics is order $n^{3/2}$.

**Theorem 2.** *If $\beta = 1$, then there are constants $C_1, C_2 > 0$ such that for the Glauber dynamics for the Ising model on $K_n$,*

$$C_1 n^{3/2} \leq t_{\text{mix}}(n) \leq C_2 n^{3/2}.$$

Finally, we consider the low-temperature case corresponding to $\beta > 1$. To state our result, it is necessary to mention here the *normalized magnetization*, the function $S$ defined on configurations $\sigma$ by $S(\sigma) := n^{-1} \sum_{i=1}^{n} \sigma(i)$. Also, we define the set $\Omega^+$ of states with non-negative magnetization,

$$\Omega^+ := \{\omega \in X \ : \ S(\sigma) \geq 0\}.$$

By using the Cheeger inequality with estimates on the stationary distribution of the magnetization, the mixing time is seen to be at least exponential in $n$ – slow mixing indeed. Arguments for exponentially slow mixing in the high temperature regime go back at least to Griffiths, Weng and Langer (1966).

In contrast, we prove that the mixing time is of the order $n \log n$ if the chain is restricted to the set $\Omega^+$. To be precise, the restricted dynamics evolve as follows on $\Omega^+$: Generate a candidate move $\eta$ according to the usual Glauber dynamics. If $S(\eta) \geq 0$, accept $\eta$ as the new state, while if $S(\eta) < 0$, move instead to $-\eta$.

**Theorem 3.** *If $\beta > 1$ then there exist constants $C_3(\beta), C_4(\beta) > 0$ depending on $\beta$ such that, for the restricted Glauber dynamics for the Ising model on $K_n$,*

$$C_3(\beta) n \log n \leq t_{\text{mix}}(n) \leq C_4(\beta) n \log n.$$

For other work on the metastability of related models, see Bovier, Eckhoff, Gayrard, and Klein (2001, 2002), and Bovier and Manzo (2002).

The rest of the paper is organized as follows: Section 2 contains some preliminary lemmas required in our proofs. Theorems 1, 2 and 3 are proved in Sections 3, 4, and 5, respectively. Section 6 contains some conjectures and open problems.

## 2. Preliminaries

2.1. **Glauber dynamics for Ising on $K_n$.** We introduce here some notation specific to our setting of the Glauber dynamics for the Ising model on $K_n$. For a configuration $\sigma$, recall that the normalized magnetization $S(\sigma)$ is defined as

$$S(\sigma) := \frac{1}{n} \sum_{j=1}^{n} \sigma(j).$$

Given that the current state of the chain is $\sigma$ and a site $i$ has been selected for updating, the probability $p(\sigma, i)$ of updating to a positive spin, displayed in (1.3), is in this case $p_+(S(\sigma) - n^{-1}\sigma(i))$, where $p_+$ is the function given by

$$p_+(s) := \frac{e^{\beta s}}{e^{\beta s} + e^{-\beta s}} = \frac{1 + \tanh(\beta s)}{2}. \tag{2.1a}$$

Similarly, the probability of updating site $i$ to a negative spin is $p_-(S(\sigma) - n^{-1}\sigma(i))$, where

$$p_-(s) := \frac{e^{-\beta s}}{e^{\beta s} + e^{-\beta s}} = \frac{1 - \tanh(\beta s)}{2}. \tag{2.1b}$$

2.2. **Monotone coupling.** We now describe a process called the *grand coupling*, a Markov chain $(\{X_t^\sigma\}_{\sigma\in\Omega})_{t\geq 0}$ such that for each $\sigma \in \Omega$, the coordinate process $(X_t^\sigma)_{t\geq 0}$ is a version of the Glauber dynamics started at $\sigma$. It will suffice to describe one step of the dynamics. Let $I$ be drawn uniformly from the sites $\{1, 2, \ldots, n\}$, and let $U$ be a uniform random variable on $[0, 1]$, independent of $I$. For each $\sigma \in \Omega$, let $U$ determine the spin $\mathcal{S}^\sigma$ according to

$$\mathcal{S}^\sigma = \begin{cases} +1 & 0 < U \leq p_+(S(\sigma) - n^{-1}\sigma(I))), \\ -1 & p_+(S(\sigma) - n^{-1}\sigma(I)) < U \leq 1. \end{cases}$$

For each $\sigma$, generate the next state $X_1^\sigma$ according to

$$X_1^\sigma(i) = \begin{cases} \sigma(i) & i \neq I \\ \mathcal{S}^\sigma & i = I \end{cases}.$$

We write $\mathbf{P}_{\vec{\sigma}}$ and $\mathbf{E}_{\vec{\sigma}}$ for the probability measure and expectation operator on the measure space where the grand coupling is defined.

For a given pair of configurations, $\sigma$ and $\tilde{\sigma}$, the two-dimensional projection of the grand coupling, $(X_t^\sigma, X_t^{\tilde{\sigma}})_{t\geq 0}$, will be called the *monotone coupling* with starting states $\sigma$ and $\tilde{\sigma}$.

For two configurations $\sigma$ and $\sigma'$, the *Hamming distance* between $\sigma$ and $\sigma'$ is the number of sites where the two configurations disagree, that is

$$\operatorname{dist}(\sigma, \sigma') := \frac{1}{2}\sum_{i=1}^{n} |\sigma(i) - \sigma'(i)|. \tag{2.2}$$

**Proposition 2.1.** *The monotone coupling $(X_t, \tilde{X}_t)$ of the Glauber dynamics started from $\sigma$ and $\tilde{\sigma}$ satisfies*

$$\mathbf{E}_{\vec{\sigma}}\left[\operatorname{dist}(X_t, \tilde{X}_t)\right] \leq \rho^t \operatorname{dist}(\sigma, \tilde{\sigma}), \tag{2.3}$$

*where*

$$\rho := 1 - n^{-1}\left(1 - n\tanh(\beta/n)\right). \tag{2.4}$$

*Proof.* We first show that (2.3) holds with $t = 1$ provided $\mathrm{dist}(\sigma, \tilde{\sigma}) = 1$. Indeed, suppose that $\sigma$ and $\tilde{\sigma}$ agree everywhere except at $i$, where $\sigma(i) = -1$ and $\tilde{\sigma}(i) = +1$.

Recall that the vertex which is updated in all configurations in the grand coupling is denoted by $I$. If $I = i$, then the distance decreases by 1; if $I \neq i$ and the event $B(I)$ occurs, where

$$B(j) := \{p_+(S(\sigma) - \sigma(j)/n) \leq U \leq p_+(S(\tilde{\sigma}) - \tilde{\sigma}(j)/n)\},$$

then the distance increases by 1. In all other cases, the distance remains the same. Consequently,

$$\mathrm{dist}(X_1, \tilde{X}_1) = 1 - \mathbf{1}\{I = i\} + \sum_{j \neq i} \mathbf{1}\{I = j\}\mathbf{1}_{B(j)}. \tag{2.5}$$

Note that $S(\tilde{\sigma}) - \tilde{\sigma}(j)/n = S(\sigma) - \sigma(j)/n + 2/n$ for $j \neq i$. Thus, letting $\hat{s}_j/n = S(\sigma) - \sigma(j)/n$, for $j \neq i$,

$$\mathbf{P}_{\vec{\sigma}}(B(j)) = \frac{1}{2}\left[\tanh(\beta(\hat{s}_j + 2)/n)) - \tanh(\beta\hat{s}_j/n)\right] \leq \tanh(\beta/n). \tag{2.6}$$

Taking expectation in (2.5), by the independence of $U$ and $I$ together with (2.6),

$$\mathbf{E}_{\vec{\sigma}}[\mathrm{dist}(X_1, \tilde{X}_1)] \leq 1 - \frac{1}{n} + \tanh(\beta/n) = \rho \tag{2.7}$$

This establishes (2.3) for the case where $\sigma$ and $\sigma'$ are at unit distance.

Now take any two configurations $\sigma, \tilde{\sigma}$ with $\mathrm{dist}(\sigma, \tilde{\sigma}) = k$. There is a sequence of states $\sigma_0, \ldots, \sigma_k$ such that $\sigma_0 = \sigma, \sigma_k = \tilde{\sigma}$, and each neighboring pair $\sigma_i, \sigma_{i-1}$ are at unit distance. Since we proved the contraction holds for configurations at unit distance,

$$\mathbf{E}_{\vec{\sigma}}\left[\mathrm{dist}(X_1^{\sigma}, X_1^{\tilde{\sigma}})\right] \leq \sum_{i=1}^{k} \mathbf{E}_{\vec{\sigma}}\left[\mathrm{dist}(X_1^{\sigma_i}, X_1^{\sigma_{i-1}})\right] \leq \rho k = \rho\,\mathrm{dist}(\sigma, \tilde{\sigma}).$$

This establishes (2.3) for $t = 1$; iterating completes the proof. ∎

We mention another property of the monotone coupling, from which it receives its name. We write $\sigma \leq \sigma'$ to mean that $\sigma(i) \leq \sigma'(i)$ for all $i$. Given the monotone coupling $(X_t, \tilde{X}_t)$, if $X_t \leq \tilde{X}_t$, then $X_s \leq \tilde{X}_s$ for all $s \geq t$. This is obvious from the definition of the grand coupling, since the function $p_+$ is non-decreasing.

2.3. **Magnetization chain.** Let $S_t := S(X_t)$, and note that $(S_t)$ is itself a Markov chain on $\Omega_S := \{-1, -1 + 2/n, \ldots, 1 - 2/n, 1\}$. The increments

$S_{t+1} - S_t$ take values in $\{-2/n, 0, 2/n\}$, and the transition probabilities are

$$P_M(s, s') = \begin{cases} \frac{1+s}{2} p_-(s - n^{-1}) & s' = s - 2/n, \\ \frac{1-s}{2} p_+(s + n^{-1}) & s' = s + 2/n, \\ 1 - \frac{1+s}{2} p_-(s - n^{-1}) - \frac{1-s}{2} p_+(s + n^{-1}) & s' = s, \end{cases} \tag{2.8}$$

for $s \in \Omega_S$, where $p_+(s)$ and $p_-(s)$ are as in (2.1).

*Remark* 2. It is easily verified that $P_M(-s, -s') = P_M(s, s')$, so the distribution of the chain $(S_t)$ started from $s$ is the same as the distribution of $(-S_t)$ started from $-s$.

*Remark* 3. Let $(X_t^+)$ be the Glauber dynamics restricted to $\Omega^+$, and define $S_t^+ := S(X_t^+)$. The chain $(S_t^+)$ has the same transition probabilities as the chain $|S_t|$.

In the remainder of this subsection, we collect some facts about the Markov chain $(S_t)$ which will be needed in our proofs.

If $(X_t, \tilde{X}_t)$ is a coupling of the Glauber dynamics, we will always write $S_t$ and $\tilde{S}_t$ for $S(X_t)$ and $S(\tilde{X}_t)$, respectively.

**Lemma 2.2.** *Let $\rho$ be as defined in (2.4). If $(X_t, \tilde{X}_t)$ is the monotone coupling, started from states $\sigma$ and $\tilde{\sigma}$, then*

$$\mathbf{E}_{\sigma,\tilde{\sigma}}\left[|S_t - \tilde{S}_t|\right] \leq \left(\frac{2}{n}\right)\rho^t \mathrm{dist}(\sigma, \tilde{\sigma}) \leq 2\rho^t. \tag{2.9}$$

*Proof.* Using the triangle inequality, we see that $|S_t - \tilde{S}_t| \leq (2/n)\mathrm{dist}(X_t, \tilde{X}_t)$. An application of Proposition 2.1 completes the proof. ∎

**Lemma 2.3.** *For the magnetization chain $(S_t)$, for any two states $s$ and $\tilde{s}$ in $\Omega_S$ with $s \geq \tilde{s}$,*

$$0 \leq \mathbf{E}_s[S_1] - \mathbf{E}_{\tilde{s}}[S_1] \leq \rho(s - \tilde{s}). \tag{2.10}$$

*Also, for any two states $s$ and $\tilde{s}$,*

$$|\mathbf{E}_s[S_1] - \mathbf{E}_{\tilde{s}}[S_1]| \leq \rho|s - \tilde{s}|. \tag{2.11}$$

*Proof.* Let $(X_t, \tilde{X}_t)$ be the monotone coupling, started from $(\sigma, \tilde{\sigma})$, where $\sigma \geq \tilde{\sigma}$ and $S(\sigma) = s, S(\tilde{\sigma}) = \tilde{s}$. In this case, $s - \tilde{s} = (2/n)\mathrm{dist}(\sigma, \tilde{\sigma})$, and

$$\mathbf{E}_{\sigma,\tilde{\sigma}}[|S_1 - \tilde{S}_1|] = \mathbf{E}_{\sigma,\tilde{\sigma}}[(2/n)\mathrm{dist}(X_1, \tilde{X}_1)] \leq \frac{2}{n}\rho \mathrm{dist}(\sigma, \tilde{\sigma}) = \rho(s - \tilde{s}).$$

By monotonicity, $X_1 \geq \tilde{X}_1$ and so $S_1 \geq \tilde{S}_1$. Thus, $\mathbf{E}_\sigma[S_1] - \mathbf{E}_{\tilde{\sigma}}[\tilde{S}_1] = \mathbf{E}_{\sigma,\tilde{\sigma}}[|S_1 - \tilde{S}_1|]$, which, together with the preceding inequality, proves that

$$\mathbf{E}_\sigma[S_1] - \mathbf{E}_{\tilde{\sigma}}[\tilde{S}_1] \leq \rho(s - \tilde{s}). \tag{2.12}$$

The left-hand side of (2.12) equals $\mathbf{E}_s[S_1]-\mathbf{E}_{\tilde{s}}[\tilde{S}_1]$, because $(S_t)$ is a Markov chain. Moreover, the left-hand side does not depends at all on the coupling. This proves (2.10). An analogous bound in the case $S(\tilde{\sigma}) \geq S(\sigma)$ establishes (2.11). ∎

We now study the drift of $(S_t)$ in some detail. From (2.8),

$$\mathbf{E}[S_{t+1} - S_t \mid S_t = s] = \frac{2}{n}\left(\frac{1-s}{2}\right)p_+(s+n^{-1}) - \frac{2}{n}\left(\frac{1+s}{2}\right)p_-(s-n^{-1}),$$

and hence

$$\mathbf{E}[S_{t+1} - S_t \mid S_t = s] = \frac{1}{n}\left[f_n(s) - s + \theta_n(s)\right], \tag{2.13}$$

where

$$f_n(s) := \frac{1}{2}\left\{\tanh[\beta(s+n^{-1})] + \tanh[\beta(s-n^{-1})]\right\}$$
$$\theta_n(s) := \frac{-s}{2}\left\{\tanh[\beta(s+n^{-1})] - \tanh[\beta(s-n^{-1})]\right\}.$$

The approximation

$$\mathbf{E}[S_{t+1} - S_t \mid S_t = s] \approx \frac{1}{n}\left[\tanh(\beta s) - s\right] \tag{2.14}$$

will play an important role in our proofs, and we will need to control the error fairly precisely. For the moment, let us observe that (2.14) is valid exactly as an inequality for $s \geq 0$:

$$\mathbf{E}[S_{t+1} - S_t \mid S_t = s] \leq \frac{1}{n}\left[\tanh(\beta s) - s\right]. \tag{2.15}$$

This follows from the concavity of the hyperbolic tangent, together with the fact that the term $\theta_n(s)$ in (2.13) is negative. By Remark 2, for $s \leq 0$,

$$\mathbf{E}[S_{t+1} - S_t \mid S_t = s] \geq \frac{1}{n}\left[\tanh(\beta s) - s\right]. \tag{2.16}$$

Since $(S_t)$ does not change sign when $|S_t| > n^{-1}$, and because tanh is an odd function, putting together (2.15) and (2.16) shows that, for $|S_t| > n^{-1}$,

$$\mathbf{E}\left[|S_{t+1}| \,\middle|\, S_t\right] \leq |S_t| + \frac{1}{n}\left[\tanh(\beta|S_t|) - |S_t|\right]. \tag{2.17}$$

Since $\tanh(x) \leq x$ for $x \geq 0$, when $\beta \leq 1$ equation (2.15) implies that, for $s \geq 0$,

$$\mathbf{E}[S_{t+1} - S_t \mid S_t = s] \leq \frac{s(\beta-1)}{n}. \tag{2.18}$$

Define

$$\tau_0 := \inf\{t \geq 0 \,:\, |S_t| \leq 1/n\}. \tag{2.19}$$

Note that, for $n$ even, $|S_{\tau_0}| = 0$, while for $n$ odd, $|S_{\tau_0}| = 1/n$. The notation $\tau_0$ will be used with the same meaning throughout the paper.

On several occasions, we will need an upper bound on the probability that an unbiased random walk remains positive for at least $u$ steps. The following lemma gives a classical estimate.

**Lemma 2.4.** *Let $(W_t)_{t\geq 0}$ be a random walk with $\mathbf{E}[W_{t+1} - W_t \mid W_t] = 0$ and $|W_{t+1} - W_t| < B$ for some constant $B$. Then there is a constant $c > 0$ such that, for all $u$,*

$$\mathbf{P}_k(|W_1| > 0, \dots, |W_u| > 0) \leq \frac{c|k|}{\sqrt{u}}. \tag{2.20}$$

*(Here $\mathbf{P}_k$ indicates probabilities for the random walk started with $W_0 = k$.)*

Lemma 2.4 can be proved using hitting estimates in Feller (1971); alternatively, it can be seen to be a special case of equation (3.9) in Bender, Lawler, Pemantle, and Wilf (2003/04).

The following lemma is proved for $n$ even. The proof can be modified to deal with the case of $n$ odd by replacing 0 with $1/n$; we omit the details.

**Lemma 2.5.** *Let $\beta \leq 1$, and suppose that $n$ is even. There exists a constant $c$ such that, for all $s$ and for all $u, t \geq 0$,*

$$\mathbf{P}(\,|S_u| > 0, \dots, |S_{u+t}| > 0 \mid S_u = s) \leq \frac{cn|s|}{\sqrt{t}}. \tag{2.21}$$

*Proof.* It will suffice to prove (2.21) for $s > 0$, in which case the absolute values may be removed.

By (2.18), $\mathbf{E}[S_{t+1} - S_t \mid S_t] \leq 0$ for $S_t \geq 0$. Also, there exists a constant $b > 0$ such that $\mathbf{P}(S_{t+1} - S_t \neq 0 \mid S_t) \geq b$ for all times $t$, uniformly in $n$. It follows that $(S_t)$ can be coupled with an unbiased nearest-neighbor random walk $(W_t)$ on $\mathbb{Z}$ satisfying

- $\mathbf{P}(W_1 - W_0 \neq 0 \mid W_0 = w) = b$ for all $w$,
- $W_0 = ns/2$,
- $nS_t/2 \leq W_t$ for $t$ less than the first time $u$ when $S_u \leq n^{-1}$.

From Lemma 2.4, there exists a constant $c > 0$ such that

$$\begin{aligned}\mathbf{P}(S_{u+1} > 0, \dots, S_{u+t} > 0 \mid S_u = s) &\leq \mathbf{P}(W_1 > 0, \dots, W_t > 0 \mid W_0 = ns/2)\\ &\leq \frac{cns}{\sqrt{t}}.\end{aligned}$$

■

## 2.4. Variance bound.

**Lemma 2.6.** *Let $(Z_t)$ be a Markov chain taking values in $\mathbb{R}$ and with transition matrix $P$. We will write $\mathbf{P}_z$ and $\mathbf{E}_z$ for its probability measure and expectation, respectively, when $Z_0 = z$. Suppose that there is some $0 < \rho < 1$ such that for all pairs of starting states $(z, \tilde{z})$,*

$$|\,\mathbf{E}_z[Z_t] - \mathbf{E}_{\tilde{z}}[Z_t]\,| \leq \rho^t |z - \tilde{z}|. \tag{2.22}$$

*Then* $v_t := \sup_{z_0} \mathrm{Var}_{z_0}(Z_t)$ *satisfies*

$$v_t \le v_1 \min\{t, (1-\rho^2)^{-1}\}.$$

*Remark* 4. Suppose that, for every pair $(z, \tilde{z})$, there is a coupling $(Z_1, \tilde{Z}_1)$ of $P(z, \cdot)$ and $P(\tilde{z}, \cdot)$ such that

$$\mathbf{E}_{z,\tilde{z}}\left[\, |Z_1 - \tilde{Z}_1| \,\right] \le \rho |z - \tilde{z}|. \tag{2.23}$$

By iterating (2.23),

$$\left|\mathbf{E}_z[Z_t] - \mathbf{E}_{\tilde{z}}[\tilde{Z}_t]\right| \le \mathbf{E}_{z,\tilde{z}}[|Z_t - \tilde{Z}_t|] \le \rho^t |z - \tilde{z}|.$$

The left-hand side does not depend at all on the coupling, and in particular, (2.22) holds. Moreover, if the state-space of $(Z_t)$ is discrete with a path metric and (2.23) holds for all neighboring pairs $z, \tilde{z}$, then it holds for all pairs of states; see Bubley and Dyer (1997).

*Proof.* Let $(Z_t)$ and $(Z_t^\star)$ be *independent* copies of the chain, both started from $z_0$. By the Markov property and (2.22),

$$\begin{aligned}\left|\mathbf{E}_{z_0}[Z_t \mid Z_1 = z_1] - \mathbf{E}_{z_0}[Z_t^\star \mid Z_1^\star = z_1^\star]\right| &= \left|\mathbf{E}_{z_1}[Z_{t-1}] - \mathbf{E}_{z_1^\star}[Z_{t-1}^\star]\right| \\ &\le \rho^{t-1}|z_1 - z_1^\star|.\end{aligned}$$

Hence, letting $\varphi(z) = \mathbf{E}_z[Z_{t-1}]$, we see that

$$\begin{aligned}\mathrm{Var}_{z_0}\left(\mathbf{E}_{z_0}[Z_t \mid Z_1]\right) &= \frac{1}{2}\mathbf{E}_{z_0}\left[[\varphi(Z_1) - \varphi(Z_1^\star)]^2\right] \\ &\le \frac{1}{2}\mathbf{E}_{z_0}\left[\rho^{2(t-1)}|Z_1 - Z_1^\star|^2\right] \\ &\le v_1 \rho^{2(t-1)}. \end{aligned} \tag{2.24}$$

By the "total variance" formula, for every $z_0$,

$$\mathrm{Var}_{z_0}(Z_t) = \mathbf{E}_{z_0}\left[\mathrm{Var}_{z_0}(Z_t \mid Z_1)\right] + \mathrm{Var}_{z_0}\left(\mathbf{E}_{z_0}[Z_t \mid Z_1]\right),$$

so that

$$v_t \le \sup_{z_0}\left\{\mathbf{E}_{z_0}[\mathrm{Var}_{z_0}(Z_t \mid Z_1)] + \mathrm{Var}_{z_0}\left(\mathbf{E}_{z_0}[Z_t \mid Z_1]\right)\right\}. \tag{2.25}$$

Now, $\mathrm{Var}_{z_0}(Z_t \mid Z_1 = z_1) \le v_{t-1}$ for every $z_1$, and so

$$\mathbf{E}_{z_0}\left[\mathrm{Var}_{z_0}(Z_t \mid Z_1)\right] \le v_{t-1}. \tag{2.26}$$

Thus we have shown that $v_t \le v_{t-1} + v_1\rho^{2(t-1)}$, whence

$$v_t \le v_1 \sum_{i=0}^{t-1} \rho^{2(i-1)} \le v_1 \min\left\{\, (1-\rho^2)^{-1},\; t\right\}.$$

∎

**Proposition 2.7.** *If* $\beta < 1$*, then* $\mathrm{Var}(S_t) = O(n^{-1})$ *as* $n \to \infty$*. If* $\beta = 1$*, then* $\mathrm{Var}(S_t) = O(t/n^2)$ *as* $n \to \infty$*.*

*Proof.* The conclusion follows from combining Lemma 2.3 with Lemma 2.6, and observing that $v_1$ is bounded by $(4/n)^2$ since the incrememnts of $(S_t)$ are at most $2/n$ in absolute value. ∎

2.5. **Expected spin value.** In order to establish the cutoff at high temperature, not only do we need to consider the magnetization chain, but also the number of positive and negative spins among subsets of the vertices.

**Lemma 2.8.** *Let $\beta < 1$.*

(i) *For all $\sigma \in \Omega$ and every $i = 1, 2, \dots, n$,*
$$|\mathbf{E}_\sigma[S_t]| \le 2e^{-(1-\beta)t/n}, \quad \textit{and} \quad |\mathbf{E}_\sigma[X_t(i)]| \le 2e^{-(1-\beta)t/n}.$$

(ii) *For any subset $A$ of vertices, if*
$$M_t(A) := \frac{1}{2}\sum_{i\in A} X_t(i), \tag{2.27}$$
*then $|\mathbf{E}_\sigma[M_t(A)]| \le |A|e^{-(1-\beta)t/n}$ and $\mathrm{Var}(M_t(A)) \le cn$ for some constant $c > 0$.*

(iii) *For any subset $A$ of vertices and all $\sigma \in \Omega$,*
$$\mathbf{E}_\sigma\left[|M_t(A)|\right] \le ne^{-(1-\beta)t/n} + O(\sqrt{n}). \tag{2.28}$$

*Proof.* Let $\mathbf{1}$ denote the configuration of all plus spins, and let $(X_t^T, \tilde{X}_t)$ be the monotone coupling with $X_0^T = \mathbf{1}$ and such that $\tilde{X}_0$ has distribution $\mu$. (Note that then $\tilde{X}_t$ has distribution $\mu$ for all $t \ge 0$, by stationarity.) From Lemma 2.2, because $\mathbf{E}_\mu[\tilde{S}_t] = 0$, we have
$$\mathbf{E}_{\mathbf{1}}\left[S_t^T\right] \le \mathbf{E}_{\mathbf{1},\mu}\left[|S_t^T - \tilde{S}_t|\right] + \mathbf{E}_\mu\left[\tilde{S}_t\right] \le 2e^{-t(1-\beta)/n}.$$
By symmetry, $\mathbf{E}_{\mathbf{1}}\left[X_t^T(i)\right] \le 2e^{-(1-\beta)t/n}$ for all $i$. By monotonicity, for any $\sigma$,
$$\mathbf{E}_\sigma[X_t(i)] \le \mathbf{E}_{\mathbf{1}}[X_t^T(i)] \le 2e^{-(1-\beta)t/n}.$$
Because the chain $(-S_t)$ started from $-\sigma$ has the same distribution as the chain $(S_t)$ started from $\sigma$,
$$-2e^{(1-\beta)t/n} \le \mathbf{E}_\sigma[X_t(i)].$$

For part (ii), the bound on the expectation follows from (i). As for the variance, since the spins are positively correlated,
$$\mathrm{Var}\left(\sum_{i\in A} X_t(i)\right) \le \mathrm{Var}\left(\sum_{i=1}^{n} X_t(i)\right) \le n^2\,\mathrm{Var}(S_t) \le cn, \tag{2.29}$$
where the last inequality follows from Proposition 2.7.

For part (iii), let $(X_t, \tilde{X}_t)$ be the monotone coupling with $X_0 = \sigma$ and the distribution of $\tilde{X}_0$ equal to $\mu$. From the triangle inequality,
$$\mathbf{E}_\sigma[|M_t(A)|] \le \mathbf{E}_{\sigma,\mu}\left[|\tilde{M}_t(A) - M_t(A)|\right] + \mathbf{E}_\mu\left[|\tilde{M}_t(A)|\right].$$

By the Cauchy-Schwartz inequality and since $|\tilde{M}_t(A) - M_t(A)| \leq \text{dist}(X_t, \tilde{X}_t)$,

$$\mathbf{E}_\sigma[|M_t(A)|] \leq \mathbf{E}_{\sigma,\mu}\left[\text{dist}(X_t, \tilde{X}_t)\right] + \sqrt{\mathbf{E}_\mu\left[\tilde{M}_t(A)^2\right]}.$$

Applying Proposition 2.1 shows that

$$\mathbf{E}_\sigma[|M_t(A)|] \leq n\rho^t + \sqrt{\mathbf{E}_\mu\left[\tilde{M}_t(A)^2\right]}. \tag{2.30}$$

Since the variables $\{\tilde{X}_t(i)\}_{i=1}^n$ are positively correlated under $\mu$,

$$\mathbf{E}_\mu\left[\tilde{M}_t(A)^2\right] \leq \frac{n^2}{4}\mathbf{E}_\mu\left[\tilde{S}_t^2\right] = \frac{n^2}{4}\text{Var}_\mu(\tilde{S}_t) = O(n), \tag{2.31}$$

where the last equality follows from Proposition 2.7. Using (2.31) in (2.30) shows that

$$\mathbf{E}_\sigma\left[|M_t(A)|\right] \leq ne^{-(1-\beta)t/n} + O(\sqrt{n}). \tag{2.32}$$

■

2.6. **Coupling of chains with the same magnetization.** The following lemma holds at all temperatures, though we will only be using it for $\beta \geq 1$. It shows that once the magnetizations of two copies of the Glauber dynamics agree, the two copies can be coupled in such a way that the entire configurations agree after at most another $O(n \log n)$ steps. Note that this simple coupling is not fast enough to show cutoff (where we need that once the magnetizations agree, only order $n$ steps are required to fully couple). A more sophisticated coupling for this purpose is given in Section 3.

For any coupling $(X_t, \tilde{X}_t)$, we will let $\tau$ denote the coupling time:

$$\tau := \min\{t \geq 0 \,:\, X_t = \tilde{X}_t\}.$$

**Lemma 2.9.** *Let $\sigma, \tilde{\sigma} \in \Omega$ be such that $S(\sigma) = S(\tilde{\sigma})$. There exists a coupling $(X_t, \tilde{X}_t)$ of the Glauber dynamics with initial states $X_0 = \sigma$ and $\tilde{X}_0 = \tilde{\sigma}$ such that*

$$\limsup_{n\to\infty} \mathbf{P}_{\sigma,\tilde{\sigma}}(\tau > c_0(\beta) n \log n) = 0,$$

*for some constant $c_0(\beta)$ large enough.*

*Proof.* To update the configuration $X_t$ at time $t$, proceed as follows: Pick a site $I \in \{1, 2, \ldots, n\}$ uniformly at random, and generate a random spin $\mathcal{S}$ according to

$$\mathcal{S} = \begin{cases} +1 & \text{with probability } p_+(S_t - X_t(I)/n), \\ -1 & \text{with probability } p_-(S_t - X_t(I)/n). \end{cases}$$

Set

$$X_{t+1}(i) = \begin{cases} X_t(i) & i \neq I, \\ \mathcal{S} & i = I. \end{cases}$$

As for updating $\tilde{X}_t$, if $X_t(I) = \tilde{X}_t(I)$, then let

$$\tilde{X}_{t+1}(i) = \begin{cases} \tilde{X}_t(i) & i \neq I, \\ \mathcal{S} & i = I. \end{cases}$$

If $X_t(I) \neq \tilde{X}_t(I)$, then we pick a vertex $\tilde{I}$ uniformly at random from the set

$$\{i \,:\, \tilde{X}_t(i) \neq X_t(i), \text{ and } \tilde{X}_t(i) = X_t(I)\},$$

and set

$$\tilde{X}_{t+1}(i) = \begin{cases} \tilde{X}_t(i) & i \neq \tilde{I}, \\ \mathcal{S} & i = \tilde{I}. \end{cases}$$

Let $D_t = \sum_{i=1}^{n} |X_t(i) - \tilde{X}_t(i)|/2$ be the number of differing coordinates between $\tilde{X}_t$ and $X_t$.

There exists a constant $c_1 = c_1(\beta) > 0$ such that $p_+(s) \wedge p_-(s) \geq c_1$ uniformly over all $s \in \{-1, \ldots, 1\}$ and all $n$. If $X_t(I) = \tilde{X}_t(I)$, then $D_{t+1} - D_t = 0$ while if $X_t(I) \neq \tilde{X}_t(I)$, then $D_{t+1} - D_t = -2$. It follows that

$$\mathbf{E}[D_{t+1} - D_t \mid X_t, \tilde{X}_t] \leq -\frac{2c_1 D_t}{n},$$

so $Y_t = D_t(1 - 2c_1/n)^{-t}$ is a non-negative supermartingale, whence

$$\mathbf{E}[D_t] \leq \mathbf{E}[D_0]\left(1 - \frac{2c_1}{n}\right)^t \leq ne^{-2c_1 t/n}.$$

Taking $t = c_0 n \log n$ for a sufficiently large constant $c_0 = c_0(\beta)$, we can make the right hand side less than $1/n$, say. Markov's inequality yields

$$\mathbf{P}_\sigma\left(\tau > c_0 n \log n\right) \leq \mathbf{P}_\sigma\left(D_{c_0 n \log n} \geq 1\right) \leq \mathbf{E}_\sigma[D_{c_0 n \log n}] \leq \frac{1}{n}.$$

∎

## 3. Cutoff for the Glauber dynamics at high temperature

In this section we prove Theorem 1. As always, $(X_t)$ will denote the Glauber dynamics, and $S_t = S(X_t) = n^{-1} \sum_{i=1}^{n} X_t(i)$ is the normalized magnetization chain. Recall the definitions

$$\begin{aligned} t_n &= [2(1-\beta)]^{-1} n \log n, \\ \rho &= 1 - (1-\beta)/n, \\ \tau_0 &= \min\{t \geq 0 \,:\, |S_t| \leq 1/n\}. \end{aligned}$$

3.1. **Upper bound.** For convenience, we restate the upper bound part of Theorem 1:

**Theorem 3.1.** *If* $\beta < 1$*, then*

$$\lim_{\gamma\to\infty} \limsup_{n\to\infty} d_n\left([2(1-\beta)]^{-1} n \log n + \gamma n\right) = 0. \tag{3.1}$$

Our strategy is to first construct a coupling of the dynamics so that the magnetizations agree with high probability after $t_n + O(n)$ steps.

**Lemma 3.2.** *Let* $\sigma$ *and* $\tilde{\sigma}$ *be any two configurations. There is a coupling* $(X_t, \tilde{X}_t)$ *of the Glauber dynamics with* $X_0 = \sigma$ *and* $\tilde{X}_0 = \tilde{\sigma}$ *such that, if*

$$\tau_{\text{mag}} := \min\{t \geq 0 \,:\, S_t = \tilde{S}_t\}, \tag{3.2}$$

*then for some constant* $c > 0$ *not depending on* $\sigma, \tilde{\sigma}$ *or n,*

$$\mathbf{P}_{\sigma,\tilde{\sigma}}(\tau_{\text{mag}} > t_n + \gamma n) \leq \frac{c}{\sqrt{\gamma}}. \tag{3.3}$$

*Proof.* Assume without loss of generality that $S(\sigma) > S(\tilde{\sigma})$. Let $(X_t, \tilde{X}_t)$ be the monotone coupling of Section 2.2. Define $\Delta_t := (n/2)|S_t - \tilde{S}_t|$. By Lemma 2.2, for some $c_1 > 0$,

$$\mathbf{E}_{\sigma,\tilde{\sigma}}\left[\Delta_{t_n}\right] \leq c_1 \sqrt{n}. \tag{3.4}$$

Define $\tau_1 := \min\{t \geq t_n \,:\, |\Delta_t| \leq 1\}$. For $t_n \leq t < \tau_1$, allow $(X_t)$ and $(\tilde{X}_t)$ to run independently.

Since $S_t \geq \tilde{S}_t$ for $t \leq \tau_1$, from Lemma 2.3, the process $(S_t - \tilde{S}_t)_{t_n \leq t < \tau_1}$ has non-positive drift. Moreover, since $(X_t)_{t_n \leq t < \tau_1}$ and $(\tilde{X}_t)_{t_n \leq t < \tau_1}$ are independent given $X_{t_n}, \tilde{X}_{t_n}$, for $t > t_n$ the conditional probability that $S_t - \tilde{S}_t$ is non-zero is bounded away from zero uniformly. Thus there is a random walk $(W_t)_{t \geq t_n}$ defined on the same probability space as $(X_t, \tilde{X}_t)$ and satisfying: the increments $W_{t+1} - W_t$ are mean-zero and bounded, $n(S_t - \tilde{S}_t) \leq W_t$ on $[t_n, \tau_1)$, and $n(S_{t_n} - \tilde{S}_{t_n}) = W_{t_n}$.

By Lemma 2.4,

$$\begin{aligned}\mathbf{P}_{\sigma,\tilde{\sigma}}(\tau_1 > t_n + \gamma n \mid X_{t_n}, \tilde{X}_{t_n}) &\leq \mathbf{P}_{\sigma,\tilde{\sigma}}(W_{t_n+1} > 0, \ldots, W_{t_n+\gamma n} > 0 \mid X_{t_n}, \tilde{X}_{t_n}) \\ &\leq \frac{n|S_{t_n} - \tilde{S}_{t_n}|}{\sqrt{\gamma n}}.\end{aligned}$$

Taking expectation above, (3.4) shows that

$$\mathbf{P}_{\sigma,\tilde{\sigma}}(\tau_1 > t_n + \gamma n) \leq O\left(\gamma^{-1/2}\right).$$

The number of plus spins in $X_{\tau_1}$ is either one more than, or the same as, the number of plus spins in $\tilde{X}_{\tau_1}$. Match each plus spin in $\tilde{X}_{\tau_1}$ with a plus spin in $X_{\tau_1}$, and match the remaining spins arbitrarily. From time $\tau_1$ onwards, run a modified version of the monotone coupling, where matched vertices are updated together in the two chains. Define dist′ as the number of disagreements between matched vertices. The conclusion of Lemma 2.2 now holds for this modified monotone coupling, with the distance dist′ replacing dist in (2.9). Thus,

$$\begin{aligned}\mathbf{P}_{\sigma,\tilde\sigma}(\tau_{\mathrm{mag}} > \tau_1 + \gamma' n \mid X_{\tau_1}, \tilde X_{\tau_1}) &\le \mathbf{P}_{\sigma,\tilde\sigma}(\Delta_{\tau_1+\gamma' n} > 1 \mid X_{\tau_1}, \tilde X_{\tau_1}) \\ &\le \mathbf{E}_{\sigma,\tilde\sigma}[\Delta_{\tau_1+\gamma' n} \mid X_{\tau_1}, \tilde X_{\tau_1}] \\ &\le \left(1 - \frac{1-\beta}{n}\right)^{\gamma' n} \\ &\le e^{-(1-\beta)\gamma'}.\end{aligned}$$

We conclude that

$$\mathbf{P}_{\sigma,\tilde\sigma}(\tau_{\mathrm{mag}} \le t_n + \gamma n + \gamma' n) \ge 1 - O\left(\gamma^{-1/2}\right).$$

■

3.2. **Good starting states.** To show the cut-off upper bound, we will start by running the Glauber dynamics for an initial burn-in period. This will ensure that the chain is with high probability in a ‘nice’ configuration required for the coupling argument in Section 3.3. The following lemma is required:

**Lemma 3.3.** *For any a subset $\Omega_0 \subset \Omega$,*

$$\begin{aligned} d(t_0 + t) &= \max_{\sigma\in\Omega} \|\mathbf{P}_\sigma(X_{t_0+t} \in \cdot) - \pi\|_{\mathrm{TV}} \\ &\le \max_{\sigma_0\in\Omega_0} \|\mathbf{P}_{\sigma_0}(X_t \in \cdot) - \pi\|_{\mathrm{TV}} + \max_{\sigma\in\Omega} \mathbf{P}_\sigma(X_{t_0} \notin \Omega_0). \end{aligned} \tag{3.5}$$

*Proof.* For $A \subset \Omega$, we can bound $|\mathbf{P}_\sigma(X_{t_0+t} \in A) - \pi(A)|$ above by

$$\Bigg| \sum_{\sigma_0\in\Omega_0} \left[\mathbf{P}_\sigma(X_{t_0+t} \in A \mid X_{t_0} = \sigma_0) - \pi(A)\right] \mathbf{P}_\sigma(X_{t_0} = \sigma_0) \\ + \left[\mathbf{P}_\sigma(X_{t_0+t} \in A \mid X_{t_0} \notin \Omega_0) - \pi(A)\right] \mathbf{P}_\sigma(X_{t_0} \notin \Omega_0) \Bigg|.$$

Using the triangle inequality, the preceding displayed quantity is bounded above by

$$\sum_{\sigma_0\in\Omega_0} |\mathbf{P}_\sigma(X_{t_0+t} \in A \mid X_{t_0} = \sigma_0) - \pi(A)| \mathbf{P}_\sigma(X_{t_0} = \sigma_0) + \mathbf{P}_\sigma(X_{t_0} \notin \Omega_0).$$

Taking a maximum over subsets $A$ shows that

$$\begin{aligned} &\|\mathbf{P}_\sigma(X_{t_0+t} \in \cdot) - \pi\|_{\mathrm{TV}} \\ &\quad\le \sum_{\sigma_0\in\Omega_0} \|\mathbf{P}_\sigma(X_{t_0+t} \in \cdot \mid X_{t_0} = \sigma_0) - \pi\|_{\mathrm{TV}} \mathbf{P}_\sigma(X_{t_0} = \sigma_0) + \mathbf{P}_\sigma(X_{t_0} \notin \Omega_0). \end{aligned}$$

By the Markov property, $\mathbf{P}_\sigma(X_{t_0+t} \in \cdot \mid X_{t_0} = \sigma_0) = \mathbf{P}_{\sigma_0}(X_t \in \cdot)$, and bounding the average above by the maximum term yields

$$\|\mathbf{P}_\sigma(X_{t_0+t} \in \cdot) - \pi\|_{\mathrm{TV}} \leq \max_{\sigma_0 \in \Omega_0} \|\mathbf{P}_{\sigma_0}(X_t \in \cdot) - \pi\|_{\mathrm{TV}} + \mathbf{P}_\sigma(X_{t_0} \notin \Omega_0).$$

Taking a maximum over $\sigma \in \Omega$ establishes (3.5). ∎

In the proof of Theorem 3.1, we apply Lemma 3.3 with

$$\Omega_0 = \{\sigma \in \Omega \,:\, |S(\sigma)| \leq 1/2\}.$$

For a configuration $\sigma_0 \in \Omega$ define

$$\bar{u}_0 := |\{i \,:\, \sigma_0(i) = 1\}|, \quad \bar{v}_0 := |\{i \,:\, \sigma_0(i) = -1\}|,$$

the number of positive and negative spins, respectively, in $\sigma_0$. Also, define $\Lambda_0 := \{(u, v) \,:\, n/4 \leq u, v \leq 3n/4\}$. Note that

$$\sigma_0 \in \Omega_0 \text{ if and only if } (\bar{u}_0, \bar{v}_0) \in \Lambda_0. \tag{3.6}$$

By Lemma 2.8, there is a constant $\theta_0 > 0$ such that $|\mathbf{E}_\sigma[S_{\theta_0 n}]| \leq 1/4$, whence, for $n$ large enough,

$$\begin{aligned}\mathbf{P}_\sigma(X_{\theta_0 n} \notin \Omega_0) &= \mathbf{P}_\sigma(|S_{\theta_0 n}| > 1/2) \\ &\leq \mathbf{P}_\sigma\left(\left|S_{\theta_0 n} - \mathbf{E}_\sigma[S_{\theta_0 n}]\right| > 1/4\right) \\ &\leq 16\,\mathrm{Var}_\sigma(S_{\theta_0 n}) = O(n^{-1}).\end{aligned} \tag{3.7}$$

The last equality follows from Proposition 2.7.

3.3. **Two-coordinate chain.** Fix a configuration $\sigma_0 \in \Omega_0$. For $\sigma \in \Omega$, define

$$\begin{aligned}U_{\sigma_0}(\sigma) &:= |\{i \in \{0, 1, \ldots, n\} \,:\, \sigma(i) = \sigma_0(i) = 1\}| \\ V_{\sigma_0}(\sigma) &:= |\{i \in \{0, 1, \ldots, n\} \,:\, \sigma(i) = \sigma_0(i) = -1\}|.\end{aligned}$$

In what follows, we shall usually omit the subscript, writing simply $U(\sigma)$ for $U_{\sigma_0}(\sigma)$ and $V(\sigma)$ for $V_{\sigma_0}(\sigma)$.

For a copy of the Glauber dynamics $(X_t)$, the process $(U_t, V_t)_{t \geq 0}$ defined by

$$U_t = U(X_t), \quad \text{and} \quad V_t = V(X_t) \tag{3.8}$$

is a Markov chain on $\{0, 1, \ldots, u_0\} \times \{0, 1, \ldots, v_0\}$ (with transition probabilities depending on the designated configuration $\sigma_0$). We will refer to the chain $(U_t, V_t)$ as the *two-coordinate chain*, and its stationary measure will be denoted by $\pi_2$. Note also that $(U_t, V_t)$ determines the magnetization chain, as we can write

$$S_t = \frac{2(U_t - V_t)}{n} - \frac{\bar{u}_0 - \bar{v}_0}{n}. \tag{3.9}$$

It turns out that, by symmetry, the distance of the law of $X_t$ to $\mu$ equals the distance of the law of $(U_t, V_t)$ to $\pi_2$, as established in the following lemma:

**Lemma 3.4.** *If $(X_t)$ is the Glauber dynamics started from $\sigma_0$ and $(U_t, V_t)$ is the chain defined by (3.8) started from $(\bar{u}_0, \bar{v}_0)$, then*

$$\|\mathbf{P}_{\sigma_0}(X_t \in \cdot) - \mu\|_{\mathrm{TV}} = \|\mathbf{P}_{(\bar{u}_0,\bar{v}_0)}((U_t, V_t) \in \cdot) - \pi_2\|_{\mathrm{TV}}. \tag{3.10}$$

*Proof.* Let

$$\Omega(u, v) := \{\sigma \in \Omega \,:\, (U(\sigma), V(\sigma)) = (u, v)\}.$$

Since both $\mu(\cdot \mid \Omega(u, v))$ and

$$\mathbf{P}_{\sigma_0}(X_t \in \cdot \mid (U_t, V_t) = (u, v))$$

are uniform over $\Omega(u, v)$, it follows that

$$\mathbf{P}_{\sigma_0}(X_t = \eta) - \mu(\eta) = \sum_{u,v} \frac{\mathbf{1}\{\eta \in \Omega(u, v)\}}{|\Omega(u, v)|} \left[\mathbf{P}_{\sigma_0}((U_t, V_t) = (u, v)) - \mu(\Omega(u, v))\right].$$

Applying the triangle inequality, summing over $\eta$, and changing the order of summations shows that

$$\|\mathbf{P}_{\sigma_0}(X_t \in \cdot) - \mu\|_{\mathrm{TV}} \leq \|\mathbf{P}_{(\bar{u}_0,\bar{v}_0)}((U_t, V_t) \in \cdot) - \pi_2\|_{\mathrm{TV}}.$$

The reverse inequality holds since $(U_t, V_t)$ is a function of $(X_t)$. ■

Identity (3.10) implies that it suffices to bound from above the distance to stationarity of the two-coordinate chain.

**Lemma 3.5.** *Suppose two configuration $\sigma$ and $\tilde{\sigma}$ satisfy $S(\sigma) = S(\tilde{\sigma})$ and $R_0 = U(\tilde{\sigma}) - U(\sigma) > 0$. Define*

$$\Xi_1 := \{\sigma \,:\, \min\{U(\sigma),\, \bar{u}_0 - U(\sigma),\, V(\sigma),\, \bar{v}_0 - V(\sigma)\} \geq n/16\}. \tag{3.11}$$

*There exists a Markovian coupling $(X_t, \tilde{X}_t)$ of the Glauber dynamics with starting states $X_0 = \sigma$ and $\tilde{X}_0 = \tilde{\sigma}$ such that the following hold:*

(i) *$S(X_t) = S(\tilde{X}_t)$ for all $t \geq 0$.*
(ii) *If $R_t := U(\tilde{X}_t) - U(X_t)$ and $R_0 \geq 0$, then $R_t \geq 0$ and for all $t$ and*

$$\mathbf{E}_{\sigma,\tilde{\sigma}}\left[R_{t+1} - R_t \mid X_t, \tilde{X}_t\right] \leq 0. \tag{3.12}$$

(iii) *There exists a constant $c$ not depending on $n$ so that on the event $\{X_t \in \Xi_1,\ \tilde{X}_t \in \Xi_1\}$,*

$$\mathbf{P}_{\sigma,\tilde{\sigma}}(R_{t+1} - R_t \neq 0 \mid X_t, \tilde{X}_t) \geq c. \tag{3.13}$$

| | 1 | 2 | 3 | | | | | | ········· | | | | $n$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\sigma_0$ | + | + | + | + | + | + | - | - | - | - | - | - | - |
| | | | $u_0$ | | | | | | $v_0$ | | | | |
| $X_t$ | + | + | + | - | - | - | + | + | + | + | - | - | - |
| | | $A(X_t)$ | | | $B(X_t)$ | | | $C(X_t)$ | | | | $D(X_t)$ | |
| $\tilde{X}_t$ | + | + | + | + | - | - | + | + | + | - | - | - | - |
| | | $A(\tilde{X}_t)$ | | | $B(\tilde{X}_t)$ | | | $C(\tilde{X}_t)$ | | | $D(\tilde{X}_t)$ | | |

FIGURE 1. The vertices in $X_t$ and $\tilde{X}_t$ are partitioned into four categories.

*Proof.* Given the coupling $(X_t, \tilde{X}_t)$, we define $\tilde{U}_t := U(\tilde{X}_t)$ and $\tilde{V}_t := V(\tilde{X}_t)$, and note that $\tilde{U}_t = U_t + R_t$ and $\tilde{V}_t = V_t + R_t$.

For any configuration $\sigma$, we divide the vertices into four sets:

$$\begin{aligned} A(\sigma) &= \{i \in \{1, 2, \dots n\} \,:\, \sigma_0(i) = +1,\ \sigma(i) = +1\}, \\ B(\sigma) &= \{i \in \{1, 2, \dots n\} \,:\, \sigma_0(i) = +1,\ \sigma(i) = -1\}, \\ C(\sigma) &= \{i \in \{1, 2, \dots n\} \,:\, \sigma_0(i) = -1,\ \sigma(i) = +1\}, \\ D(\sigma) &= \{i \in \{1, 2, \dots n\} \,:\, \sigma_0(i) = -1,\ \sigma(i) = -1\}, \end{aligned} \tag{3.14}$$

and so

$$|A(\sigma)| = U(\sigma),\ |B(\sigma)| = \bar{u}_0 - U(\sigma),\ |C(\sigma)| = \bar{v}_0 - V(\sigma),\ |D(\sigma)| = V(\sigma).$$

See Figure 1 for a schematic representation of this partition for $X_t$ and $\tilde{X}_t$.

Our coupling is as follows: To update $X_t$, select a uniformly random $I \in \{1, 2 \dots, n\}$, and generate a random spin $\mathcal{S}$ for $I$ according to the distribution

$$\mathcal{S} = \begin{cases} +1 & \text{with probability } p_+(S_t - X_t(I)/n), \\ -1 & \text{with probability } p_-(S_t - X_t(I)/n). \end{cases}$$

Set

$$X_{t+1}(i) = \begin{cases} X_t(i) & i \neq I, \\ \mathcal{S} & i = I. \end{cases}$$

For $\tilde{X}_t$, we select $\tilde{I}$ uniformly at random from $\{i \,:\, \tilde{X}_t(i) = X_t(I)\}$, and let

$$\tilde{X}_{t+1}(i) = \begin{cases} \tilde{X}_t(i) & i \neq \tilde{I}, \\ \mathcal{S} & i = \tilde{I}. \end{cases}$$

The difference $R_{t+1} - R_t$ is determined by the values of $I, \tilde{I}$ and $\mathcal{S}$ according to the following table:

| $I$ | $\tilde{I}$ | $\mathcal{S}$ | $R_{t+1} - R_t$ |
|---|---|---|---|
| $I \in B(X_t)$ | $\tilde{I} \in D(\tilde{X}_t)$ | $+1$ | $-1$ |
| $I \in C(X_t)$ | $\tilde{I} \in A(\tilde{X}_t)$ | $-1$ | $-1$ |
| $I \in A(X_t)$ | $\tilde{I} \in C(\tilde{X}_t)$ | $-1$ | $+1$ |
| $I \in D(X_t)$ | $\tilde{I} \in B(\tilde{X}_t)$ | $+1$ | $+1$ |
| all other combinations | | | 0 |

It follows that

$$\mathbf{P}_{\sigma,\tilde{\sigma}}(R_{t+1} - R_t = -1 \mid X_t, \tilde{X}_t) = a(U_t, V_t, R_t),$$
$$\mathbf{P}_{\sigma,\tilde{\sigma}}(R_{t+1} - R_t = +1 \mid X_t, \tilde{X}_t) = b(U_t, V_t, R_t),$$

where (using the identities $\tilde{U}_t = U_t + R_t$ and $\tilde{V}_t = V_t + R_t$)

$$\begin{aligned} a(U_t, V_t, R_t) &= \left(\frac{\bar{v}_0 - V_t}{n}\right)\left(\frac{U_t + R_t}{\bar{v}_0 + U_t - V_t}\right) p_-(S_t - 1/n) \\ &\quad + \left(\frac{\bar{u}_0 - U_t}{n}\right)\left(\frac{V_t + R_t}{\bar{u}_0 - U_t + V_t}\right) p_+(S_t + 1/n), \\ b(U_t, V_t, R_t) &= \left(\frac{U_t}{n}\right)\left(\frac{\bar{v}_0 - V_t - R_t}{\bar{v}_0 + U_t - V_t}\right) p_-(S_t - 1/n) \\ &\quad + \left(\frac{V_t}{n}\right)\left(\frac{\bar{u}_0 - U_t - R_t}{\bar{u}_0 - U_t + V_t}\right) p_+(S_t + 1/n). \end{aligned}$$

We obtain

$$\begin{aligned} \mathbf{E}_{\sigma,\tilde{\sigma}}\left[R_{t+1} - R_t \mid X_t, \tilde{X}_t\right] &= b(U_t, V_t, R_t) - a(U_t, V_t, R_t) \\ &= \frac{-R_t}{n}\left[p_-(S_t - 1/n) + p_+(S_t + 1/n)\right], \end{aligned}$$

so, in particular,

$$\mathbf{E}_{\sigma,\tilde{\sigma}}[R_{t+1} - R_t \mid X_t, \tilde{X}_t] \le 0. \tag{3.15}$$

Furthermore, on the event $\{X_t \in \Xi_1, \tilde{X}_t \in \Xi_1\}$,

$$\mathbf{P}_{\sigma,\tilde{\sigma}}(R_{t+1} - R_t \neq 0 \mid X_t, \tilde{X}_t) \ge b(U_t, V_t, R_t) \ge c$$

for some constant $c > 0$, uniformly in $n$, since the functions $p_+$ and $p_-$ are uniformly bounded away from 0 and 1. ∎

*Proof of Theorem 3.1.* Applying Lemma 3.3 with $t_0 = \theta_0 n$, together with the bound (3.7), shows that

$$d_n(\theta_0 n + t) \le \max_{\sigma_0 \in \Omega_0} \|\mathbf{P}_{\sigma_0}(X_t \in \cdot) - \mu\|_{\mathrm{TV}} + O(n^{-1}). \tag{3.16}$$

Hence, using Lemma 3.4 and (3.6),

$$d_n(\theta_0 n + t) \leq \max_{(\bar{u}_0,\bar{v}_0)\in\Lambda_0} \|\mathbf{P}_{(\bar{u}_0,\bar{v}_0)}((U_t, V_t) \in \cdot) - \pi_2\|_{\mathrm{TV}} + O(n^{-1}), \tag{3.17}$$

recalling that $\Lambda_0 = \{(u, v) : n/4 \leq u, v \leq 3n/4\}$.

We will call a pair of chains $(U_t, V_t)_{t\geq 0}$ and $(\tilde{U}_t, \tilde{V}_t)_{t\geq 0}$ a *coupling of the two-coordinate chain* with initial values $(\bar{u}_0, \bar{v}_0)$ and $(\tilde{u}, \tilde{v})$ if

- The two chains are defined on a common probability space,
- Each of $(U_t, V_t)$ and $(\tilde{U}_t, \tilde{V}_t)$ has the same transition probabilities as $(U(X_t), V(X_t))$, where $(X_t)$ is the Glauber dynamics,
- $(U_0, V_0) = (\bar{u}_0, \bar{v}_0)$ and $(\tilde{U}, \tilde{V}) = (\tilde{u}, \tilde{v})$.

We will always consider couplings which have $(\bar{u}_0, \bar{v}_0) \in \Lambda_0$, but $(\tilde{u}, \tilde{v})$ will not be so constrained.

For a given coupling of the two-coordinate chain as above, we let

$$\tau_c := \min\{t \geq 0 : (U_t, V_t) = (\tilde{U}_t, \tilde{V}_t)\}.$$

For a coupling with initial states $(\bar{u}_0, \bar{v}_0)$ and $(\tilde{u}, \tilde{v})$,

$$\|\mathbf{P}_{\bar{u}_0,\bar{v}_0}((U_t, V_t) \in \cdot) - \mathbf{P}_{\tilde{u},\tilde{v}}\left((\tilde{U}, \tilde{V}_t) \in \cdot\right)\|_{\mathrm{TV}} \leq \mathbf{P}_{(\bar{u}_0,\bar{v}_0),(\tilde{u},\tilde{v})}(\tau_c > t). \tag{3.18}$$

(See, for example, Lindvall (2002, Equation 2.8).) A simple calculation shows that

$$\max_{(\bar{u}_0,\bar{v}_0)\in\Lambda_0} \|\mathbf{P}_{\bar{u}_0,\bar{v}_0}((U_t, V_t) \in \cdot) - \pi_2\|_{\mathrm{TV}} \leq \max_{\substack{(\bar{u}_0,\bar{v}_0)\in\Lambda_0,\\ (\tilde{u},\tilde{v})}} \|\mathbf{P}_{\bar{u}_0,\bar{v}_0}((U_t, V_t) \in \cdot) - \mathbf{P}_{\tilde{u},\tilde{v}}((\tilde{U}_t, \tilde{V}_t) \in \cdot)\|_{\mathrm{TV}}. \tag{3.19}$$

We say that $f(n, t)$ is a *uniform coupling bound* if for any initial states $(\bar{u}_0, \bar{v}_0) \in \Lambda_0$ and $(\tilde{u}, \tilde{v})$, there is a coupling of the two-coordinate chain with

$$\mathbf{P}_{(\bar{u}_0,\bar{v}_0),(\tilde{u},\tilde{v})}(\tau_c > t) \leq f(n, t).$$

If $f(n, t)$ is a uniform coupling bound, then combining (3.18) with (3.19) shows that

$$\max_{(\bar{u}_0,\bar{v}_0)\in\Lambda_0} \|\mathbf{P}_{\bar{u}_0,\bar{v}_0}((U_t, V_t) \in \cdot) - \pi_2\|_{\mathrm{TV}} \leq f(n, t),$$

and by (3.17),

$$d_n(\theta_0 n + t) \leq f(n, t) + O(n^{-1}).$$

Recall that $t_n = [2(1 - \beta)]^{-1}(n \log n)$. For any $\gamma > 0$, let $t_n(\gamma) := t_n + \gamma n$. The theorem will be proved if we can establish a uniform coupling bound $f(n, t)$ such that

$$\lim_{\gamma\to\infty} \limsup_{n\to\infty} f(n, t_n(\gamma)) = 0.$$

Fix $(\bar{u}_0, \bar{v}_0) \in \Lambda_0$ and arbitrary $(\tilde{u}, \tilde{v})$. Let $\sigma_0$ be any configuration with $(U(\sigma_0), V(\sigma_0)) = (\bar{u}_0, \bar{v}_0)$, and let $\tilde{\sigma}$ be any configuration with $(U(\tilde{\sigma}), V(\tilde{\sigma})) =$

$(\tilde{u}, \tilde{v})$. We will construct, in two phases, a coupling $(X_t, \tilde{X}_t)$ of the full Glauber dynamics with initial states $X_0 = \sigma_0$ and $\tilde{X}_0 = \tilde{\sigma}$. Given such a coupling, the projections

$$(U_t, V_t) := (U(X_t), V(X_t)), \quad \text{and} \quad (\tilde{U}_t, \tilde{V}_t) := (U(\tilde{X}_t), V(\tilde{X}_t))$$

are a coupling of the two-coordinate chains, started from $(\bar{u}_0, \bar{v}_0)$ and $(\tilde{u}, \tilde{v})$.

The magnetization coupling phase, lasting from time 0 to time $t_n(\gamma)$ will ensure that $S_{t_n(\gamma)} = \tilde{S}_{t_n(\gamma)}$ with high probability, and that

$$\mathbf{E}_{\sigma_0, \tilde{\sigma}}\left[|\tilde{U}_{t_n(\gamma)} - U_{t_n(\gamma)}|\right] = O(\sqrt{n}).$$

During the two-coordinate coupling phase, from time $t_n(\gamma)$ to time $t_n(2\gamma)$, with high probability the chains $(U_t)$ and $(\tilde{U}_t)$ coalesce. To facilitate coalescence, we must ensure that throughout the second phase with high probability $X_t \in \Xi_1$ and $\tilde{X}_t \in \Xi_1$, where $\Xi_1$ is as defined in (3.11). Also, the coupling will ensure $S_t = \tilde{S}_t$ for all $t \in [t_n(\gamma), t_n(2\gamma)]$.

*(i) Magnetization coupling.* Recall that $\tau_{\text{mag}}$, defined in (3.2), is the first time the normalized magnetizations agree. Let $H_1 := \{\tau_{\text{mag}} \leq t_n(\gamma)\}$ be the event that the magnetizations couple by time $t_n(\gamma)$. By Lemma 3.2, there exists a constant $c$ not depending on $\sigma_0$ or $\tilde{\sigma}$ such that

$$\mathbf{P}_{\sigma_0, \tilde{\sigma}}(H_1^c) \leq c\gamma^{-1/2}.$$

*(ii) Two-coordinate chain coupling phase.* Assume that $\tilde{U}_{t_n} > U_{t_n}$; if this is not the case, just reverse the roles of $X_t$ and $\tilde{X}_t$ in what follows. On the event $H_1$, for $t \geq t_n(\gamma)$, use the coupling constructed in Lemma 3.5. On the event $H_1^c$, we let the two chains run independently for $t \geq t_n(\gamma)$.

The outline of the remainder of the proof is as follows: By (3.12), the drift of the difference $\tilde{U}_t - U_t$ is non-positive, so it can be dominated by a process with independent and unbiased increments with values in $\{-1, 0, 1\}$, until $\tilde{U}_t - U_t$ hits zero. Provided that the increments of $\tilde{U}_t - U_t$ are non-zero with probability bounded away from 0 uniformly in $n$, the dominated process can be taken to be an unbiased random walk. We will establish that at time $t_n(\gamma)$, the beginning of the second coupling phase, the expected difference $\mathbf{E}_{\sigma_0, \tilde{\sigma}}[\tilde{U}_{t_n(\gamma)} - U_{t_n(\gamma)}]$ is order $\sqrt{n}$. Thus by comparison with random walk, the two-coordinate process will couple in $O(n)$ more steps.

We begin by showing that, if $H_2(t) := \{X_t \in \Xi_1, \tilde{X}_t \in \Xi_1\}$, then

$$\mathbf{P}_{\sigma_0, \tilde{\sigma}}\left(\bigcup_{t_n(\gamma) \leq t \leq t_n(2\gamma)} H_2(t)^c\right) = O(n^{-1}). \tag{3.20}$$

(Note that the bound above depends on $\gamma$. This does not pose a problem, because the limit in $n$ is taken before the limit in $\gamma$ in (3.1).)

Recall the definition of $M_t(A)$ in (2.27). We introduce the following definitions:

$$\begin{aligned} A_0 &:= \{i \,:\, \sigma_0(i) = 1\}, \\ B^\star &:= \bigcup_{t\in[t_n+\gamma n,\, t_n+2\gamma n]} \{|M_t(A_0)| \ge n/32\}\,, \\ Y &:= \sum_{t\in[t_n+\gamma n,\, t_n+2\gamma n]} \mathbf{1}\{|M_t(A_0)| > n/64\}. \end{aligned}$$

(Note that $|A_0| = \bar{u}_0$.) Since $M_t(A_0)$ has increments in $\{-1, 0, 1\}$, if $|M_{t_0}(A_0)| > n/32$, then $|M_t(A_0)| > n/64$ for all $t$ in any interval of length $n/64$ containing $t_0$. Consequently, $B^\star \subset \{Y > n/64\}$ and

$$\mathbf{P}_{\sigma_0,\tilde{\sigma}}(B^\star) \le \mathbf{P}_{\sigma_0,\tilde{\sigma}}(Y > n/64) \;\le \frac{c_0\mathbf{E}_{\sigma_0,\tilde{\sigma}}[Y]}{n}.$$

By Lemma 2.8(ii), $\mathbf{P}_{\sigma_0,\tilde{\sigma}}(|M_t(A_0)| > n/64) = O(n^{-1})$ for $t \ge t_n$, so $\mathbf{E}_{\sigma,\tilde{\sigma}}[Y] = O(1)$ and

$$\mathbf{P}_{\sigma_0,\tilde{\sigma}}(B^\star) = O(n^{-1}).$$

Making analogous definitions and deductions for the chain $(\tilde{X}_t)$ shows that

$$\mathbf{P}_{\sigma_0,\tilde{\sigma}}(\tilde{B}^\star) = O(n^{-1}).$$

If $U_t \le n/16$, then $\bar{u}_0 - U_t \ge 3n/16$, since we are assuming that $\bar{u}_0 \ge n/4$. Consequently, if $U_t \le n/16$, then

$$|M_t(A_0)| = |U_t - (\bar{u}_0 - U_t)| \ge (\bar{u}_0 - U_t) - U_t \ge \frac{n}{8}.$$

Similarly, $\bar{u}_0 - U_t \ge n/16$ implies that $|M_t(A_0)| \ge 1/8$. An analogous argument applied to $V_t$ and $\bar{v}_0 - V_t$ shows that if either $V_t$ or $\bar{v}_0 - V_t$ does not exceed $n/16$, then $|M_t(A_0)| \ge n/8$, since $|V_t - (\bar{v}_0 - V_t)| = |\bar{v}_0| \ge n/4$. Finally, the same implications are obtained for the chains $(\tilde{X}_t), (\tilde{U}_t)$ and $(\tilde{V}_t)$. To summarize,

$$H_2(t)^c \subset \{|M_t(A_0)| \ge n/16\} \cup \{|\tilde{M}_t(A_0)| \ge n/16\}.$$

Thus,

$$\mathbf{P}_{\sigma_0,\tilde{\sigma}}\left( \bigcup_{t_n(\gamma)\le t\le t_n(2\gamma)} H_2(t)^c \right) \le \mathbf{P}_{\sigma_0,\tilde{\sigma}}(B^\star) + \mathbf{P}_{\sigma_0,\tilde{\sigma}}(\tilde{B}^\star) = O(n^{-1}).$$

Recall that $R_t = |\tilde{U}_t - U_t|$, and let $H_2 := \bigcap_{t_n(\gamma)\le t\le t_n(2\gamma)} H_2(t)$. On the event $H_2$, the process $R_t$ can be dominated by a nearest-neighbor random walk, with delay, until the first time when $(R_t)$ visits 0. Then by Lemma 2.4, on $H_1$,

$$\mathbf{P}_{\sigma_0,\tilde{\sigma}}\left(\{\tau_c > t_n(2\gamma)\} \cap H_2 \,\middle|\, X_{t_n(\gamma)}, \tilde{X}_{t_n(\gamma)}\right) \le \frac{c_1|R_{t_n(\gamma)}|}{\sqrt{n\gamma}}.$$

Taking expectation gives

$$\mathbf{P}_{\sigma_0,\tilde{\sigma}}\left(\{\tau_c > t_n(2\gamma)\} \cap H_2 \cap H_1\right) \le \frac{c_1 \mathbf{E}_{\sigma_0,\tilde{\sigma}}[|R_{t_n(\gamma)}|]}{\sqrt{n\gamma}}. \qquad (3.21)$$

Observe that

$$U_t = M_t(A_0) + \bar{u}_0/2, \quad \text{and} \quad \tilde{U}_t = \tilde{M}_t(A_0) + \bar{u}_0/2,$$

whence

$$|U_t - \tilde{U}_t| = |M_t(A_0) - \tilde{M}_t(A_0)| \le |M_t(A_0)| + |\tilde{M}_t(A_0)|.$$

Taking expectation shows that

$$\mathbf{E}_{\sigma_0,\tilde{\sigma}}[|R_t|] \le \mathbf{E}_{\sigma_0}[|M_t(A_0)|] + \mathbf{E}_{\tilde{\sigma}}[|M_t(A_0)|].$$

Applying Lemma 2.8(iii) shows that $\mathbf{E}_{\sigma_0,\tilde{\sigma}}[|R_{t_n(\gamma)}|] = O(\sqrt{n})$ .

Using this estimate in (3.21), we conclude that

$$\begin{aligned} \mathbf{P}_{\sigma_0,\tilde{\sigma}}\left(\tau_c > t_n(2\gamma)\right) &\le \mathbf{P}_{\sigma_0,\tilde{\sigma}}\left(\{\tau_c > t_n(2\gamma)\} \cap H_2 \cap H_1\right) \\ &\quad + \mathbf{P}_{\sigma_0,\tilde{\sigma}}(H_2^c) + \mathbf{P}_{\sigma_0,\tilde{\sigma}}(H_1^c) \\ &\le \frac{c_2}{\sqrt{\gamma}} + O(n^{-1}). \end{aligned}$$

This gives the uniform coupling bound required. ∎

3.4. **Lower bound.** Recall $t_n = [2(1-\beta)^{-1}] n \log n$, and $\rho = 1 - (1-\beta)/n$. Let us first restate the lower bound part of Theorem 1.

**Theorem 3.6.** *If $\beta < 1$, then*

$$\lim_{\gamma\to\infty} \liminf_{n\to\infty} d_n\left(t_n - \gamma n\right) = 1.$$

*Proof.* It is enough to produce a suitable lower bound on the distance of the distribution of $S_t$ from its stationary distribution, since the chain $(S_t)$ is a projection of the chain $(X_t)$.

Since $\theta_n(s) = O(n^{-2})$, expanding $\tanh[\beta(s + n^{-1})]$ around $\beta s$ in $f_n(s)$ and using equation (2.13) shows that, for $s \ge 0$,

$$\mathbf{E}_{s_0}[S_{t+1} \mid S_t = s] \ge \rho s - \frac{s^3}{2n} - O(n^{-2}). \qquad (3.22)$$

By Remark 2, if $|S_t| > n^{-1}$,

$$\mathbf{E}_{s_0}\left[|S_{t+1}| \,\middle|\, S_t\right] \ge \rho|S_t| - \frac{|S_t|^3}{2n} - O(n^{-2}). \qquad (3.23)$$

This also clearly holds for $|S_t| = 0$ or $|S_t| = n^{-1}$. (In the latter case, $|S_{t+1}| \ge 1/n$.)

Take the initial state $S_0$ to be $s_0 = s_0(\beta)$; we will specify the value of $s_0$ later. Define $Z_t := |S_t|\rho^{-t}$, whence $Z_0 = S_0 = s_0$. Since $\rho^{-1} \le 2$ for large $n$, from (3.23) it follows that

$$\mathbf{E}_{s_0}[Z_{t+1} \mid Z_t] \ge Z_t - \frac{\rho^{-t}[|S_t|^3 + O(1/n)]}{n},$$

for $n$ large enough. Since $0 \le |S_t| \le 1$,

$$\mathbf{E}_{s_0}[Z_t - Z_{t+1} \mid Z_t] \le \frac{\rho^{-t}[|S_t|^3 + O(1/n)]}{n} \le \frac{\rho^{-t}[|S_t|^2 + O(1/n)]}{n}. \tag{3.24}$$

Applying Lemma 2.8(iii) with $A = \{1, 2, \dots, n\}$, we find that

$$\mathbf{E}_{s_0}[|S_t|] \le |s_0|\rho^t + c_1 n^{-1/2}. \tag{3.25}$$

Here and below, the constants $c_i$ depend only on $\beta$.

Using the variance bound $\mathrm{Var}(S_t) \le c_2 n^{-1}$ (c.f. Proposition 2.7) together with the inequality (3.25) shows that

$$\mathbf{E}_{s_0}\left[S_t^2\right] = (\mathbf{E}_{s_0}[S_t])^2 + \mathrm{Var}(S_t) \le s_0^2\rho^{2t} + 2c_1 n^{-1/2}|s_0|\rho^t + c_3 n^{-1}. \tag{3.26}$$

Taking expectations in (3.24) and using (3.26) yields

$$\mathbf{E}_{s_0}[Z_t - Z_{t+1}] \le \frac{1}{n}\left[s_0^2\rho^t + 2c_1 n^{-1/2}|s_0| + c_3\rho^{-t}/n\right] + O(n^{-2}).$$

Let $t^\star = t_n - \alpha n/(1-\beta)$. Adding the increments $\mathbf{E}_{s_0}[Z_t] - \mathbf{E}_{s_0}[Z_{t+1}]$ for $t = 0, \dots, t^\star - 1$, the above inequality gives that

$$s_0 - \mathbf{E}_{s_0}[Z_{t^\star}] \le \frac{s_0^2}{n(1-\rho)} + \frac{2c_1|s_0|t^\star}{n^{3/2}} + c_3\frac{\rho^{-t^\star}}{n^2(1-\rho)} + O(t^\star n^{-2}).$$

Since $\rho^{-t^\star} \le n^{1/2}$, we deduce that

$$s_0 - \mathbf{E}_{s_0}[Z_{t^\star}] \le \frac{s_0^2}{1-\beta} + \frac{2c_2\log(n)}{n^{1/2}} + c_4 n^{-1/2}. \tag{3.27}$$

If $s_0 < (1-\beta)/3$ and $n$ is large enough, then the right-hand side of (3.27) is less than $s_0/2$. Thus

$$\mathbf{E}_{s_0}[|S_{t^\star}|] \ge \frac{s_0\rho^{t^\star}}{2} \ge B := \frac{s_0 e^\alpha}{2n^{1/2}}.$$

By Proposition 2.7, $\max\{\mathrm{Var}_{s_0}(S_t), \mathrm{Var}_\mu(S)\} \le c_5/n$. Thus

$$B/2 \le \mathbf{E}_{s_0}[S_{t^\star}] - \frac{s_0 e^\alpha}{4c_5}\sqrt{\mathrm{Var}_{s_0}(S_{t^\star})},$$

$$B/2 \ge \mathbf{E}_\mu[S] + \frac{s_0 e^\alpha}{4c_5}\sqrt{\mathrm{Var}_\mu(S)}.$$

Let $\pi_S$ be the stationary distribution of $(S_t)$, and let $A := [-B/2, B/2]$. Then

$$\|\mathbf{P}_{s_0}(S_{t^\star} \in \cdot) - \pi_S\|_{\mathrm{TV}} \ge \pi_S(A) - \mathbf{P}_{s_0}(|S_{t^\star}| \in A) \ge 1 - 32c_5^2 e^{-2\alpha}/s_0^2,$$

where the last inequality follows from application of Chebyshev's inequality. The right-hand side clearly tends to 1 as $\alpha \to \infty$. ■

## 4. Critical Case

In this section, we analyze the mixing time of the Glauber dynamics in the critical case $\beta = 1$, proving Theorem 2. We consider the upper and lower bounds separately.

### 4.1. Upper bound.

**Theorem 4.1.** *If $\beta = 1$, then $t_{\mathrm{mix}} = O(n^{3/2})$.*

Recall the definition of $\tau_0$ in (2.19): $\tau_0 := \min\{t \geq 0 \,:\, |S_t| \leq 1/n\}$.

*Proof.* We show that we can couple Glauber dynamics so that the magnetizations agree in order $n^{3/2}$ steps, and then appeal to Lemma 2.9 to show the configurations can be made to agree in another order $n \log n$ steps.

*Step 1:* Our first goal is to prove that $\lim_{c\to\infty} \mathbf{P}_\sigma(\tau_0 > cn^{3/2}) = 0$, uniformly in $n$.

Recall the inequality (2.17): For $|S_t| > n^{-1}$,

$$\mathbf{E}_\sigma\left[|S_{t+1}| \,\middle|\, S_t\right] \leq \left(1 - \frac{1}{n}\right)|S_t| + \frac{1}{n}\tanh(|S_t|).$$

Multiply both sides above by $\mathbf{1}\{\tau_0 > t\}$ and use the fact that $\tanh(0) = 0$ to find that

$$\mathbf{E}_\sigma\left[|S_{t+1}|\mathbf{1}\{\tau_0 > t\} \,\middle|\, S_t\right] \leq \left(1 - \frac{1}{n}\right)|S_t|\mathbf{1}\{\tau_0 > t\} + \frac{1}{n}\tanh(|S_t|\mathbf{1}\{\tau_0 > t\}).$$

Since $\mathbf{1}\{\tau_0 > t+1\} \leq \mathbf{1}\{\tau_0 > t\}$,

$$\mathbf{E}_\sigma\left[|S_{t+1}|\mathbf{1}\{\tau_0 > t+1\} \,\middle|\, S_t\right] \leq \left(1 - \frac{1}{n}\right)|S_t|\mathbf{1}\{\tau_0 > t\} + \frac{1}{n}\tanh(|S_t|\mathbf{1}\{\tau_0 > t\}).$$

Define $\xi_t^+ := \mathbf{E}_\sigma[|S_t|\mathbf{1}\{\tau_0 > t\}]$. Take expectation above and apply Jensen's inequality to the concave function tanh restricted to the non-negative axis, to see that

$$\xi_{t+1}^+ \leq \left(1 - \frac{1}{n}\right)\xi_t^+ + \frac{1}{n}\tanh(\xi_t^+). \tag{4.1}$$

Thus, there exists a constant $c_\varepsilon > 0$ such that, if $\xi_t^+ \geq \varepsilon$, then

$$\xi_{t+1}^+ - \xi_t^+ \leq -\frac{c_\varepsilon}{n}.$$

We conclude that there exists a time $t_\star = t_\star(n) = O(n)$ such that $\xi_t^+ \leq 1/4$ for all $t \geq t_\star$.

Expand $\tanh(x)$ in a Taylor series and use (4.1) to obtain

$$\xi^+_{t+1} \leq \xi^+_t - \frac{(\xi^+_t)^3}{4n} + O(n^{-2}),$$

for $t \geq t_\star$.

This shows that, for $n$ sufficiently large, $\xi^+_t$ is decreasing for $t \geq t_\star$. We will assume from now on that $n$ is large enough for this to hold. Given a decreasing sequence of numbers

$$1/4 \geq b_1 > b_2 > \cdots > 0,$$

let $u_i := \min\{t \geq t_\star : \xi^+_t \leq b_i\}$. Since $\xi^+_t$ is decreasing, $b_{i+1} < \xi^+_t \leq b_i$ for all times $u_i \leq t < u_{i+1}$. Let $b_i = (1/4)2^{-i}$. For $t \in (u_i, u_{i+1}]$,

$$\xi^+_{t+1} \leq \xi^+_t - \frac{b_i^3}{32n} + O(n^{-2}).$$

It follows that

$$u_{i+1} - u_i \leq \frac{16n}{b_i^2}\left[1 + O(b_i^{-3}n^{-1})\right]$$

Let $i_0 = \min\{i : b_i \leq n^{\alpha-1}\}$, where $\alpha$ is a parameter to be chosen below. If $\alpha > 2/3$, then $b_i \geq n^{-1/3+\delta}$ for $i < i_0$, for some $\delta > 0$. In particular, $b_i^{-3} \leq n^{1-\delta}$ and $O(b_i^{-3}n^{-1}) = o(n)$ for $i < i_0$. Thus for $n$ large enough, for $0 \leq i < i_0$,

$$u_{i+1} - u_i \leq \frac{32n}{b_i^2}.$$

Summing the above,

$$u_{i_0} - u_0 \leq \sum_{i=0}^{i_0-1} \frac{32n}{b_i^2} \leq \frac{c_0 n}{b_{i_0-1}^2} = O(n^{3-2\alpha}),$$

so

$$u_{i_0} \leq O(n^{3-2\alpha}) + O(n),$$

where the second inequality follows since $u_0 = t^\star = O(n)$. To summarize, provided $1 \geq \alpha > 2/3$, there is a constant $c_1$ such that $\xi^+_t \leq n^{\alpha-1}$ for $t \geq c_1 n^{3-2\alpha}$. In particular, letting $r_n = c_1 n^{3-2\alpha}$, there is a constant $c_2 > 0$ such that

$$\mathbf{E}_\sigma\left[|S^+_{r_n}|\mathbf{1}\{\tau_0 > r_n\}\right] \leq c_2 n^{\alpha-1}. \tag{4.2}$$

By the Markov property and Lemma 2.5, for some constant $c_3$,

$$\mathbf{P}_\sigma(\tau_0 > r_n + \gamma n^{2\alpha} \mid X_{r_n}) \leq \frac{c_3 n |S_{r_n}|}{\sqrt{\gamma} n^\alpha}.$$

Multiplying both sides by $\mathbf{1}\{\tau_0 > r_n\}$, taking expectation, and then using (4.2) shows that

$$\mathbf{P}_\sigma(\tau_0 > r_n + \gamma n^{2\alpha}) = O(\gamma^{-1/2}).$$

Choosing $\alpha = 3/4 > 2/3$, we see that

$$\mathbf{P}_\sigma(\tau_0 > (c_1 + \gamma)n^{3/2}) = O(\gamma^{-1/2}).$$

*Step 2: Construction of coupling*. We now describe how to build a Markovian coupling $(X_t, \tilde{X}_t)$ of the Glauber dynamics such that the following holds: There are constants $c_1 > 0$ and $b < 1$ such that, if $\tau_{\mathrm{mag}}$ is as defined in (3.2), then for *any* two configurations $\sigma$ and $\tilde{\sigma}$,

$$\mathbf{P}_{\sigma,\tilde{\sigma}}(\tau_{\mathrm{mag}} > c_1 n^{3/2}) \leq b. \tag{4.3}$$

This is sufficient, since we only desire to prove $t_{\mathrm{mix}} = O(n^{3/2})$.

Fix two configurations $\sigma$ and $\tilde{\sigma}$, and suppose without loss of generality that $|S(\sigma)| > |S(\tilde{\sigma})|$. Define the stopping time $\tau_{\mathrm{abs}}$ to be the first time the two chains cross over one another, i.e.

$$\tau_{\mathrm{abs}} := \min\{t \geq 0 \,:\, |S_t| \leq |\tilde{S}_t|\},$$

and let $G_1 := \{|S_{\tau_{\mathrm{abs}}+1}| = |\tilde{S}_{\tau_{\mathrm{abs}}+1}|\}$ be the event that the two chains meet one step after $\tau_{\mathrm{abs}}$. There is a constant $c_4 > 0$, not depending on $n$, such that $\mathbf{P}_{\sigma,\tilde{\sigma}}(G_1) \geq c_4$.

On $G_1^c$, couple the two chains independently. On $G_1$, we divide into two cases:

*Case* $S_{\tau_{abs}+1} = \tilde{S}_{\tau_{abs}+1}$. If this situation occurs, then couple such that the magnetizations continue to agree. To do so, if a site $I$ is selected to update $X_t$ with a spin $\mathcal{S}$, then pick a site in $\tilde{X}_t$ at random from those with the same spin as $X_t(I)$, and update this site also with spin $\mathcal{S}$.

*Case* $S_{\tau_{abs}+1} = -\tilde{S}_{\tau_{abs}+1}$. In this case, we use the *reflection coupling*: Suppose state $I$ is selected to update $X_t$, and the spin used to update is $\mathcal{S}$. Then pick a site in $\tilde{X}_t$ at random from those with spin $-X_t(I)$, and update with spin $-\mathcal{S}$. In this case, the process $(S_t)$ and $(\tilde{S}_t)$ will be reflections of one another for $t \geq \tau_{\mathrm{abs}}$.

If $n$ is even, in either situation the magnetizations agree at time $\tau_0$, so $\tau_{\mathrm{mag}} \leq \tau_0$. For even $n$, run the chains together after $\tau_0$. If $n$ is odd, at time $\tau_0$ run the chains independently of one another for a single step.

By Step 1 of the proof, there exists a constants $c_\star$ and $c_6 > 0$ such that, for all $\sigma$,

$$\mathbf{P}_\sigma(\tau_0 + 1 \leq c_\star n^{3/2}) \geq c_6. \tag{4.4}$$

Let $G_2 = \{\tau_0 + 1 \leq c_\star n^{3/2}\}$.

Let $G_3$ be the event that the two chains couple at time $\tau_0 + 1$. There exists some $c_5 > 0$ not depending on $n$ such that $\mathbf{P}_\sigma(G_3 \mid G_1 \cap G_2) \geq c_5$. (If $n$ is even, this probability is one.)

Then

$$\mathbf{P}_\sigma(G_1 \cap G_2 \cap G_3) \leq \mathbf{P}_\sigma(\tau_c \leq c_\star n^{3/2}).$$

The probability on the left is uniformly bounded away from zero, completing the proof. ■

### 4.2. **Lower bound.**

**Theorem 4.2.** *Suppose $\beta = 1$. There is a constant $C_1 > 0$ such that $t_{\text{mix}} \geq C_1 n^{3/2}$.*

*Proof.* It will suffice to prove a lower bound on the mixing time of the magnetization chain $(S_t)$.

As usual, $S$ denotes the normalized magnetization in equilibrium. The sequence $n^{1/4}S$ converges to a non-trivial limit law as $n \to \infty$. (This is proved in Simon and Griffiths (1973); see also Ellis (1985, Theorem V.9.5).) Take $A > 0$ such that

$$\mu\left(|S| \leq An^{-1/4}\right) \geq 3/4. \tag{4.5}$$

Take $s_0 = 2An^{-1/4}$. Let $(\tilde{S}_t)$ be a chain with the same transition probabilities as $(S_t)$, except at $s_0$. At $s_0$, the $\tilde{S}$-chain remains at $s_0$ with probability equal to the probability that the $S$-chain either moves up or remains in place at $s_0$. The two chains can be coupled so that $\tilde{S}_t \leq S_t$ when both are started from $s_0$. In particular, for all $s$, the inequality $\mathbf{P}_{s_0}(S_t \leq s) \leq \mathbf{P}_{s_0}(\tilde{S}_t \leq s)$ holds.

Let $Z_t = \tilde{S}_0 - \tilde{S}_{t\wedge\tau}$, where $\tau := \min\{t \geq 0 : \tilde{S}_t \leq An^{-1/4}\}$. Note that $(Z_t)$ is non-negative.

We will now show that if $\mathcal{F}_t$ is the sigma-algebra generated by $Z_1, \ldots, Z_t$, then there is a constant $c_A$ so that

$$\mathbf{E}_{s_0}[Z_{t+1}^2 - Z_t^2 \mid \mathcal{F}_t] \leq \frac{c_A}{n^2}. \tag{4.6}$$

The equation (4.6) is clearly satisfied when $Z_t = 0$. On the event $\tilde{S}_t = s$, where $An^{-1/4} < s < s_0$, the conditional distribution of $\tilde{S}_{t+1}$ is the same as the conditional distribution of $S_{t+1}$ given $S_t = s$. Thus

$$\mathbf{E}_{s_0}[\tilde{S}_{t+1} \mid \tilde{S}_t = s] = \mathbf{E}_{s_0}[S_{t+1} \mid S_t = s] \geq s - c_0\frac{s^3}{n}, \tag{4.7}$$

for a constant $c_0$. The inequality is obtained by expanding tanh in (2.13). From (4.7), it follows that

$$\mathbf{E}_{s_0}[Z_{t+1} \mid \mathcal{F}_t] \leq Z_t + \frac{c_0}{n}\tilde{S}_t^3. \tag{4.8}$$

We decompose the conditional second moment of $Z_{t+1}$ as

$$\mathbf{E}_{s_0}[Z_{t+1}^2 \mid \mathcal{F}_t] = \mathrm{Var}(Z_{t+1} \mid \mathcal{F}_t) + \left(\mathbf{E}_{s_0}[Z_{t+1} \mid \mathcal{F}_t]\right)^2 . \tag{4.9}$$

Since $|Z_{t+1} - Z_t| \le 2/n$,

$$\mathrm{Var}(Z_{t+1} \mid \mathcal{F}_t) = \mathrm{Var}(Z_{t+1} - Z_t + Z_t \mid \mathcal{F}_t) = \mathrm{Var}(Z_{t+1} - Z_t \mid \mathcal{F}_t) \le \frac{4}{n^2}. \tag{4.10}$$

By (4.8), for $t < \tau$, there is a constant $c_1$ (depending on $A$) so that

$$\mathbf{E}_{s_0}^2[Z_{t+1} \mid \mathcal{F}_t] \le Z_t^2 + 2\frac{c_0}{n} Z_t \tilde{S}_t^3 + \frac{c_0^2 \tilde{S}_t^6}{n^2} \le Z_t^2 + c_1 n^{-2}. \tag{4.11}$$

Using the bounds (4.10) and (4.11) in (4.9) establishes (4.6). We conclude that

$$\mathbf{E}_{s_0}[Z_t^2] \le c_A n^{-2} t. \tag{4.12}$$

Note that

$$\mathbf{E}_{s_0}[Z_t^2] \ge \mathbf{E}_{s_0}[Z_t^2 \mathbf{1}\{\tau \le t\}] \ge \frac{A^2}{n^{1/2}} \mathbf{P}_{s_0}(\tau \le t),$$

which together with (4.12) shows that

$$\mathbf{P}_{s_0}(\tau \le t) \le \frac{c_A t}{A^2 n^{3/2}}.$$

Taking $t = (A^2/4c_A) n^{3/2}$ above shows that

$$\mathbf{P}_{s_0}(S_t \le A n^{-1/4}) \le \frac{1}{4}.$$

This, together with the bound (4.5), proves that $d(c_3 n^{3/2}) \ge 1/2$, where $c_3 = A^2/4c_A$. That is, $t_{\mathrm{mix}} \ge c_3 n^{3/2}$. ∎

## 5. Truncated Dynamics for Low Temperature

We now consider the case $\beta > 1$. As stated in the introduction, the mixing time for the full Glauber dynamics is exponential in $n$. This is proved via an upper bound on the Cheeger constant, defined as

$$\Phi := \min_{A \,:\, \mu(A) \le 1/2} \frac{\sum_{x \in A, y \notin A} \mu(x) P(x,y)}{\mu(A)},$$

where $P$ is the transition matrix for the Glauber dynamics. By taking $A = \{\sigma \;:\; \mu(\sigma) \ge 0\}$ and estimating $\left[\sum_{x \in A, x \notin A} \mu(x) P(x,y)\right] / \mu(A)$, when $\beta > 1$ there are positive constants $c_1$ and $c_2$ such that $\Phi \le c_1 e^{-c_2 n}$. The spectral gap of $P$ is bounded below by $c_3/\Phi$ (see, for example, Sinclair (1993).) The mixing time, in turn, is bounded below by the spectral gap (see, for example, Aldous and Fill (in progress).) The details of this standard argument can be found in the forthcoming book Levin, Peres, and Wilmer (2007). That the Glauber dynamics is slow mixing for $\beta > 1$ was understood as far back as

Griffiths, Weng, and Langer (1966), although they lacked the tool of the Cheeger inequality to make a complete proof.

Here we study the Glauber dynamics confined to the configurations where the magnetization is non-negative, and show that the restricted Glauber dynamics has a mixing time of order $n \log n$.

We remind the reader of the exact mechanism for restricting the dynamics. The usual dynamics are run from a state with non-negative magnetization. If a move to a state $\eta$ is proposed, and $\eta$ has negative magnetization, then the chain moves to $-\eta$ instead.

To establish an $O(n \log n)$ upper bound on the mixing time, we need to estimate the hitting times of the normalized magnetization chain.

**Lemma 5.1.** *Let $\beta > 1$. Let $s^\star$ denote the unique positive solution to* $\tanh(\beta s) = s$*, and for $\alpha > 0$ define*

$$\tau^\star = \tau^\star(\alpha) := \inf\{t \geq 0 : S_t^+ \leq s^\star + \alpha n^{-1/2}\}. \tag{5.1}$$

*There exists a constant $c > 0$, depending on $\alpha$ and $\beta$, such that*

$$\lim_{n\to\infty} \mathbf{P}_\sigma(\tau^\star > cn \log n) = 0.$$

*Proof.* Let $\gamma^\star := \beta \cosh^{-2}(\beta s^\star)$. First, we show that

$$\mathbf{E}_\sigma[S_{t+1}^+ - s^\star \mid S_t^+ = s] \leq \left[1 - \frac{(1-\gamma^\star)}{n}\right](s - s^\star). \tag{5.2}$$

By Remark 3 and (2.17), for $S_t^+ > 1/n$

$$\mathbf{E}_\sigma\left[S_{t+1}^+ - S_t^+ \,\middle|\, S_t^+\right] \leq \frac{1}{n}\left[\tanh(\beta S_t^+) - S_t^+\right].$$

Since $\beta > 1$, it follows that $\gamma^\star = \beta \cosh^{-2}(\beta s^\star) < 1$. By the mean-value theorem, for $y > 0$,

$$\tanh[\beta(s^\star + y)] - \tanh(\beta s^\star) = \frac{\beta}{\cosh^2(\bar{s})} y,$$

for some $\bar{s} \in [s^\star, s^\star + y]$. Since $\cosh(x)$ is increasing for $x \geq 0$, the right-hand side is bounded above by $\gamma^\star y$. Thus, for $y \geq 0$,

$$\tanh[\beta(s^\star + y)] \leq s^\star + \gamma^\star y. \tag{5.3}$$

Hence,

$$\mathbf{E}_\sigma[S_{t+1}^+ - S_t^+ \mid S_t^+ = s] \leq -(s - s^\star)\frac{(1-\gamma^\star)}{n},$$

from which (5.2) follows.

By (5.2),

$$Y_t := \left[1 - \frac{(1-\gamma^\star)}{n}\right]^{-t} (S_t^+ - s^\star)$$

defines a non-negative supermartingale for $t < \tau^\star$. By optional stopping,

$$1 \geq \mathbf{E}_\sigma[Y_{\tau^\star \wedge t}] \geq \mathbf{E}_\sigma\left[(1-(1-\gamma^\star)/n)^{-t\wedge\tau^\star}(S^+_{\tau^\star\wedge t} - s^\star)\right]$$
$$\geq c_1 n^{-1/2}[1-(1-\gamma^\star)/n]^{-t}\mathbf{P}_\sigma(\tau^\star > t).$$

Hence $\mathbf{P}_\sigma(\tau^\star > t) \leq c_1 n^{-1/2}[1-(1-\gamma)/n]^t$, and the lemma is proved. ∎

**Proposition 5.2.** *Let* $\beta > 1$. *For* $c_3 > 0$, *if*

$$\tau_\star = \tau_\star(c_3) := \min\{t \geq 0 \,:\, S^+_t \geq s^\star + c_3 n^{-1/2}\},$$

*then*

$$\mathbf{E}_0[\tau_\star] = O(n \log n). \tag{5.4}$$

Proposition 5.2 is proved in Section 5.2. Meanwhile, we state and prove Theorem 5.3 below, which establishes the upper bound.

**Theorem 5.3.** *Let* $\beta > 1$. *There is a constant* $c(\beta)$ *so that* $t_{\mathrm{mix}}(n) \leq c(\beta) n \log n$ *for the Glauber dynamics restricted to* $\Omega^+$.

*Proof.* We show that there is a coupling $(X^+_t, \tilde{X}^+_t)$ of the restricted Glauber dynamics started from states $\sigma$ and $\tilde{\sigma}$ such that, if $\tau_{\mathrm{mag}}$ is the first time $t$ with $S^+_t = \tilde{S}^+_t$, then

$$\limsup_{n\to\infty} \mathbf{P}_{\sigma,\tilde{\sigma}}(\tau_{\mathrm{mag}} > cn\log n) \to 0 \quad \text{as } c \to \infty.$$

An application of Lemma 2.9 will then complete the proof.

By monotonicity, it is enough to consider the the starting positions 0 and 1. The "top" chain with starting position 1 we denote by $(S^T_t)$, and the "bottom" chain with starting position 0 we denote by $(S^B_t)$. Let $\mu^+$ be the stationary distribution of the restricted magnetization chain, and let $(S_t)$ be a stationary copy of the restricted magnetization chain, that is, started with initial distribution $\mu^+$.

Initially, all the chains are independent of one another. Given constants $c_1 \leq c_2$, let

$$\tau_1 = \min\{t \geq 0 \,:\, S^T_t \leq s^\star + c_1 n^{-1/2}\},$$
$$\tau_2 = \min\{t \geq 0 \,:\, S^B_t \geq s^\star + c_2 n^{-1/2}\}.$$

Suppose that $\tau_1 \leq \tau_2$. On the event $S_{\tau_1} \geq s^\star + c_1 n^{-1/2}$, for $t \geq \tau_1$ we couple together monotonically the $S$-chain and the $S^T$-chain (that is, such that $S_t \geq S^T_t$ for all $t \geq \tau_1$), and continue to evolve the $S^B$-chain independently of $S_t$ and $S^T_t$. On the event $S_{\tau_1} < s^\star + c_1 n^{-1/2}$, we continue to run all three chains independently. Then at time $\tau_2$, on the event that $S_{\tau_2} \leq s^\star + c_2 n^{-1/2}$, couple together all three chains monotonically (so that $S^T_t \leq S_t \leq S^B_t$ for all $t \geq \tau_2$). If $S_{\tau_2} > s^\star + c_2$, just let the chains run independently. The case $\tau_2 < \tau_1$ is handled analogously.

Note that, since $(S_t)$ is independent of $(S^T_t)$ until after time $\tau_1$, the random variable $S_{\tau_1}$ is independent of $\tau_1$ and hence still stationary.

Let $c_3 > 0$ be a constant, and define events $H_1, H_2$ by

$$H_1 = \{\tau_1 \leq c_3 n \log n\} \cap \{S_{\tau_1} \geq s^\star + c_1 n^{-1/2}\},$$
$$H_2 = \{\tau_2 \leq c_3 n \log n\} \cap \{S_{\tau_2} \leq s^\star + c_2 n^{-1/2}.\}.$$

Then

$$\mathbf{P}_{\sigma,\tilde{\sigma}}(H_1^c) \leq \mathbf{P}_{\sigma,\tilde{\sigma}}(\tau_1 > c_3 n \log n) + \mu^+(0, s^\star + c_1 n^{-1/2}), \tag{5.5}$$

and

$$\mathbf{P}_{\sigma,\tilde{\sigma}}(H_2^c) \leq \mathbf{P}_{\sigma,\tilde{\sigma}}(\tau_2 > c_3 n \log n) + \mu^+(s^\star + c_2 n^{-1/2}, 1). \tag{5.6}$$

Now observe that on the event $H_1 \cap H_2$ the chains $(S^T_t)$ and $(S^B_t)$ have crossed over by the time $c_3 n \log n$, and that by (5.5) and (5.6),

$$\mathbf{P}_{\sigma,\tilde{\sigma}}(H_1 \cap H_2) \geq 1 - \mathbf{P}_{\sigma,\tilde{\sigma}}(\tau_1 > c_3 n \log n) - \mathbf{P}_{\sigma,\tilde{\sigma}}(\tau_2 > c_3 n \log n) - \mu^+(I^c),$$

where $I = (s^\star + c_1 n^{-1/2}, s^\star + c_2 n^{-1/2})$.

Since, as a consequence of Theorem 2.4 of Ellis, Newman, and Rosen (1980), the stationary magnetization satisfies a central limit theorem, $\mu^+(I^c) < 1$ uniformly in $n$. Further,

$$\lim_{n\to\infty} \mathbf{P}_{\sigma,\tilde{\sigma}}(\tau_1 > c_3 n \log n) = 0 \quad \text{and} \quad \lim_{n\to\infty} \mathbf{P}_{\sigma,\tilde{\sigma}}(\tau_2 > c_3 n \log n) = 0,$$

by Lemma 5.1 and Proposition 5.2, respectively. Hence the probability that $S^T$ and $S^B$ will have crossed by the time $c_3 n \log n$ stays bounded away from 0 as $n \to \infty$.

Finally, observe that, whenever the two chains cross, they coalesce with probability bounded away from 0 uniformly in $n$, which completes the proof. ■

5.1. **Hitting times for birth-and-death chains.** A *birth-and-death chain* on $\{0, 1, \ldots, N\}$ is a Markov chain $(Z_t)$ on $\mathbb{Z}^+$ with transitions $Z_{t+1} - Z_t$ contained in the set $\{-1, 0, 1\}$.

This section contains a few standard results concerning the hitting times of birth-and-death chains. We shall use these in the proof of Proposition 5.2 in the next section.

Define

$$\begin{aligned}
p_k &= \mathbf{P}(Z_{t+1} - Z_t = +1 \mid Z_t = k) && k = 0, 1, \ldots, N-1,\\
q_k &= \mathbf{P}(Z_{t+1} - Z_t = -1 \mid Z_t = k) && k = 1, \ldots, N,\\
r_k &= \mathbf{P}(Z_{t+1} - Z_t = 0 \mid Z_t = k) && k = 0, \ldots, N.
\end{aligned}$$

Clearly, $p_k + q_k + r_k = 1$ for all $k$ if we define $q_0 = p_N = 0$. Using $\pi$ to denote the stationary distribution of the chain, we have

$$\pi(1) = C_{p,q,r},$$
$$\pi(k) = C_{p,q,r} \prod_{j=1}^{k} \frac{p_{j-1}}{q_j}, \qquad k = 1, \ldots, N,$$

where $C_{p,q,r} = [1 + \sum_{k=1}^{n} p_{j-1} q_j^{-1}]^{-1}$ is a normalizing constant.

Now, let $\ell < N$ be a positive integer, and let $Z_t^{(\ell)}$ be a restriction of $Z_t$ to the set $\{0, \ldots, \ell\}$. In other words, when at $k \in \{0, \ldots, \ell - 1\}$, the chain makes transitions from $k$ as the original chain, but when at $\ell$, it moves to $\ell - 1$ with probability $q_\ell$ and stays at $\ell$ with probability $p_\ell + r_\ell$. Let $\pi^{(\ell)}$ be the stationary measure of $Z_t^{(\ell)}$. It is easy to verify that there is a constant $C_{p,q,r}^{\ell}$ such that

$$\pi^{(\ell)}(k) = C_{p,q,r}^{\ell} \pi(k) \quad \text{for } k = 0, 1, \ldots, \ell.$$

In other words, under the stationary measure of the restricted chain, the states $0, 1, \ldots, k$ each have the same relative weights as in the unrestricted chain.

For $k \in \{0, 1, \ldots, N\}$ let

$$\tau_k = \inf\{t \geq 0 \,:\, Z_t = k\},$$
$$\tau_k^+ = \inf\{t > 0 \,:\, Z_t = k\}.$$

Then (see for instance Levin, Peres, and Wilmer (2007)) for $k = 0, 1, \ldots, N-1$,

$$\frac{1}{\pi^{(\ell)}(\ell)} = \mathbf{E}_\ell^{(\ell)}[\tau_\ell^+] = 1 + q_\ell \mathbf{E}_{\ell-1}(\tau_\ell). \tag{5.7}$$

In the above, $\mathbf{E}_j$ and $\mathbf{E}_j^{\ell}$ respectively denote the expectation operators corresponding to the unrestricted and restricted chain starting in $j$. We shall now apply identity (5.7) to the Glauber dynamics magnetization chain.

## 5.2. **Hitting time for magnetization.**

*Proof of Proposition 5.2.* Here it is more convenient to work with $M_t = nS(X_t^+)/2$, which is a birth-and-death chain with values in $\{0, \ldots, n/2 - 1, n/2\}$. Note that, if $n$ is odd, this chain is not integer-valued, but this causes no difficulties, as one can simply shift all states by -1/2.

Let $\ell^\star = \lfloor ns^\star \rfloor$. Let $c > 0$ be a constant. Also, throughout the calculation, $C$ will denote a generic positive constant whose value may be adjusted between inequalities. In the notation of Section 5.1, we have for

$\ell \in \{1, \ldots, \lceil ns^\star + cn^{1/2} \rceil\}$,

$$\mathbf{E}_{\ell-1}[\tau_\ell] \leq \frac{1}{q_\ell \pi^{(\ell)}(\ell)}.$$

The probability of moving left, $q_\ell$, is bounded away from 0, uniformly in $\ell \in \{1, \ldots, n/2\}$. Consequently, writing $\ell = nx$ and $j = ny$, we obtain the upper bound

$$\mathbf{E}_{\ell-1}[\tau_\ell] \leq C \frac{\sum_{j=0}^{\ell} \binom{n}{n/2+ny} \exp\left(\beta 2ny^2\right)}{\binom{n}{n/2+nx} \exp\left(2\beta nx^2\right)}.$$

Applying Stirling's formula, the right-hand side is bounded above by

$$C \frac{\sum_{j=0}^{\ell} (1+y)^{-(1+2y)n/2}(1-2y)^{-(1-2y)n/2}(1-4y^2)^{-1/2} \exp\left(2\beta ny^2\right)}{(1+2x)^{-(1+2x)n/2}(1-2x)^{-(1-2x)n/2}(1-4x^2)^{-1/2} \exp\left(2\beta nx^2\right)},$$

which can be rewritten as

$$C \frac{\sum_{j=0}^{\ell} \exp\left[-nf(y)\right](1-4y^2)^{-1/2}}{\exp\left[-nf(x)\right](1-4x^2)^{-1/2}} = C \sum_{j=0}^{\ell} \exp\left[n(f(x) - f(y)\right] \left(\frac{1-4x^2}{1-4y^2}\right)^{1/2},$$

where

$$f(z) = \frac{1}{2}(1+2z)\log(1+2z) + \frac{1}{2}(1-2z)\log(1-2z) - 2\beta z^2.$$

Since $\ell/n \leq (\ell^\star + O(\sqrt{n}))/n < 1$ uniformly in $n$, we can bound

$$\sup_n \sup_{0 \leq y \leq s^\star = \ell^\star/n} \left(\frac{1-4x^2}{1-4y^2}\right)^{1/2} \leq C.$$

It follows that the behavior of each term in the sum is dominated by the behavior of the exponential factor $\exp\left[n(f(x) - f(y))\right]$, and so it is enough to upper bound the expression

$$\sum_{j=0}^{\ell} \exp\left[n(f(x) - f(y))\right].$$

We then need to look for stationary points of $f$ in the interval $[0, 1]$; we have

$$f'(z) = \log(1+2z) - \log(1-2z) - 4\beta z$$

$$f''(z) = \frac{1}{1-4z^2} - 4\beta,$$

so $f'(z) = 0$ if and only if

$$\frac{1+2z}{1-2z} = e^{4\beta z}, \tag{5.8}$$

or, equivalently,

$$2z = \tanh(2\beta z).$$

When $\beta < 1$, the unique maximum of $f$ is at $x = 0$. When $\beta > 1$, there is a local maximum of $f$ at $s = 0$, and as mentioned earlier, there is a unique $0 < s^\star < 1$ minimizing $f$. As before, we write $\ell^* = \lfloor ns^\star \rfloor$.

By the above, when $x < s^\star$,

$$\mathbf{E}_{\ell-1}[\tau_\ell] \le C \sum_{j=0}^{\ell} \exp\left[n(f(x) - f(y))\right],$$

and $f(x) \le f(y)$ for all $y \le x$.

Throughout the calculation below, we shall use the fact that $f'(y) < 0$ for all $y \in [0, s^\star)$, and that the second derivative $f''(y)$ exists and is uniformly bounded in that range, as $s^\star < 1/2$.

Suppose $x = O(n^{-1/2})$, i.e. $\ell = O(\sqrt{n})$. Then

$$\begin{aligned}
\mathbf{E}_{\ell-1}[\tau_\ell] &\le C \sum_{j=0}^{\ell} \exp\left[2f'(x)(nx - ny) + O(n(x-y)^2)\right] \\
&\le C \sum_{j=0}^{\ell} \exp\left[(f'(\ell/n)(\ell - j)\right] \\
&\le \sqrt{n}\left[1 + O(n^{-1/2})\right],
\end{aligned}$$

valid for $1 \le \ell \le C_1\sqrt{n}$. The final bound is valid as $f'(\ell/n) < 0$, and so each term is bounded by a constant.

Similarly (taking $C_1 = 20$) we have, for $20\sqrt{n} \le \ell \le \ell^\star/2$,

$$\begin{aligned}
\mathbf{E}_{\ell-1}[\tau_\ell] &\le C \sum_{j=0}^{\ell} \exp\left[f'(c_{\ell,y})(\ell - j) + O(n(x-y)^2)\right] \\
&\le C \sum_{j=0}^{\ell} \exp\left[f'(c_{\ell,y})(\ell - j)\right],
\end{aligned}$$

where $c_{\ell,y}$ is between $x$ and $y$ (we could take $c_{\ell,y} = x$, for each $y$, by the uniform boundedness of the second derivative). There exists a constant $c_1 > 0$ such that, if $j \ge \ell/2$, then $f'(c_{\ell,y}) \le -c_1\ell/n$. Then there exists a constant $c_2 > 0$ such that, for $j \le \ell/2$,

$$f(j/n) - f(\ell/n) \le -c_2.$$

This in turn implies that the sum of remaining terms is negligible. More precisely,

$$\sum_{j=0}^{\ell/2} \exp\left[n(f(\ell/n) - f(j/n))\right] \leq n\exp(-c_2 n).$$

It follows that

$$\begin{aligned}
\mathbf{E}_{\ell-1}[\tau_\ell] &\leq \sum_{j=\lfloor \ell/2 \rfloor}^{\ell} \exp\left[-c_1 \ell n^{-1}(\ell - j)\right] + n\exp(-c_2 n) \\
&\leq \frac{1}{1-\exp(-c_1\ell/n)} + n\exp(-c_2 n) \\
&\leq \frac{Cn}{\ell},
\end{aligned}$$

for some constant $C > 0$, uniformly in $n$.

Now suppose that $\ell^\star/2 \leq \ell \leq \ell^\star - 20\sqrt{n}$. Then, for some constant $\tilde{c}_1 > 0$, $f'(c_{\ell,y}) \leq -\tilde{c}_1(\ell^\star - \ell)/n$, as long as $j = yn \geq \ell/2$. Also, there exists a constant $\tilde{c}_2 > 0$ such that, for $j \leq \ell/2$,

$$f(j/n) - f(\ell/n) \leq -\tilde{c}_2,$$

and so the contribution due to the terms with $j \leq \ell/2$ is negligible.

Then a calculation similar to that for $20\sqrt{n} \leq \ell \leq \ell^*/2$ above implies that there is a constant $C > 0$ such that

$$\mathbf{E}_{\ell-1}[\tau_\ell] \leq \frac{Cn}{\ell^* - \ell},$$

uniformly in $n$. Similarly, if $\ell^\star - 20\sqrt{n} \leq \ell \leq \lceil ns^* + c\sqrt{n}\rceil$, then we see that

$$\mathbf{E}_{\ell-1}[\tau_\ell] = O(\sqrt{n}).$$

Summing over $\ell$, we obtain an upper bound on the expected hitting time of $\lceil ns^\star + c\sqrt{n}\rceil$ starting from 0, as follows:

$$\begin{aligned}
\mathbf{E}_0[\tau_{\ell^\star + c\sqrt{n}}] &= \sum_{\ell=0}^{\ell^\star + c\sqrt{n}} \mathbf{E}_{\ell-1}[\tau_\ell] \\
&\leq C\left(\sqrt{n} \times \sqrt{n} + \sum_{\ell=1}^{n} \frac{n}{\ell} + \sum_{\ell=\ell^\star-1}^{\ell^\star/2} \frac{n}{\ell^\star - \ell}\right) \\
&\leq C(n + n\log n),
\end{aligned}$$

where $C$ is once again a generic constant, and was changed to $2C$ in the last inequality. ■

Related results on the magnetization chain can be found in Olivieri and Vares (2005).

### 5.3. **Lower bound.**

**Theorem 5.4.** *Assume that $\beta > 1$. For the Glauber dynamics restricted to configurations with non-negative magnetization, $t_{\mathrm{mix}}(n) \geq (1/4)n \log n$.*

The Glauber dynamics restricted to configurations with non-negative magnetization will be denoted by $(X_t^+)$.

*Proof.* Recall again that $s^\star$ is the unique positive solution to $\tanh(\beta s^\star) = s^\star$.

Since we are proving a lower bound, it suffices to consider any specific starting state; we take $X_0^+$ to be the all plus configuration.

We let $(X_t^+, \tilde{X}_t^+)$ be the monotone coupling, where $X_0^+$ is the all plus configuration and $\tilde{X}_0^+$ has the stationary distribution $\mu^+$. We write $\mathbf{P}_{\mathbf{1},\mu^+}$ and $\mathbf{E}_{\mathbf{1},\mu^+}$ for the probability measure and expectation operator on the space where $(X_t^+, \tilde{X}_t^+)$ is defined.

Let $\mathcal{B}(\sigma) := \{i : \sigma(i) = -1\}$, and $B(\sigma) := |\mathcal{B}(\sigma)|$.

By the central limit theorem for the stationary magnetization, (c.f. Ellis, Newman, and Rosen (1980)), for some $0 < c_1 < 1$,

$$\mathbf{P}_{\mathbf{1},\mu^+}\left(B(\tilde{X}_0^+) \leq c_1 n\right) = \mu^+(\{\sigma : B(\sigma) \leq c_1 n\}) = o(1).$$

Let $N_t$ be the number of the sites in $\mathcal{B}(\tilde{X}_0^+)$ which have not been updated by time $t$. By writing $N_t$ as a sum of indicators,

$$\mathbf{E}_{\mathbf{1},\mu^+}\left[N_t \,\middle|\, B(\tilde{X}_0^+)\right] = B(\tilde{X}_0^+)[1 - n^{-1}]^t,$$

and so, for some $c_2 > 0$,

$$\mathbf{E}_{\mathbf{1},\mu^+}\left[N_{t_n^\star} \,\middle|\, B(\tilde{X}_0^+)\right] \geq c_2 B(\tilde{X}_0^+) n^{-1/4},$$

where $t_n^\star = (1/4)n \log n$. Also, since these indicators are negatively correlated, $\mathrm{Var}_{\mathbf{1},\mu^+}(N_t) \leq n$ for all $t$. Applying Chebyshev's inequality shows that, for some $c_3 > 0$, on the event $\{B(\tilde{X}_0^+) > c_1 n\}$,

$$\mathbf{P}_{\mathbf{1},\mu^+}\left(N_{t_n^\star} \leq c_3 n^{3/4} \,\middle|\, B(\tilde{X}_0^+)\right) = o(1),$$

where the $o(1)$ bound is uniform in $B$. We conclude that

$$\begin{aligned}
\mathbf{P}_{\mathbf{1},\mu^+}\left(N_{t_n^\star} \leq c_3 n^{3/4}\right) &\leq \mathbf{P}_{\mathbf{1},\mu^+}\left(B(\tilde{X}_0^+) \leq c_1 n\right) \\
&\quad + \mathbf{P}_{\mathbf{1},\mu^+}\left(N_{t_n^\star} \leq c_3 n^{3/4} \text{ and } B(\tilde{X}_0^+) > c_1 n\right) \\
&= o(1).
\end{aligned}$$

Suppose now that $N_{t_n^\star} > c_3 n^{3/4}$. It follows that $S_{t_n^\star} \geq \tilde{S}_{t_n^\star} + c_4 n^{-1/4}$ for some $c_4 > 0$. Thus, if $S_{t_n^\star} \leq s^\star + c_5 n^{-1/4}$ for a small constant $c_5 > 0$, then $\tilde{S}_{t_n^\star} \leq s^\star + (c_5 - c_4)n^{-1/4}$. Therefore,

$$\begin{aligned}
\mathbf{P}_{\mathbf{1},\mu^+}\left(S_{t_n^\star} \leq s^\star + c_5 n^{-1/4}\right) &\leq o(1) + \mathbf{P}_{\mathbf{1},\mu^+}\left(N_{t_n^\star} > c_3 n^{3/4} \text{ and } S_{t_n^\star} \leq s^\star + c_5 n^{-1/4}\right) \\
&\leq o(1) + \mathbf{P}_{\mathbf{1},\mu^+}\left(\tilde{S}_{t_n^\star} \leq s^\star + (c_5 - c_4)n^{-1/4}\right).
\end{aligned}$$

Again by the central limit theorem, the probability on the right-hand side above tends to 0 as $n \to \infty$, provided we choose $c_5 < c_4$.

On the other hand, appealing one final time to the central limit theorem,

$$\mu^+(\{\sigma : S(\sigma) > s^\star + c_5 n^{-1/4}\}) = o(1).$$

Consequently,

$$\begin{aligned} d_n(t_n^\star) &\geq \mathbf{P}_{\mathbf{1},\mu^+}\left(S_{t_n^\star} > s^\star + c_5 n^{-1/4}\right) \\ &\qquad - \mu^+(\{\sigma : S(\sigma) > s^\star + c_5 n^{-1/4}\}) \\ &= 1 - o(1), \end{aligned}$$

and so $t_{\mathrm{mix}}(n) \geq (1/4) n \log n$ for $n$ large. ■

## 6. Conjectures

We believe the results proven in this paper should be generic for Glauber dynamics on transitive graphs.

To be concrete, consider the $d$-dimensional torus $(\mathbb{Z}/n\mathbb{Z})^d$. Let $\beta_c$ be the critical temperature for uniqueness of Gibbs measures on $\mathbb{Z}^d$.

We make the following conjectures:

(i) For $\beta < \beta_c$, there is a cut-off.
(ii) For $\beta = \beta_c$, the mixing time is polynomial in $n$. A stronger conjecture is that there is a critical dimension $d_c$ such that for $d \geq d_c$, the mixing time $t_{\mathrm{mix}}$ is $O(|V_n|^{3/2})$.
(iii) For $\beta > \beta_c$, if the dynamics are suitably truncated, the mixing time is polynomial in $n$. A stronger version is that again there is a critical dimension $d_c$ such that for $d > d_c$, the mixing time is $O(|V_n| \log |V_n|)$.

## Acknowledgments

This research was initiated during a visit by DAL and MJL to Microsoft Research, and continued during both the DIMACS-Georgia Tech Phase Transition workshop and the Park City Mathematics Institute. Also, while working on this project, MJL was first supported in part by the Nuffield Foundation, and then in part by the Humboldt Foundation. We thank Elchanan Mossel for useful discussions at an early stage of this work, and Jian Ding and Eyal Lubetzky for comments on an early draft.

## References

Aizenman, M. and R. Holley. 1987. *Rapid convergence to equilibrium of stochastic Ising models in the Dobrushin Shlosman regime*, (Minneapolis, Minn., 1984), IMA Vol. Math. Appl., vol. 8, Springer, New York, pp. 1–11.

Aldous, D. and J. Fill. *Reversible Markov chains and random walks on graphs*, in progress. Manuscript available at http://www.stat.berkeley.edu/~aldous/RWG/book.html.

Bender, E. A., G. F. Lawler, R. Pemantle, and H. S. Wilf. 2003/04. *Irreducible compositions and the first return to the origin of a random walk*, Sém. Lothar. Combin. **50**, Art. B50h, 13 pp. (electronic), available at arxiv:math/0404253v1.

Bovier, A., M. Eckhoff, V. Gayrard, and M. Klein. 2001. *Metastability in stochastic dynamics of disordered mean-field models*, Probab. Theory Related Fields **119**, no. 1, 99–161.

______. 2002. *Metastability and low lying spectra in reversible Markov chains*, Comm. Math. Phys. **228**, no. 2, 219–255.

Bovier, A. and F. Manzo. 2002. *Metastability in Glauber dynamics in the low-temperature limit: beyond exponential asymptotics*, J. Statist. Phys. **107**, no. 3-4, 757–779.

Bubley, R. and M. Dyer. 1997. *Path coupling: A technique for proving rapid mixing in Markov chains*, Proceedings of the 38th Annual Symposium on Foundations of Computer Science, pp. 223–231.

Diaconis, P. 1996. *The cutoff phenomenon in finite Markov chains*, Proc. Nat. Acad. Sci. U.S.A. **93**, no. 4, 1659–1664.

Diaconis, P. and L. Saloff-Coste. 2006. *Separation cut-offs for birth and death chains*, Ann. Appl. Probab. **16**, no. 4, 2098–2122.

Ellis, R. S. 1985. *Entropy, large deviations, and statistical mechanics*, Grundlehren der Mathematischen Wissenschaften, vol. 271, Springer-Verlag, New York.

Ellis, R. S., C. M. Newman, and J. S. Rosen. 1980. *Limit theorems for sums of dependent random variables occurring in statistical mechanics. II. Conditioning, multiple phases, and metastability*, Z. Wahrsch. Verw. Gebiete **51**, no. 2, 153–169.

Feller, W. 1971. *An Introduction to Probability Theory and its Applications*, second edition, Vol. 2, J. Wiley & Sons, New York.

Griffiths, R. B., C.-Y. Weng, and J. S. Langer. 1966. *Relaxation Times for Metastable States in the Mean-Field Model of a Ferromagnet*, Phys. Rev. **149**, 301 –305.

Levin, D., Y. Peres, and E. Wilmer. 2007. *Markov Chains and Mixing Times*. In preparation, available at http://www.uoregon.edu/~dlevin/MARKOV/.

Lindvall, T. 2002. *Lectures on the coupling method*, Dover Publications Inc., Mineola, NY. Corrected reprint of the 1992 original.

Olivieri, E. and M. E. Vares. 2005. *Large deviations and metastability*, Encyclopedia of Mathematics and its Applications, vol. 100, Cambridge University Press, Cambridge.

Simon, B. and R. B. Griffiths. 1973. *The* $(\varphi^4)_2$ *field theory as a classical Ising model*, Comm. Math. Phys. **33**, 145–164.

Sinclair, A. 1993. *Algorithms for random generation and counting*, Progress in Theoretical Computer Science, Birkhäuser Boston Inc., Boston, MA. A Markov chain approach.

Department of Mathematics, University of Oregon, Eugene, Oregon 97402-1222
*E-mail address*: dlevin@uoregon.edu
*URL*: http://www.uoregon.edu/~dlevin

Department of Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom
*E-mail address*: m.j.luczak@lse.ac.uk
*URL*: http://www.lse.ac.uk/people/m.j.luczak@lse.ac.uk/

Microsoft Research and University of California, Berkeley
*E-mail address*: peres@microsoft.com