THE MIXING TIME EVOLUTION OF GLAUBER DYNAMICS FOR THE MEAN-FIELD ISING MODEL

JIAN DING, EYAL LUBETZKY AND YUVAL PERES

ABSTRACT. We consider Glauber dynamics for the Ising model on the complete graph on n vertices, known as the Curie-Weiss model. It is well-known that the mixing-time in the high temperature regime ($\beta < 1$) has order $n \log n$, whereas the mixing-time in the case $\beta > 1$ is exponential in n. Recently, Levin, Luczak and Peres proved that for any fixed $\beta < 1$ there is cutoff at time $\frac{1}{2(1-\beta)} n \log n$ with a window of order n, whereas the mixing-time at the critical temperature $\beta = 1$ is $\Theta(n^{3/2})$. It is natural to ask how the mixing-time transitions from $\Theta(n \log n)$ to $\Theta(n^{3/2})$ and finally to $\exp(\Theta(n))$. That is, how does the mixing-time behave when $\beta = \beta(n)$ is allowed to tend to 1 as $n \to \infty$.

In this work, we obtain a complete characterization of the mixing-time of the dynamics as a function of the temperature, as it approaches its critical point $\beta_c = 1$. In particular, we find a scaling window of order $1/\sqrt{n}$ around the critical temperature. In the high temperature regime, $\beta = 1 - \delta$ for some $0 < \delta < 1$ so that $\delta^2 n \to \infty$ with n, the mixing-time has order $(n/\delta) \log(\delta^2 n)$, and exhibits cutoff with constant $\frac{1}{2}$ and window size n/δ. In the critical window, $\beta = 1 \pm \delta$ where $\delta^2 n$ is $O(1)$, there is no cutoff, and the mixing-time has order $n^{3/2}$. At low temperature, $\beta = 1 + \delta$ for $\delta > 0$ with $\delta^2 n \to \infty$ and $\delta = o(1)$, there is no cutoff, and the mixing time has order $\frac{n}{\delta} \exp\left((\frac{3}{4} + o(1))\delta^2 n\right)$.

1. INTRODUCTION

The *Ising Model* on a finite graph $G = (V, E)$ with parameter $\beta \geq 0$ and no external magnetic field is defined as follows. Its set of possible *configurations* is $\Omega = \{1, -1\}^V$, where each configuration $\sigma \in \Omega$ assigns positive or negatives *spins* to the vertices of the graph. The probability that the system is at a given configuration σ is given by the *Gibbs distribution*

$$\mu_G(\sigma) = \frac{1}{Z(\beta)} \exp\left(\beta \sum_{xy \in E} \sigma(x)\sigma(y)\right),$$

where $Z(\beta)$ (the partition function) serves as a normalizing constant. The parameter β represents the inverse temperature: the higher β is (the lower the temperature is), the more μ_G favors configurations where neighboring
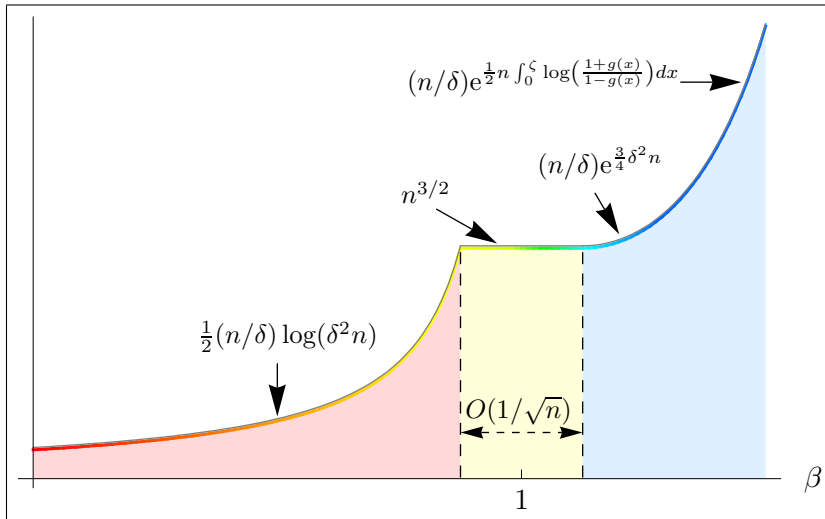
FIGURE 1. Illustration of the mixing time evolution as a function of the inverse-temperature β, with a scaling window of order $1/\sqrt{n}$ around the critical point. We write $\delta = |\beta - 1|$ and let ζ be the unique positive root of $g(x) := \frac{\tanh(\beta x) - x}{1 - x\tanh(\beta x)}$. Cutoff only occurs at high temperature.

spins are aligned. At the extreme case $\beta = 0$ (infinite temperature), the spins are totally independent and μ_G is uniform over Ω.

The *Curie-Weiss* model corresponds to the case where the underlying geometry is the complete graph on n vertices. The study of this model (see, e.g., [7],[8],[9],[11]) is motivated by the fact that its behavior approximates that of the Ising model on high-dimensional tori. It is convenient in this case to re-scale the parameter β, so that the stationary measure μ_n satisfies

$$\mu_n(\sigma) \propto \exp\left(\frac{\beta}{n}\sum_{x<y}\sigma(x)\sigma(y)\right) . \tag{1.1}$$

The *heat-bath Glauber dynamics* for the distribution μ_n is the following Markov Chain, denoted by (X_t). Its state space is Ω, and at each step, a vertex $x \in V$ is chosen uniformly at random, and its spin is updated as follows. The new spin of x is randomly chosen according to μ_n conditioned on the spins of all the other vertices. It can easily be shown that (X_t) is an aperiodic irreducible chain, which is reversible with respect to the stationary distribution μ_n.

We require several definitions in order to describe the mixing-time of the chain (X_t). For any two distributions ϕ, ψ on Ω, the *total-variation distance*

of ϕ and ψ is defined to be

$$\|\phi - \psi\|_{\mathrm{TV}} := \sup_{A \subset \Omega} |\phi(A) - \psi(A)| = \frac{1}{2} \sum_{\sigma \in \Omega} |\phi(\sigma) - \psi(\sigma)| \ .$$

The (worst-case) total-variation distance of (X_t) to stationarity at time t is

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

where $\mathbf{P}_\sigma$ denotes the probability given that $X_0 = \sigma$. The total-variation *mixing-time* of (X_t), denoted by $t_{\mathrm{MIX}}(\varepsilon)$ for $0 < \varepsilon < 1$, is defined to be

$$t_{\mathrm{MIX}}(\varepsilon) := \min\{t : d_n(t) \le \varepsilon\} \ .$$

A related notion is the spectral-gap of the chain, $\mathtt{gap} := 1 - \lambda$, where λ is the largest absolute-value of all nontrivial eigenvalues of the transition kernel.

Consider an infinite family of chains $(X_t^{(n)})$, each with its corresponding worst-distance from stationarity $d_n(t)$, its mixing-times $t_{\mathrm{MIX}}^{(n)}$, etc. We say that $(X_t^{(n)})$ exhibits *cutoff* iff for some sequence $w_n = o\big(t_{\mathrm{MIX}}^{(n)}(\frac{1}{4})\big)$ we have the following: for any $0 < \varepsilon < 1$ there exists some $c_\varepsilon > 0$, such that

$$t_{\mathrm{MIX}}^{(n)}(\varepsilon) - t_{\mathrm{MIX}}^{(n)}(1 - \varepsilon) \le c_\varepsilon w_n \quad \text{for all } n \ . \tag{1.2}$$

That is, there is a sharp transition in the convergence of the given chains to equilibrium at time $(1 + o(1))t_{\mathrm{MIX}}^{(n)}(\frac{1}{4})$. In this case, the sequence w_n is called a *cutoff window*, and the sequence $t_{\mathrm{MIX}}^{(n)}(\frac{1}{4})$ is called a *cutoff point*.

It is well known that for any fixed $\beta > 1$, the Glauber dynamics (X_t) mixes in exponential time (cf., e.g., [10]), whereas for any fixed $\beta < 1$ (high temperature) the mixing time has order $n \log n$ (see [2] and also [3]). Recently, Levin, Luczak and Peres [11] established that the mixing-time at the critical point $\beta = 1$ has order $n^{3/2}$, and that for fixed $0 < \beta < 1$ there is cutoff at time $\frac{1}{2(1-\beta)} n \log n$ with window n. It is therefore natural to ask how the phase transition between these states occurs around the critical $\beta_c = 1$: abrupt mixing at time $(\frac{1}{2(1-\beta)} + o(1)) n \log n$ changes to a mixing-time of $\Theta(n^{3/2})$ steps, and finally to exponentially slow mixing.

In this work, we determine this phase transition, and characterize the mixing-time of the dynamics as a function of the parameter β, as it approaches its critical value $\beta_c = 1$ both from below and from above. The scaling window around the critical temperature β_c has order $1/\sqrt{n}$, as formulated by the following theorems, and illustrated in Figure 1.

Theorem 1 (Subcritical regime)**.** *Let $\delta = \delta(n) > 0$ be such that $\delta^2 n \to \infty$ with n. The Glauber dynamics for the mean-field Ising model with parameter $\beta = 1 - \delta$ exhibits cutoff at time $\frac{1}{2}(n/\delta) \log(\delta^2 n)$ with window size n/δ. In addition, the spectral gap of the dynamics in this regime is $(1 + o(1))\delta/n$, where the $o(1)$-term tends to 0 as $n \to \infty$.*

Theorem 2 (Critical window)**.** *Let $\delta = \delta(n)$ satisfy $\delta = O(1/\sqrt{n})$. The mixing time of the Glauber dynamics for the mean-field Ising model with parameter $\beta = 1 \pm \delta$ has order $n^{3/2}$, and does not exhibit cutoff. In addition, the spectral gap of the dynamics in this regime has order $n^{-3/2}$.*

Theorem 3 (Supercritical regime)**.** *Let $\delta = \delta(n) > 0$ be such that $\delta^2 n \to \infty$ with n. The mixing-time of the Glauber dynamics for the mean-field Ising model with parameter $\beta = 1 + \delta$ does not exhibit cutoff, and has order*

$$t_{\exp}(n) := \frac{n}{\delta} \exp\left(\frac{n}{2} \int_0^\zeta \log\left(\frac{1 + g(x)}{1 - g(x)}\right) dx\right) \ ,$$

where $g(x) := (\tanh(\beta x) - x) / (1 - x \tanh(\beta x))$, and ζ is the unique positive root of g. In particular, in the special case $\delta \to 0$, the order of the mixing time is $\frac{n}{\delta} \exp\left((\frac{3}{4} + o(1))\delta^2 n\right)$, where the $o(1)$-term tends to 0 as $n \to \infty$. In addition, the spectral gap of the dynamics in this regime has order $1/t_{\exp}(n)$.

As we further explain in Section 2, the key element in the proofs of the above theorems is understanding the behavior of the sum of all spins (known as the *magnetization chain*) at different temperatures. This function of the dynamics turns out to be an ergodic Markov chain as well, and namely a *birth-and-death* chain (a one-dimensional chain, where only moves between neighboring positions are permitted). In fact, the reason for the exponential mixing at low-temperature is essentially that this magnetization chain has two centers of mass, $\pm\zeta n$ (where ζ is as defined in Theorem 3), with an exponential commute time between them. Figure 2 demonstrates how the single center of mass around 0 that this chain (rescaled) has at high near-critical temperature proceeds to split into two symmetric centers of mass that drift further and further apart as the temperature decreases.

In light of this, a natural question that rises is whether the above mentioned bottleneck between the two centers of mass at $\pm\zeta n$ is the *only* reason for the exponential mixing-time at low temperatures. Indeed, as shown in [11] for the strictly supercritical regime, $\beta > 1$ fixed, if one restricts the Glauber dynamics to non-negative magnetization (known as the *censored* dynamics), the mixing-time becomes $\Theta(n \log n)$ just like in the subcritical regime. Formally, the censored dynamics is defined as follows: at each step, a new state σ is generated according to the original rule of the Glauber dynamics, and if a negative magnetization is reached ($S(\sigma) < 0$) then σ is replaced by $-\sigma$. Interestingly, this simple modification suffices to boost the mixing-time back to order $n \log n$, just as in the high temperature case, and thus raises the question of whether the symmetry between the high temperature regime and the low temperature censored regime applies also to the existence of cutoff.
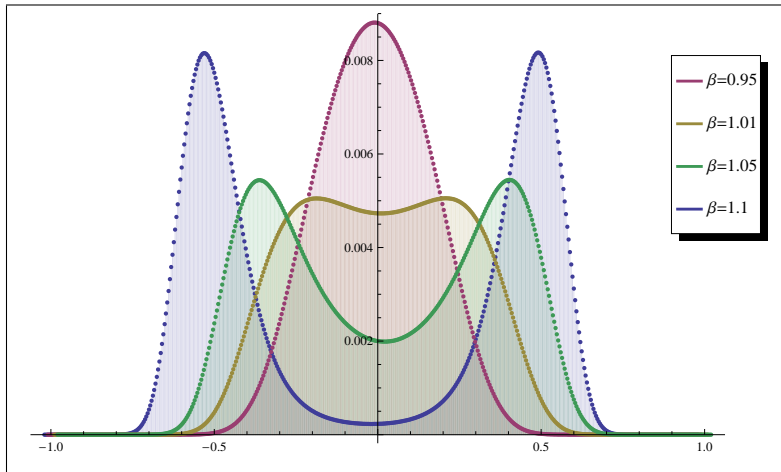
FIGURE 2. The stationary distribution of the normalized magnetization chain (average of all spins) for the dynamics on $n = 500$ vertices. The center of mass at high temperatures (see $\beta = 0.95$) is at 0. Low temperatures feature two centers of mass at $\pm\zeta$ (where ζ is the unique positive solution of $\tanh(\beta x) = x$), leading to the exponential mixing time.

In a companion paper [6], we strengthen the result of [11] by showing that the scaling window of $1/\sqrt{n}$ exists also for the censored low temperature case, beyond which cutoff indeed occurs (yet at a different location than in the symmetric high temperature point).

Theorem 4. *Let $\delta > 0$ be such that $\delta^2 n \to \infty$ arbitrarily slowly with n. Then the* censored *Glauber dynamics for the mean field Ising model with parameter $\beta = 1 + \delta$ has a cutoff at*

$$t_n = \left(\frac{1}{2} + \frac{1}{2(\zeta^2\beta/\delta - 1)}\right)\frac{n}{\delta}\log(\delta^2 n)$$

with a window of order n/δ. In the special case of the dynamics started from the all-plus configuration, the cutoff constant is $[2(\zeta^2\beta/\delta - 1)]^{-1}$ (the order of the cutoff point and the window size remain the same).

Theorem 5. *Let $\delta > 0$ be such that $\delta^2 n \to \infty$ arbitrarily slowly with n. Then the* censored *Glauber dynamics for the mean field Ising model with parameter $\beta = 1 + \delta$ has a spectral gap of order δ/n.*

Recalling Theorem 1, the above confirms that there is a symmetric scaling window of order $1/\sqrt{n}$ around the critical temperature, beyond which there is cutoff both at high and at low temperatures, with the same order of mixing-time (yet with a different constant), cutoff window and spectral gap.

The rest of this paper is organized as follows. Section 2 contains a brief outline of the proofs of the main theorems. Several preliminary facts on the Curie-Weiss model and on one-dimensional chains appear in Section 3. Sections 4, 5 and 6 address the high temperature regime (Theorem 1), critical temperature regime (Theorem 2) and low temperature regime (Theorem 3) respectively.

2. Outline of proof

In what follows, we present a sketch of the main ideas and arguments used in the proofs of the main theorems. We note that the analysis of the critical window relies on arguments similar to those used for the subcritical and supercritical regimes. Namely, to obtain the order of the mixing-time in Theorem 2 (critical window), we study the magnetization chain using the arguments that appear in the proof of Theorem 1 (high temperature regime). It is then straightforward to show that the mixing-time of the entire Glauber dynamics has the very same order. In turn, the spectral-gap in the critical window is obtained using arguments similar to those used in the proof of Theorem 3 (low temperature regime). In light of this, the following sketch will focus on the two non-critical temperature regimes.

2.1. **High temperature regime.**

Upper bound for mixing. As mentioned above, a key element in the proof is the analysis of the *normalized magnetization* chain, (S_t), which is the average spin in the system. That is, for a given configuration σ, we define $S(\sigma)$ to be $\frac{1}{n} \sum_i \sigma(i)$, and it is easy to verify that this function of the dynamics is an irreducible and aperiodic Markov chain. Clearly, a necessary condition for the mixing of the dynamics is the mixing of its magnetization, but interestingly, in our case the converse essentially holds as well. For instance, as we later explain, in the special case where the starting state is the all-plus configuration, by symmetry these two chains have precisely the same total variation distance from equilibrium at any given time.

In order to determine the behavior of the chain (S_t), we first keep track of its expected value along the Glauber dynamics. To simplify the sketch of the argument, suppose that our starting configuration is somewhere near the all-plus configuration. In this case, one can show that $\mathbf{E}S_t$ is monotone decreasing in t, and drops to order $\sqrt{1/\delta n}$ precisely at the cutoff point. Moreover, if we allow the dynamics to perform another $\Theta(n/\delta)$ steps (our cutoff window), then the magnetization will hit 0 (or $\frac{1}{n}$, depending on the parity of n) with probability arbitrarily close to 1. At that point, we essentially achieve the mixing of the magnetization chain.

It remains to extend the mixing of the magnetization chain to the mixing of the entire Glauber dynamics. Roughly, keeping in mind the above comment on the symmetric case of the all-plus starting configuration, one can apply a similar argument to an arbitrary starting configuration σ, by separately treating the set of spins which were initially positive and those which were initially negative. Indeed, it was shown in [11] that the following holds for $\beta < 1$ fixed (strictly subcritical regime). After a "burn-in" period of order n steps, the magnetization typically becomes not too biased. Next, if one runs two instances of the dynamics, from two such starting configurations (where the magnetization is not too biased), then by the time it takes their magnetization chains to coalesce, the entire configurations become relatively similar. This was established by a so-called *Two Coordinate Chain* analysis, where the two coordinates correspond to the current sum of spins along the set of sites which were initially either positive or negative respectively.

By extending the above Two Coordinate Chain Theorem to the case of $\beta = 1 - \delta$ where $\delta = \delta(n)$ satisfies $\delta^2 n \to \infty$, and combining it with second moment arguments and some additional ideas, we were able to show that the above behavior holds throughout this mildly subcritical regime. The burn-in time required for the typical magnetization to become "balanced" now has order n/δ, and so does the time it takes the full dynamics of two chains to coalesce once their magnetization chains have coalesced. Thus, these two periods are conveniently absorbed in our cutoff window, making the cutoff of the magnetization chain the dominant factor in the mixing of the entire Glauber dynamics.

Lower bound for mixing. While the above mentioned Two Coordinate Chain analysis was required in order to show that the entire Glauber dynamics mixes fairly quickly once its magnetization chain reaches equilibrium, the converse is immediate. Thus, we will deduce the lower bound on the mixing time of the dynamics solely from its magnetization chain.

The upper bound in this regime relied on an analysis of the first and second moments of the magnetization chain, however this approach is too coarse to provide a precise lower bound for the cutoff. We therefore resort to establishing an upper bound on the *third* moment of the magnetization chain, using which we are able to fine-tune our analysis of how its first moment changes along time. Examining the state of the system order n/δ steps before the alleged cutoff point, using concentration inequalities, we show that the magnetization chain is typically substantially far from 0. Recalling Figure 2, this implies a lower bound on the total variation distance of the magnetization chain to stationarity, as required.

Spectral gap analysis. In the previous arguments, we stated that the magnetization chain essentially dominates the mixing-time of the entire dynamics. An even stronger statement holds for the spectral gap: the Glauber dynamics and its magnetization chain have precisely the same spectral gap, and it is in both cases attained by the second largest eigenvalue. We therefore turn to establish the spectral gap of (S_t).

The lower bound follows directly from the contraction properties of the chain in this regime. To obtain a matching upper bound, we use the Dirichlet representation for the spectral gap, combined with an appropriate bound on the *fourth* moment of the magnetization chain.

2.2. **Low temperature regime.**

Exponential mixing. As mentioned above, the exponential mixing in this regime follows directly from the behavior of the magnetization chain, which has a bottleneck between $\pm\zeta$. To show this, we analyze the effective resistance between these two centers of mass, and obtain the precise order of the commute time between them. Additional arguments show that the mixing time of the entire Glauber dynamics in this regime has the same order.

Spectral gap analysis. In the above mentioned proof of the exponential mixing, we establish that the commute time of the magnetization chain between 0 and ζ has the same order as the hitting time from 1 to 0. We can therefore apply a recent result of [5] for general birth-and-death chains, which implies that in this case the inverse of the spectral-gap (known as the relaxation-time) and the mixing-time must have the same order.

3. Preliminaries

3.1. **The magnetization chain.** The normalized magnetization of a configuration $\sigma \in \Omega$, denoted by $S(\sigma)$, is defined as

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

Suppose that the current state of the Glauber dynamics is σ, and that site i has been selected to have its spin updated. By definition, the probability of updating this site to a positive spin is given by $p^+ \left(S(\sigma) - \sigma(i)/n \right)$, where

$$p^+(s) := \frac{\mathrm{e}^{\beta s}}{\mathrm{e}^{\beta s} + \mathrm{e}^{-\beta s}} = \frac{1 + \tanh(\beta s)}{2} \ . \tag{3.1}$$

Similarly, the probability of updating the spin of site i to a negative one is given by $p^- \left(S(\sigma) - \sigma(i)/n \right)$, where

$$p^-(s) := \frac{\mathrm{e}^{-\beta s}}{\mathrm{e}^{\beta s} + \mathrm{e}^{-\beta s}} = \frac{1 - \tanh(\beta s)}{2} \ . \tag{3.2}$$

It follows that the (normalized) magnetization of the Glauber dynamics at each step is a Markov chain, (S_t), with the following transition kernel:

$$P_M(s, s') = \begin{cases} \frac{1+s}{2} p^-(s - n^{-1}) & \text{if } s' = s - \frac{2}{n}, \\ \frac{1-s}{2} p^+(s + n^{-1}) & \text{if } s' = s + \frac{2}{n}, \\ 1 - \frac{1+s}{2} p^-(s - n^{-1}) - \frac{1-s}{2} p^+(s + n^{-1}) & \text{if } s' = s . \end{cases}$$

(3.3)

An immediate important property that the above reveals is the symmetry of S_t: the distribution of $(S_{t+1} \mid S_t = s)$ is precisely that of $(-S_{t+1} \mid S_t = -s)$.

As evident from the above transition rules, the behavior of the Hyperbolic tangent will be useful in many arguments. This is illustrated in the following simple calculation, showing that the minimum over the holding probabilities of the magnetization chain is nearly $\frac{1}{2}$. Indeed, since the derivative of $\tanh(x)$ is bounded away from 0 and 1 for all $x \in [0, \beta]$ and any $\beta = O(1)$, the Mean Value Theorem gives

$$\begin{array}{rcl} P_M(s, s + \frac{2}{n}) & = & \frac{1-s}{4} (1 + \tanh(\beta s)) + O(n^{-1}) , \\ P_M(s, s - \frac{2}{n}) & = & \frac{1+s}{4} (1 - \tanh(\beta s)) + O(n^{-1}) , \\ P_M(s, s) & = & \frac{1}{2} (1 + s \tanh(\beta s)) - O(n^{-1}) . \end{array}$$

(3.4)

Therefore, the holding probability in state s is at least $\frac{1}{2} - O\left(\frac{1}{n}\right)$. In fact, since $\tanh(x)$ is monotone increasing, $P_M(s, s) \leq \frac{1}{2} + \frac{1}{2} s \tanh(\beta s)$ for all s, hence these probability are also bounded from above by $\frac{1}{2}(1 + \tanh(\beta)) < 1$.

Using the above fact, the next lemma will provide an upper bound for the coalescence time of two magnetization chains, S_t and $\tilde{S}_t$, in terms of the hitting time τ_0, defined as $\tau_0 := \min\{t : |S_t| \leq n^{-1}\}$.

Lemma 3.1. *Let (S_t) and $(\tilde{S}_t)$ denote two magnetization chains, started from two arbitrary states. Then for any $\varepsilon > 0$ there exists some $c_\varepsilon > 0$, such that the following holds: if $T > 0$ satisfies $\mathbf{P}_1(\tau_0 \geq T) < \varepsilon$ then S_t and $\tilde{S}_t$ can be coupled in a way such that they coalesce within at most $c_\varepsilon T$ steps with probability at least $1 - \varepsilon$.*

Proof. Assume without loss of generality that $|\tilde{S}_0| < |S_0|$, and by symmetry, that $\sigma = |S_0| \geq 0$. Define

$$\tau := \min \left\{ t : |S_t| \leq |\tilde{S}_t| + \frac{2}{n} \right\} .$$

Recalling the definition of τ_0, clearly we must have $\tau < \tau_0$. Next, since the holding probability of S_t at any state s is bounded away from 0 and 1 for large n (by the discussion preceding the lemma), there clearly exists a constant $0 < b < 1$ such that

$$\mathbf{P}\left(S_{t+1} = \tilde{S}_{t+1} \,\big|\, |S_t - \tilde{S}_t| \leq \tfrac{2}{n} \right) > b > 0$$

(for instance, one may choose $b = \frac{1}{10}(1 - \tanh(\beta))$ for a sufficiently large n). It therefore follows that $|S_{\tau+1}| = |\tilde{S}_{\tau+1}|$ with probability at least b.

Condition on this event. We claim that in this case, the coalescence of (S_t) and $(\tilde{S}_t)$ (rather than just their absolute values) occurs at some $t \le \tau_0 + 1$ with probability at least b. The case $S_{\tau+1} = \tilde{S}_{\tau+1}$ is immediate, and it remains to deal with the case $S_{\tau+1} = -\tilde{S}_{\tau+1}$. Let us couple (S_t) and $(\tilde{S}_t)$ so that the property $S_t = -\tilde{S}_t$ is maintained henceforth. Thus, at time $t = \tau_0$ we obtain $|S_t - \tilde{S}_t| = 2|S_t| \le \frac{2}{n}$, and with probability b this yields $S_{t+1} = \tilde{S}_{t+1}$.

Clearly, our assumption on T and the fact that $0 \le \sigma \le 1$ together give

$$\mathbf{P}_\sigma(\tau_0 \ge T) \le \mathbf{P}_1(\tau_0 \ge T) < \varepsilon .$$

Thus, with probability at least $(1 - \varepsilon)b^2$, the coalescence time of (S_t) and $(\tilde{S}_t)$ is at most T. Repeating this experiment a sufficiently large number of times then completes the proof.   ∎

In order to establish cutoff for the magnetization chain (S_t), we will need to carefully track its moments along the Glauber dynamics. By definition (see (3.3)), the behavior of these moments is governed by the Hyperbolic tangent function, as demonstrated by the following useful form for the conditional expectation of S_{t+1} given S_t (see also [11, (2.13)]).

$$\mathbf{E}\,[S_{t+1} \mid S_t = s] = \left(s + \tfrac{2}{n}\right)P_M\!\left(s, s + \tfrac{2}{n}\right) + sP_M(s, s) + \left(s - \tfrac{2}{n}\right)P_M\!\left(s, s - \tfrac{2}{n}\right)$$
$$= (1 - n^{-1})s + \varphi(s) - \psi(s) , \tag{3.5}$$

where

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

3.2. From magnetization equilibrium to full mixing.

The motivation for studying the magnetization chain is that its mixing essentially dominates the full mixing of the Glauber dynamics. This is demonstrated by the next straightforward lemma (see also [11, Lemma 3.4]), which shows that in the special case where the starting point is the all-plus configuration, the mixing of the magnetization is *precisely* equivalent to that of the entire dynamics.

Lemma 3.2. *Let (X_t) be an instance of the Glauber dynamics for the mean field Ising model starting from the all-plus configuration, namely, $\sigma_0 = \mathbf{1}$, and let $S_t = S(X_t)$ be its magnetization chain. Then*

$$\|\mathbf{P_1}(X_t \in \cdot) - \mu_n\|_{\mathrm{TV}} = \|\mathbf{P_1}(S_t \in \cdot) - \pi_n\|_{\mathrm{TV}} , \tag{3.6}$$

where π_n is the stationary distribution of the magnetization chain.

Proof. For any $s \in \{-1, -1 + \frac{2}{n}, \dots, 1 - \frac{2}{n}, 1\}$, let $\Omega_s := \{\sigma \in \Omega : S(\sigma) = s\}$. Since by symmetry, both $\mu_n(\cdot \mid \Omega_s)$ and $\mathbf{P_1}(X_t \in \cdot \mid S_t = s)$ are uniformly distributed over Ω_s, the following holds:

$$
\begin{aligned}
\|\mathbf{P_1}(X_t \in \cdot) - \mu_n\|_{\mathrm{TV}} &= \frac{1}{2} \sum_s \sum_{\sigma \in \Omega_s} |\mathbf{P_1}(X_t = \sigma) - \mu_n(\sigma)| \\
&= \frac{1}{2} \sum_s \sum_{\sigma \in \Omega_s} \Big| \frac{\mathbf{P_1}(S_t = s)}{|\Omega_s|} - \frac{\mu_n(\Omega_s)}{|\Omega_s|} \Big| \\
&= \|\mathbf{P_1}(S_t \in \cdot) - \pi_n\|_{\mathrm{TV}} . \qquad \blacksquare
\end{aligned}
$$

In the general case where the Glauber dynamics starts from an arbitrary configuration σ_0, though the above equivalence (3.6) no longer holds, the magnetization still dominates the full mixing of the dynamics in the following sense. The full coalescence of two instances of the dynamics occurs within order $n \log n$ steps once the magnetization chains have coalesced.

Lemma 3.3 ([11, Lemma 2.9])**.** *Let $\sigma, \tilde{\sigma} \in \Omega$ be such that $S(\sigma) = S(\tilde{\sigma})$. For a coupling $(X_t, \tilde{X}_t)$, define the coupling time $\tau_{X,\tilde{X}} := \min\{t \geq 0 : X_t = \tilde{X}_t\}$. Then for a sufficiently large $c_0 > 0$ 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*

$$
\limsup_{n \to \infty} \mathbf{P}_{\sigma, \tilde{\sigma}} \left(\tau_{X,\tilde{X}} > c_0 n \log n \right) = 0 .
$$

Though Lemma 3.3 holds for any temperature, it will only prove useful in the critical and low temperature regimes. At high temperature, using more delicate arguments, we will establish full mixing within order of $\frac{n}{\delta}$ steps once the magnetization chains have coalesced. That is, the extra steps required to achieve full mixing, once the magnetization chain cutoff had occurred, are absorbed in the cutoff window. Thus, in this regime, the entire dynamics has cutoff precisely when its magnetization chain does (with the same window).

3.3. **Contraction and one-dimensional Markov chains.** We say that a Markov chain, assuming values in $\mathbb{R}$, is *contracting*, if the expected distance between two chains after a single step decreases by some factor bounded away from 0. As we later show, the magnetization chain is contracting at high temperatures, a fact which will have several useful consequences. One example of this is the following straightforward lemma of [11], which provides a bound on the variance of the chain. Here and throughout the paper, the notation $\mathbf{P}_z$, $\mathbf{E}_z$ and Var_z will denote the probability, expectation and variance respectively given that the starting state is z.

Lemma 3.4 ([11, Lemma 2.6])**.** *Let (Z_t) be a Markov chain taking values in $\mathbb{R}$ and with transition matrix P. Suppose that there is some $0 < \rho < 1$*

such that for all pairs of starting states $(z, \tilde{z})$,

$$| \mathbf{E}_z[Z_t] - \mathbf{E}_{\tilde{z}}[Z_t] | \leq \rho^t |z - \tilde{z}|. \tag{3.7}$$

Then $v_t := \sup_{z_0} \mathrm{Var}_{z_0}(Z_t)$ satisfies $v_t \leq v_1 \min \{t, 1/(1 - \rho^2)\}$.

Remark. By following the original proof of the above lemma, one can readily extend it to the case $\rho \geq 1$ and get the following bound:

$$v_t \leq v_1 \cdot \rho^{2t} \min \{t, 1/(\rho^2 - 1)\} . \tag{3.8}$$

This bound will prove to be effective for reasonably small values of t in the critical window, where although the magnetization chain is not contracting, ρ is only slightly larger than 1.

Another useful property of the magnetization chain in the high temperature regime is its drift towards 0. As we later show, in this regime, for any $s > 0$ we have $\mathbf{E}[S_{t+1}|S_t = s] < s$, and with probability bounded below by a constant we have $S_{t+1} < S_t$. We thus refer to the following lemma of [12]:

Lemma 3.5 ([12, Chapter 18]). *Let $(W_t)_{t \geq 0}$ be a non-negative supermartingale and τ be a stopping time such*

  (i) $W_0 = k$,
  (ii) $W_{t+1} - W_t \leq B$,
  (iii) $\mathrm{Var}(W_{t+1} \mid \mathcal{F}_t) > \sigma^2 > 0$ *on the event* $\tau > t$.
If $u > 4B^2/(3\sigma^2)$, then $\mathbf{P}_k(\tau > u) \leq \frac{4k}{\sigma \sqrt{u}}$.

This lemma, together with the above mentioned properties of (S_t), yields the following immediate corollary:

Corollary 3.6 ([11, 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 $u, t \geq 0$,*

$$\mathbf{P}(\, |S_u| > 0, \ldots, |S_{u+t}| > 0 \mid S_u = s) \leq \frac{cn|s|}{\sqrt{t}} . \tag{3.9}$$

Finally, our analysis of the spectral gap of the magnetization chain will require several results concerning birth-and-death chains from [5]. In what follows and throughout the paper, the relaxation-time of a chain, t_{REL}, is defined to be $\mathtt{gap}^{-1}$, where $\mathtt{gap}$ denotes its spectral-gap. We say that a chain is *b-lazy* if all its holding probabilities are at least b, or simply *lazy* for the useful case of $b = \frac{1}{2}$. Finally, given an ergodic birth-and-death chain on $\mathcal{X} = \{0, 1, \ldots, n\}$ with stationary distribution π, the *quantile state $Q(\alpha)$*, for $0 < \alpha < 1$, is defined to be the smallest $i \in \mathcal{X}$ such that $\pi(\{0, \ldots, i\}) \geq \alpha$.

Lemma 3.7 ([5, Lemma 2.9]). *Let $X(t)$ be a lazy irreducible birth-and-death chain on $\{0, 1, \ldots, n\}$, and suppose that for some $0 < \varepsilon < \frac{1}{16}$ we have $t_{\mathrm{REL}} < \varepsilon^4 \cdot \mathbf{E}_0 \tau_{Q(1-\varepsilon)}$. Then for any fixed $\varepsilon \leq \alpha < \beta \leq 1 - \varepsilon$:*

$$\mathbf{E}_{Q(\alpha)} \tau_{Q(\beta)} \leq \frac{3}{2\varepsilon} \sqrt{t_{\mathrm{REL}} \cdot \mathbf{E}_0 \tau_{Q(\frac{1}{2})}} \; . \tag{3.10}$$

Lemma 3.8 ([5, Lemma 2.3]). *For any fixed $0 < \varepsilon < 1$ and lazy irreducible birth-and-death chain X, the following holds for any t:*

$$\|P^t(0, \cdot) - \pi\|_{\mathrm{TV}} \leq \mathbf{P}_0(\tau_{Q(1-\varepsilon)} > t) + \varepsilon \; , \tag{3.11}$$

and for all $k \in \Omega$,

$$\|P^t(k, \cdot) - \pi\|_{\mathrm{TV}} \leq \mathbf{P}_k(\max\{\tau_{Q(\varepsilon)}, \tau_{Q(1-\varepsilon)}\} > t) + 2\varepsilon \; . \tag{3.12}$$

Remark. As argued in [5] (see Theorem 3.1 and its proof), the above two lemmas also hold for the case where the birth-and-death chain is not lazy but rather b-lazy for some constant $b > 0$. The formulation for this more general case incurs a cost of a slightly different constant in (3.10), and replacing t with t/C (for some constant C) in (3.11) and (3.12). As we already established (recall (3.4)), the magnetization chain is indeed b-lazy for any constant $b < \frac{1}{2}$ and a sufficiently large n.

3.4. **Monotone coupling.** A useful tool throughout our arguments is the *monotone coupling* of two instances of the Glauber dynamics (X_t) and $(\tilde{X}_t)$, which maintains a coordinate-wise inequality between the corresponding configurations. That is, given two configurations $\sigma \geq \tilde{\sigma}$ (i.e., $\sigma(i) \geq \tilde{\sigma}(i)$ for all i), it is possible to generate the next two states σ' and $\tilde{\sigma}'$ by updating the same site in both, in a manner that ensures that $\sigma' \geq \tilde{\sigma}'$. More precisely, we draw a random variable I uniformly over $\{1, 2, \ldots, n\}$ and independently draw another random variable U uniformly over $[0, 1]$. To generate σ' from σ, we update site I to $+1$ if $U \leq p^+ \left(S(\sigma) - \frac{\sigma(I)}{n} \right)$, otherwise $\sigma'(I) = -1$. We perform an analogous process in order to generate $\tilde{\sigma}'$ from $\tilde{\sigma}$, using the same I and U as before. The monotonicity of the function p^+ guarantees that $\sigma' \geq \tilde{\sigma}'$, and by repeating this process, we obtain a coupling of the two instances of the Glauber dynamics that always maintains monotonicity.

Clearly, the above coupling induces a monotone coupling for the two corresponding magnetization chains. We say that a birth-and-death chain with a transition kernel P and a state-space $\mathcal{X} = \{0, 1, \ldots, n\}$ is *monotone* if $P(i, i+1) + P(i+1, i) \leq 1$ for every $i < n$. It is easy to verify that this condition is equivalent to the existence of a monotone coupling, and that for such a chain, if $f : \mathcal{X} \to \mathbb{R}$ is a monotone increasing (decreasing) function then so is Pf (see, e.g., [5, Lemma 4.1]).

3.5. **The spectral gap of the dynamics and its magnetization chain.**
To analyze the spectral gap of the Glauber dynamics, we establish the following lemma which reduces this problem to determining the spectral-gap of the one-dimensional magnetization chain. Its proof relies on increasing eigenfunctions, following the ideas of [13].

Proposition 3.9. *The Glauber dynamics for the mean-field Ising model and its one-dimensional magnetization chain have the same spectral gap. Furthermore, both gaps are attained by the largest nontrivial eigenvalue.*

Proof. We will first show that the one-dimensional magnetization chain has an increasing eigenfunction, corresponding to the second eigenvalue.

Recalling that S_t assumes values in $\mathcal{X} := \{-1, -1 + \frac{2}{n}, \ldots, 1 - \frac{2}{n}, 1\}$, let M denote its transition matrix, and let π denote its stationary distribution. Let $1 = \theta_0 \geq \theta_1 \geq \ldots \geq \theta_n$ be the $n + 1$ eigenvalues of M, corresponding to the eigenfunctions $f_0 \equiv 1, f_1, \ldots, f_n$. Define $\theta = \max\{\theta_1, |\theta_n|\}$, and notice that, as S_t is aperiodic and irreducible, $0 < \theta < 1$. Furthermore, by the existence of the monotone coupling for S_t and the discussion in the previous subsection, whenever a function $f : \mathcal{X} \to \mathbb{R}$ is increasing so is Mf.

Define $f : I \to \mathbb{R}$ by $f := f_1 + f_n + K\mathbb{1}$, where $\mathbb{1}$ is the identity function and $K > 0$ is sufficiently large to ensure that f is monotone increasing (e.g., $K = \frac{n}{2}\|f_1 + f_n\|_{L^\infty}$ easily suffices). Notice that, by symmetry of S_t, $\pi(x) = \pi(-x)$ for all $x \in \mathcal{X}$, and in particular $\sum_{x \in \mathcal{X}} x\pi(x) = 0$, that is to say, $\langle \mathbb{1}, f_0 \rangle_{L^2(\pi)} = 0$. Recalling that for all $i \neq j$ we have $\langle f_i, f_j \rangle_{L^2(\pi)} = 0$, it follows that for some $q_1, \ldots, q_n \in \mathbb{R}$ we have $f = \sum_{i=1}^n q_i f_i$ with $q_1 \neq 0$ and $q_n \neq 0$, and thus

$$\left(\theta^{-1}M\right)^m f = \sum_{i=1}^n q_i(\theta_i/\theta)^m f_i .$$

Next, define

$$g = \left\{ \begin{array}{ll} q_1 f_1 & \text{if } \theta = \theta_1 \\ 0 & \text{otherwise} \end{array} \right. , \quad \text{and} \quad h = \left\{ \begin{array}{ll} q_n f_n & \text{if } \theta = -\theta_n \\ 0 & \text{otherwise} \end{array} \right. ,$$

and notice that

$$\lim_{m \to \infty} \left(\theta^{-1}M\right)^{2m} f = g + h , \quad \text{and} \quad \lim_{m \to \infty} \left(\theta^{-1}M\right)^{2m+1} f = g - h .$$

As stated above, $M^m f$ is increasing for all m, and thus so are the two limits $g+h$ and $g-h$ above, as well as their sum. We deduce that g is an increasing function, and next claim that $g \not\equiv 0$. Indeed, if $g \equiv 0$ then both h and $-h$ are increasing functions, hence necessarily $h \equiv 0$ as well; this would imply that $q_1 = q_n = 0$, thus contradicting our construction of f.

We deduce that g is an increasing eigenfunction corresponding to $\theta_1 = \theta$, and next wish to show that it is strictly increasing. Recall that for all $x \in \mathcal{X}$,

$$(Mg)(x) = M\left(x, x - \frac{2}{n}\right)g\left(x - \frac{2}{n}\right) + M(x,x)g(x) + M\left(x, x + \frac{2}{n}\right)g\left(x + \frac{2}{n}\right) .$$

Therefore, if for some $x \in \mathcal{X}$ we had $g(x - \frac{2}{n}) = g(x) \geq 0$, the fact that g is increasing would imply that

$$\theta_1 g(x) = (Mg)(x) \geq g(x) \geq 0 ,$$

and analogously, if $g(x) = g(x + \frac{2}{n}) \leq 0$ we could write

$$\theta_1 g(x) = (Mg)(x) \leq g(x) \leq 0 .$$

In either case, since $0 < \theta_1 < 1$ (recall that $\theta_1 = \theta$) this would in turn lead to $g(x) = 0$. By inductively substituting this fact in the above equation for $(Mg)(x)$, we would immediately get $g \equiv 0$, a contradiction.

Let $1 = \lambda_0 \geq \lambda_1 \geq \ldots \geq \lambda_{|\Omega|-1}$ denote the eigenvalues of the Glauber dynamics, and let $\lambda := \max\{\lambda_1, |\lambda_{2^n-1}|\}$. We translate g into a function $G : \Omega \to \mathbb{R}$ in the obvious manner:

$$G(\sigma) := g(S(\sigma)) = g\left(\frac{1}{n}\sum_{i=1}^{n}\sigma(i)\right) .$$

One can verify that G is indeed an eigenfunction of the Glauber dynamics corresponding to the eigenvalue θ_1, and clearly G is strictly increasing with respect to the coordinate-wise partial order on Ω. At this point, we refer to the following lemma of [13]:

Lemma 3.10 ([13, Lemma 4]). *Let P be the transition matrix of the Glauber dynamics, and let λ_1 be its second largest eigenvalue. If P has a strictly increasing eigenfunction f, then f corresponds to λ_1.*

The above lemma immediately implies that G corresponds to the second eigenvalue of Glauber dynamics, which we denote by λ_1, and thus $\lambda_1 = \theta_1$.

It remains to show that $\lambda = \lambda_1$. To see this, first recall that all the holding probabilities of S_t are bounded away from 0, and the same applies to the entire Glauber dynamics by definition (the magnetization remains the same if and only if the configuration remains the same). Therefore, both θ_n and λ_{2^n-1} are bounded away from -1, and it remains to show that $\texttt{gap} = o(1)$ for the Glauber dynamics (and hence also for its magnetization chain).

To see this, suppose P is the transition kernel of the Glauber dynamics, and recall the Dirichlet representation for the second eigenvalue of a reversible chain (see [12, Lemma 13.7], and also [1, Chapter 3]):

$$1 - \lambda_1 = \min\left\{\frac{\mathcal{E}(f)}{\langle f, f\rangle_{\mu_n}} \ : \ f \not\equiv 0 , \ \mathbf{E}_{\mu_n}(f) = 0\right\} , \qquad (3.13)$$

where $\mathbf{E}_{\mu_n}(f)$ denotes $\langle \mathbb{1}, f \rangle_{\mu_n}$, and

$$\mathcal{E}(f) = \langle (I - P)f, f \rangle_{\mu_n} = \frac{1}{2} \sum_{\sigma, \sigma' \in \Omega} \left[f(\sigma) - f(\sigma') \right]^2 \mu_n(\sigma) P(\sigma, \sigma') .$$

By considering the sum of spins, $h(\sigma) = \sum_{i=1}^n \sigma(i)$, we get $\mathcal{E}(h) \leq 2$, and since the spins are positively correlated, $\mathrm{Var}_{\mu_n} \sum_i \sigma(i) \geq n$. It follows that

$$1 - \lambda_1 \leq 2/n ,$$

and thus $\mathtt{gap} = 1 - \lambda_1 = 1 - \theta_1$ for both the Glauber dynamics and its magnetization chain, as required. ∎

4. High temperature regime

In this section we prove Theorem 1. Subsection 4.1 establishes the cutoff of the magnetization chain, which immediately provides a lower bound on the mixing time of the entire dynamics. The matching upper bound, which completes the proof of cutoff for the Glauber dynamics, is given in Subsection 4.2. The spectral gap analysis appears in Subsection 4.3. Unless stated otherwise, assume throughout this section that $\beta = 1 - \delta$ where $\delta^2 n \to \infty$.

4.1. **Cutoff for the magnetization chain.** Clearly, the mixing of the Glauber dynamics ensures the mixing of its magnetization. Interestingly, the converse is also essentially true, as the mixing of the magnetization turns out to be the most significant part in the mixing of the full Glauber dynamics. We thus wish to prove the following cutoff result:

Theorem 4.1. *Let $\beta = 1 - \delta$, where $\delta > 0$ satisfies $\delta^2 n \to \infty$. Then the corresponding magnetization chain (S_t) exhibits cutoff at time $\frac{1}{2} \cdot \frac{n}{\delta} \log(\delta^2 n)$ with a window of order n/δ.*

Notice that Lemma 3.2 then gives the following corollary for the special case where the initial state of the dynamics is the all-plus configuration:

Corollary 4.2. *Let $\delta = \delta(n) > 0$ be such that $\delta^2 n \to \infty$ with n, and let (X_t) denote the Glauber dynamics for the mean-field Ising model with parameter $\beta = 1 - \delta$, started from the all-plus configuration. Then (X_t) exhibits cutoff at time $\frac{1}{2}(n/\delta) \log(\delta^2 n)$ with window size n/δ.*

4.1.1. *Upper bound.* Our goal in this subsection is to show the following:

$$\lim_{\gamma \to \infty} \limsup_{n \to \infty} d_n \left(\frac{1}{2} \cdot \frac{n}{\delta} \log(\delta^2 n) + \gamma \frac{n}{\delta} \right) = 0 , \qquad (4.1)$$

where $d_n(\cdot)$ is with respect to the magnetization chain (S_t) and its stationary distribution. This will be obtained using an upper bound on the coalescence time of two instances of the magnetization chain. Given the properties of

its stationary distribution (see Figure 2), we will mainly be interested in the time it takes this chain to hit near 0. The following theorem provides an upper bound for that hitting time.

Theorem 4.3. *For* $0 < \beta < 1 + O(n^{-1/2})$, *consider the magnetization chain started from some arbitrary state* s_0, *and let* $\tau_0 = \min\{t : |S_t| \leq n^{-1}\}$. *Write* $\beta = 1 - \delta$, *and for* $\gamma > 0$ *define*

$$
t_n(\gamma) = \begin{cases} \dfrac{n}{2\delta} \log(\delta^2 n) + (\gamma + 3)\dfrac{n}{\delta} & \delta^2 n \to \infty , \\[2mm] \left(200 + 6\gamma\left(1 + 6\sqrt{\delta^2 n}\right)\right) n^{3/2} & \delta^2 n = O(1) . \end{cases}
\tag{4.2}
$$

Then there exists some $c > 0$ *such that* $\mathbf{P}_{s_0}(\tau_0 > t_n(\gamma)) \leq c/\sqrt{\gamma}$.

Proof. For any $t \geq 1$, define:

$$
s_t := \mathbf{E}_{s_0}\left[|S_t|\mathbf{1}_{\{\tau_0 > t\}}\right] .
$$

Suppose $s > 0$. Recalling (3.5) and bearing in mind the concavity of the Hyperbolic tangent and the fact that $\psi(s) \geq 0$, we obtain that

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

Using symmetry for the case $s < 0$, we can then deduce that

$$
\mathbf{E}\left[|S_{t+1}| \mid S_t\right] \leq |S_t| + \frac{1}{n}\left(\tanh(\beta|S_t|) - |S_t|\right) \text{ for any } t < \tau_0 .
\tag{4.3}
$$

Hence, combining the concavity of the Hyperbolic tangent together with Jensen's inequality yields

$$
s_{t+1} \leq \left(1 - \frac{1}{n}\right)s_t + \frac{1}{n}\tanh(\beta s_t) .
\tag{4.4}
$$

Since the Taylor expansion of $\tanh(x)$ is

$$
\tanh(x) = x - \frac{x^3}{3} + \frac{2x^5}{15} - \frac{17x^7}{315} + O(x^9) ,
\tag{4.5}
$$

we have $\tanh(x) \leq x - \frac{x^3}{5}$ for $0 \leq x \leq 1$, giving

$$
\begin{aligned}
s_{t+1} &\leq \left(1 - \frac{1}{n}\right)s_t + \frac{1}{n}\tanh(\beta s_t) \leq \left(1 - \frac{1}{n}\right)s_t + \frac{1}{n}\beta s_t - \frac{(s_t)^3}{5n} \\
&= s_t - \frac{\delta}{n}s_t - \frac{(s_t)^3}{5n} .
\end{aligned}
\tag{4.6}
$$

For some $1 < a \leq 2$ to be defined later, set

$$
b_i = a^{-i}/4 , \text{ and } u_i = \min\{t : s_t \leq b_i\} .
$$

Notice that s_t is decreasing in t by (4.6), thus for every $t \in [u_i, u_{i+1}]$ we have

$$
b_i/a = b_{i+1} \leq s_t \leq b_i .
$$

It follows that

$$s_{t+1} \leq s_t - \frac{\delta}{n} \cdot \frac{b_i}{a} - \frac{b_i^3}{5a^3 n} \ ,$$

and

$$u_{i+1} - u_i \leq \left(\frac{a-1}{a} b_i \right) / \left(\frac{\delta}{n} \frac{b_i}{a} + \frac{b_i^3}{5a^3 n} \right) \leq \frac{5(a-1)a^2 n}{5\delta a^2 + b_i^2} \ . \tag{4.7}$$

For the case $\delta^2 n \to \infty$, define:

$$i_0 = \min\{i : b_i \leq 1/\sqrt{\delta n}\} \ .$$

The following holds:

$$\sum_{i=1}^{i_0}(u_{i+1} - u_i) \leq \sum_{i=1}^{i_0} \frac{5(a-1)a^2 n}{5\delta a^2 + b_i^2} \leq \sum_{\substack{i \leq i_0 \\ b_i^2 > \delta}} \frac{5(a-1)a^2 n}{b_i^2} + \sum_{\substack{i \leq i_0 \\ b_i^2 \leq \delta}} \frac{5(a-1)a^2 n}{5\delta a^2}$$

$$\leq \frac{5na^2}{\delta(a+1)} + \frac{a-1}{2\log a} \cdot \frac{n}{\delta} \log(\delta^2 n) \ ,$$

where in the last inequality we used the fact that the series $\{b_i^{-2}\}$ is a geometric series with a ratio a^2, and that, as $b_i^2 \geq 1/(\delta n)$ for all $i \leq i_0$, the number of summands such that $b_i^2 \leq \delta$ is at most $\log_a(\sqrt{\delta^2 n})$. Therefore, choosing $a = 1 + n^{-1}$, we deduce that:

$$\sum_{i=1}^{i_0}(u_{i+1} - u_i) \leq \left(\frac{5}{2} + O(n^{-1}) \right) \frac{n}{\delta} + \left(\frac{1}{2} + O(n^{-1}) \right) \frac{n}{\delta} \log(\delta^2 n)$$

$$\leq 3\frac{n}{\delta} + \frac{n}{2\delta} \log(\delta^2 n) \ , \tag{4.8}$$

where the last inequality holds for any sufficiently large n. Combining the above inequality and the definition of i_0, we deduce that

$$(s_{t_n(0)} =) \ \mathbf{E}_{s_0} \left[|S_{t_n(0)}| \mathbf{1}_{\{\tau_0 > t_n(0)\}} \right] \leq 1/\sqrt{\delta n} \ . \tag{4.9}$$

Thus, by Corollary 3.6 (after taking expectation), for some fixed $c > 0$

$$\mathbf{P}(\tau_0 > t_n(\gamma)) \leq c/\sqrt{\gamma} \ .$$

For the case $\delta^2 n = O(1)$, choose $a = 2$, that is, $b_i = 2^{-(i+2)}$, and define

$$i_1 = \min\{i : b_i \leq n^{-1/4} \vee 5\sqrt{|\delta|}\} \ . \tag{4.10}$$

Substituting $a = 2$ in (4.7), while noting that $\delta > -\frac{1}{25}b_i^2$ for all $i < i_1$, gives

$$\sum_{i=1}^{i_1}(u_{i+1} - u_i) \leq \sum_{i=1}^{i_1} \frac{20n}{20\delta + b_i^2} \leq 100 \sum_{i=1}^{i_1} \frac{n}{b_i^2} \leq 200 \frac{n}{b_{i_1}^2}$$

$$\leq \left(200n^{3/2} \wedge 8\frac{n}{|\delta|} \right) \leq 200n^{3/2} \ .$$

where the last inequality in the first line incorporated the geometric sum over $\{b_i^{-2}\}$. By (4.10),

$$b_{i_1} \leq n^{-1/4} \vee 5\sqrt{|\delta|} \leq n^{-1/4}\left(1 + 5(\delta^2 n)^{1/4}\right) \ ,$$

and as in the subcritical case, we now combine the above results with an application of Corollary 3.6 (after taking expectation), and deduce that for some absolute constant $c > 0$,

$$\mathbf{P}(\tau_0 > t_n(\gamma)) \leq c/\sqrt{\gamma} \ ,$$

as required. ∎

Apart from drifting toward 0, and as we had previously mentioned, the magnetization chain at high temperatures is in fact contracting; this is a special case of the following lemma.

Lemma 4.4. *Let (S_t) and $(\tilde{S}_t)$ be the corresponding magnetization chains of two instances of the Glauber dynamics for some $\beta = 1 - \delta$ (where δ is not necessarily positive), and put $D_t := S_t - \tilde{S}_t$. The following then holds:*

$$\mathbf{E}[D_{t+1} - D_t \mid D_t] \leq -\frac{\delta}{n}D_t + \frac{|D_t|}{n^2} + O(n^{-4}) \ . \tag{4.11}$$

Proof. By definition (recall (3.5)), we have

$$\begin{aligned}
\mathbf{E}[D_{t+1} - D_t \mid D_t] &= \mathbf{E}[S_{t+1} - S_t + \tilde{S}_t - \tilde{S}_{t+1} \mid D_t] \\
&= \frac{\tilde{S}_t - S_t}{n} + \left[\varphi(S_t) - \varphi(\tilde{S}_t)\right] - \left[\psi(S_t) - \psi(\tilde{S}_t)\right] \ .
\end{aligned}$$

The Mean Value Theorem implies that

$$\varphi(S_t) - \varphi(\tilde{S}_t) \leq \frac{\beta}{n}(S_t - \tilde{S}_t) \ ,$$

and applying Taylor expansions on $\tanh(x)$ around βS_t and $\beta \tilde{S}_t$, we deduce that

$$\psi(S_t) - \psi(\tilde{S}_t) = \frac{S_t}{n^2 \cosh^2(\beta S_t)} - \frac{\tilde{S}_t}{n^2 \cosh^2(\beta \tilde{S}_t)} + O\left(\frac{1}{n^4}\right) \ .$$

Since the derivative of the function $x/\cosh^2(\beta x)$ is bounded by 1, another application of the Mean Value Theorem gives

$$\left|\psi(S_t) - \psi(\tilde{S}_t)\right| \leq \frac{|S_t - \tilde{S}_t|}{n^2} + O\left(\frac{1}{n^4}\right) \ .$$

Altogether, we obtain (4.11), as required. ∎

Indeed, the above lemma ensures that in the high temperature regime, $\beta = 1 - \delta$ where $\delta > 0$, the magnetization chain is contracting:

$$\mathbf{E}\left[|D_{t+1}| \,\middle|\, D_t\right] \le \left(1 - \frac{\delta}{2n}\right)|D_t| \quad \text{for any sufficiently large } n . \qquad (4.12)$$

We are now ready to prove that hitting near 0 essentially ensures the mixing of the magnetization.

Lemma 4.5. *Let $\beta = 1 - \delta$ for $\delta > 0$ with $\delta^2 n \to \infty$, (X_t) and $(\tilde{X}_t)$ be two instances of the dynamics started from arbitrary states σ_0 and $\tilde{\sigma}_0$ respectively, and (S_t) and $(\tilde{S}_t)$ be their corresponding magnetization chains. Let τ_{mag} denote the coalescence time $\tau_{\mathrm{mag}} := \min\{t : S_t = \tilde{S}_t\}$, and $t_n(\gamma)$ be as defined in Theorem 4.3. Then there exists some constant $c > 0$ such that*

$$\mathbf{P}\left(\tau_{\mathrm{mag}} > t_n(3\gamma)\right) \le c/\sqrt{\gamma} \quad \text{for all } \gamma > 0 . \qquad (4.13)$$

Proof. Set $T = t_n(\gamma)$. We claim that the following holds for large n:

$$|\mathbf{E}S_t| \le \frac{2}{\sqrt{\delta n}} \quad \text{and} \quad |\mathbf{E}\tilde{S}_t| \le \frac{2}{\sqrt{\delta n}} \quad \text{for all } t \ge T . \qquad (4.14)$$

To see this, first consider the case where n is even. The above inequality then follows directly from (4.9) and the decreasing property of s_t (see (4.6)), combined with the fact that $\mathbf{E}_0 S_t = 0$ (and thus $\mathbf{E}S_t = 0$ for all $t \ge \tau_0$). In fact, in case n is even, $|\mathbf{E}S_t|$ and $|\mathbf{E}\tilde{S}_t|$ are both at most $1/\sqrt{\delta n}$ for all $t \ge T$. For the case of n odd (where there is no 0 state for the magnetization chain, and τ_0 is the hitting time to $\pm\frac{1}{n}$), a simple way to show that (4.14) holds is to bound $|\mathbf{E}_{\frac{1}{n}} S_t|$. By definition, $P_M(\frac{1}{n}, \frac{1}{n}) \ge P_M(\frac{1}{n}, -\frac{1}{n})$ (see (3.3)). Combined with the symmetry of the positive and negative parts of the magnetization chain, one can then verify by induction that $P_M^t(\frac{1}{n}, \frac{k}{n}) \ge P_M^t(\frac{1}{n}, -\frac{k}{n})$ for any odd $k > 0$ and any t. Therefore, by symmetry as well as the fact that $\mathbf{E}_{s_0} S_t \le s_0$ for positive s_0, we conclude that $|\mathbf{E}_{\frac{1}{n}} S_t|$ is decreasing with t, and thus is bounded by $\frac{1}{n}$. This implies that (4.14) holds for odd n as well.

Combining (4.14) with the Cauchy-Schwartz inequality we obtain that for any $t \ge T$

$$\mathbf{E}|S_t - \tilde{S}_t| \le \mathbf{E}|S_t| + \mathbf{E}|\tilde{S}_t| \le \sqrt{\mathrm{Var}(S_t) + \frac{4}{\delta n}} + \sqrt{\mathrm{Var}(\tilde{S}_t) + \frac{4}{\delta n}} .$$

Now, combining Lemma 3.4 and Lemma 4.4 (and in particular, (4.12)), we deduce that $\mathrm{Var}\, S_t \le \frac{4}{\delta n}$, and plugging this into the above inequality gives

$$\mathbf{E}|S_t - \tilde{S}_t| \le \frac{10}{\sqrt{\delta n}} \quad \text{for any } t \ge T .$$

We next wish to show that within $2\gamma n/\delta$ additional steps, S_t and $\tilde{S}_t$ coalesce with probability at least $1 - c/\sqrt{\gamma}$ for some constant $c > 0$.

Consider time T, and let $D_t := S_t - \tilde{S}_t$. Recall that we have already established that

$$\mathbf{E}D_T \leq 10/\sqrt{\delta n} \ , \tag{4.15}$$

and assume without loss of generality that $D_T > 0$. We now run the magnetization chains S_t and $\tilde{S}_t$ independently for $T \leq t \leq \tau_1$, where

$$\tau_1 := \min \left\{ t \geq T : D_t \in \{0, -\tfrac{2}{n}\} \right\} \ ,$$

and let $\mathcal{F}_t$ be the σ-field generated by these two chains up to time t. By Lemma 4.4, we deduce that for sufficiently large values of n, if $D_t > 0$ then

$$\mathbf{E}[D_{t+1} - D_t \mid \mathcal{F}_t] \leq -\frac{\delta}{2n}D_t \leq 0 \ , \tag{4.16}$$

and D_t is a supermartingale with respect to $\mathcal{F}_t$. Hence, so is

$$W_t := D_{T+t} \cdot \frac{n}{2}\mathbf{1}_{\{\tau_1 > t\}} \ ,$$

and it is easy to verify that W_t satisfies the conditions of Lemma 3.5 (recall the upper bound on the holding probability of the magnetization chain, as well as the fact that at most one spin is updated at any given step). Therefore, for some constant $c > 0$,

$$\mathbf{P}\left(\tau_1 > t_n(2\gamma) \mid D_T\right) = \mathbf{P}(W_0 > 0, W_1 > 0, \ldots, W_{t_n(2\gamma)-T} > 0 \mid D_T)$$
$$\leq \frac{cnD_T}{\sqrt{\gamma n/\delta}} \ .$$

Taking expectation and plugging in (4.15), we get that for some constant c',

$$\mathbf{P}\left(\tau_1 > t_n(2\gamma)\right) \leq \frac{c'}{\sqrt{\gamma}} \ . \tag{4.17}$$

From time τ_1 and onward, we couple S_t and $\tilde{S}_t$ using a monotone coupling, thus D_t becomes a non-negative supermartingale with $D_{\tau_1} \leq \frac{2}{n}$. By (4.16),

$$\mathbf{E}\left[D_{t+1} - D_t \mid \mathcal{F}_t\right] \leq -\frac{\delta}{n^2} \ \text{ for } \tau_1 \leq t < \tau_{\mathrm{mag}} \ ,$$

and therefore, the Optional Stopping Theorem for non-negative supermartingales implies that, for some constant c'',

$$\mathbf{P}\left(\tau_{\mathrm{mag}} - \tau_1 \geq n/\delta\right) \leq \frac{\mathbf{E}(\tau_{\mathrm{mag}} - \tau_1)}{n/\delta} \leq \frac{c''}{\gamma} \ . \tag{4.18}$$

Combining (4.17) and (4.18) we deduce that for some constant c,

$$\mathbf{P}\left(\tau_{\mathrm{mag}} > t_n(3\gamma)\right) \leq \frac{c}{\sqrt{\gamma}} \ ,$$

completing the proof. ∎

4.1.2. *Lower bound.* We need to prove that the following statement holds for the distance of the magnetization at time t from stationarity:

$$\lim_{\gamma \to \infty} \liminf_{n \to \infty} d_n \left(\frac{1}{2} \cdot \frac{n}{\delta} \log(\delta^2 n) - \gamma \frac{n}{\delta} \right) = 1 \ . \tag{4.19}$$

The idea is to show that, at time $\frac{1}{2} \cdot \frac{n}{\delta} \log(\delta^2 n) - \gamma \frac{n}{\delta}$, the expected magnetization remains large. Standard concentration inequalities will then imply that the magnetization will typically be significantly far from 0, unlike its stationary distribution.

To this end, we shall first analyze the third moment of the magnetization chain. Recalling the transition rule (3.3) of S_t under the notations (3.1),(3.2)

$$p^+(s) = \frac{1 + \tanh(\beta s)}{2} \ , \ p^-(s) = \frac{1 - \tanh(\beta s)}{2} \ ,$$

the following holds:

$$\mathbf{E} \left[S_{t+1}^3 \mid S_t = s \right]$$

$$= \frac{1+s}{2} p^-(s - n^{-1}) \left(s - \frac{2}{n} \right)^3 + \frac{1-s}{2} p^+(s + n^{-1}) \left(s + \frac{2}{n} \right)^3$$

$$+ \left(1 - \frac{1+s}{2} p^-(s - n^{-1}) - \frac{1-s}{2} p^+(s + n^{-1}) \right) s^3$$

$$= s^3 + \frac{6s^2}{n} \cdot \frac{1}{4} \left(-2s + \tanh\left(\beta(s - n^{-1}) \right) + \tanh\left(\beta(s + n^{-1}) \right) \right.$$

$$\left. + s \left(\tanh\left(\beta(s - n^{-1}) \right) - \tanh\left(\beta(s + n^{-1}) \right) \right) \right) + c_1 \frac{s}{n^2} + \frac{c_2}{n^3} \ . \tag{4.20}$$

As $\tanh(x) \leq x$ for $x \geq 0$, for every $s > 0$ we get

$$\mathbf{E} \left[S_{t+1}^3 \mid S_t = s \right] \leq s^3 + \frac{3s^2}{2n} \left(-2s + \beta(s - n^{-1}) + \beta(s + n^{-1}) \right) + c_1 \frac{s}{n^2} + \frac{c_2}{n^3}$$

$$= s^3 - 3 \frac{\delta}{n} s^3 + \frac{c_1}{n^2} s + \frac{c_2}{n^3} \ . \tag{4.21}$$

If $s = 0$, the above also holds, since in that case $|S_{t+1}|^3 \leq (2/n)^3$. Finally, by symmetry, if $s < 0$ then the distribution of $|S_{t+1}^3| = -S_{t+1}^3$ given $S_t = s$ is the same as that of S_{t+1}^3 given $S_t = |s|$, and altogether we get:

$$\mathbf{E} \left[|S_{t+1}|^3 \mid S_t = s \right] \leq |s|^3 - 3 \frac{\delta}{n} |s|^3 + \frac{c_1}{n^2} |s| + \frac{c_2}{n^3} \ .$$

We deduce that

$$\mathbf{E}|S_{t+1}|^3 \leq \mathbf{E} \left(|S_t|^3 - 3 \frac{\delta}{n} |S_t|^3 + \frac{c_1}{n^2} |S_t| + \frac{c_2}{n^3} \right)$$

$$\leq \left(1 - \frac{3\delta}{n} \right) \mathbf{E}|S_t|^3 + \frac{c_1}{n^2} \mathbf{E}|S_t| + \frac{c_2}{n^3} \ . \tag{4.22}$$

Note that the following statement holds for the first moment of S_t:

$$\mathbf{E}_{s_0}\left[|S_t|\right] \leq \sqrt{(\mathbf{E}_{s_0} S_t)^2 + \mathrm{Var}_{s_0}(S_t)}$$
$$\leq \sqrt{(s_t)^2 + \frac{16}{\delta n}} \leq \left(1 - \frac{\delta}{n}\right)^t |s_0| + \frac{4}{\sqrt{\delta n}} \ .$$

Hence,

$$\mathbf{E}_{s_0}|S_{t+1}|^3 \leq \left(1 - \frac{3\delta}{n}\right)\mathbf{E}_{s_0}|S_t|^3 + \frac{c_1}{n^2}\left(1 - \frac{\delta}{n}\right)^t |s_0| + \frac{2}{n^2\sqrt{\delta n}} + \frac{c_2}{n^3}$$
$$= \eta^3 \mathbf{E}_{s_0}|S_t|^3 + \eta^t \frac{c_1}{n^2}|s_0| + \frac{4}{n^2\sqrt{\delta n}} + \frac{c_2'\delta^2}{n^2} \ ,$$

where $\eta = 1 - \delta/n$, and the extra error term involving c_2' absorbs the change of coefficient of $\mathbf{E}_{s_0}|S_t|^3$ and also the $1/n^3$ term. Iterating, we obtain

$$\mathbf{E}_{s_0}|S_{t+1}|^3 \leq \eta^{3t}|s_0|^3 + \eta^t \frac{c_1}{n^2}|s_0|\sum_{j=0}^{t}\eta^{2j} + \left(\frac{c_1'}{n^2\sqrt{\delta n}} + \frac{c_2'\delta^2}{n^2}\right)\sum_{j=0}^{t}\eta^{3j}$$
$$\leq \eta^{3t}|s_0|^3 + \eta^t \frac{c_1}{n^2}\cdot\frac{|s_0|}{1-\eta^2} + \left(\frac{c_1'}{n^2\sqrt{\delta n}} + \frac{c_2'\delta^2}{n^2}\right)\cdot\frac{1}{1-\eta^3}$$
$$\leq \eta^{3t}|s_0|^3 + \eta^t \frac{c_1}{\delta n}|s_0| + \frac{c_1'}{(\delta n)^{3/2}} + \frac{c_2'\delta}{n} \ . \tag{4.23}$$

Define $Z_t := |S_t|\eta^{-t}$, whence $Z_0 = |S_0| = |s_0|$. Recalling (3.5), and combining the Taylor expansion of $\tanh(x)$ given in (4.5) with the fact that $|\psi(s)| = O\left(s/n^2\right)$, we get that for $s > 0$

$$\mathbf{E}\left[|S_{t+1}| \,\middle|\, S_t = s\right] \geq \eta s - \frac{s^3}{2n} - \frac{s}{n^2} \ .$$

By symmetry, an analogous statement holds for $s < 0$, and altogether we obtain that

$$\mathbf{E}\left[|S_{t+1}| \,\middle|\, S_t\right] \geq \eta|S_t| - \frac{|S_t|^3}{2n} - \frac{|S_t|}{n^2} \ . \tag{4.24}$$

Remark. Note that (4.24) in fact holds for any temperature, having followed from the basic definition of the transition rule of (S_t), rather than from any special properties that this chain may have in the high temperature regime.

Rearranging the terms and multiplying by $\eta^{-(t+1)}$, we obtain that for any sufficiently large n,

$$\mathbf{E}\left[\left(1 - \frac{2}{n^2}\right)Z_t - Z_{t+1} \mid S_t\right] \leq \frac{1}{n}\eta^{-t}|S_t|^3 \ ,$$

where we used the fact that $\eta^{-1} \leq 2$. Taking expectation and plugging in (4.23), we deduce that

$$\mathbf{E}_{s_0}\Big[\Big(1 - \frac{2}{n^2}\Big)Z_t - Z_{t+1}\Big] \leq \frac{1}{n}\left(\eta^{2t}|s_0|^3 + \frac{c_1}{\delta n}|s_0| + \eta^{-t}\left(\frac{c_1'}{(\delta n)^{3/2}} + \frac{c_2'\delta}{n}\right)\right). \tag{4.25}$$

Set

$$\bar{t} = \frac{n}{2\delta}\log(\delta^2 n) - \gamma n/\delta ,$$

and notice that when n is sufficiently large, $\left(1 - \frac{2}{n^2}\right)^{-(t+1)} \leq 2$ for any $t \leq \bar{t}$. Therefore, multiplying (4.25) by $(1 - \frac{2}{n^2})^{-(t+1)}$ and summing over gives:

$$\begin{aligned}
|s_0| - 2\mathbf{E}_{s_0}Z_{\bar{t}} &\leq \frac{2|s_0|^3}{n(1 - \eta^2)} + \bar{t}\frac{c_1}{\delta n^2}|s_0| + \frac{2\eta^{-\bar{t}}}{n(1 - \eta)}\left(\frac{c_1'}{(\delta n)^{3/2}} + \frac{c_2'\delta}{n}\right) \\
&\leq \frac{2|s_0|^3}{\delta} + \frac{c_1\log(\delta^2 n)}{2\delta^2 n}|s_0| + \frac{c_1'}{\delta^{3/2}n} + \frac{c_2}{\delta n^{3/2}} + \frac{c_2'\delta}{\sqrt{n}} \\
&= \frac{2|s_0|^3}{\delta} + o(\sqrt{\delta} + |s_0|) ,
\end{aligned}$$

where the last inequality follows from the assumption $\delta^2 n \to \infty$. We now select $s_0 = \sqrt{\delta}/3$, which gives

$$\sqrt{\delta}/3 - 2\mathbf{E}_{s_0}Z_{\bar{t}} \leq 2\sqrt{\delta}/27 + o(\sqrt{\delta}) ,$$

and for a sufficiently large n we get

$$\mathbf{E}_{s_0}Z_{\bar{t}} \geq \sqrt{\delta}/9 .$$

Recalling the definition of Z_t, and using the well known fact that $(1 - x) \geq \exp(-x/(1 - x))$ for $0 < x < 1$, we get that for a sufficiently large n,

$$\mathbf{E}_{s_0}|S_{\bar{t}}| \geq \eta^{\bar{t}}\sqrt{\delta}/9 \geq \frac{\mathrm{e}^{\gamma/2}}{10\sqrt{\delta n}} =: L . \tag{4.26}$$

Lemma 3.4 implies that $\max\{\mathrm{Var}_{s_0}(S_t), \mathrm{Var}_{\mu_n}(\tilde{S}_t)\} \leq 16/\delta n$. Therefore, recalling that $\mathbf{E}_{\mu_n}\tilde{S}_{\bar{t}} = 0$, Chebyshev's inequality gives

$$\mathbf{P}_{s_0}(|S_{\bar{t}}| \leq L/2) \leq \mathbf{P}_{s_0}(||S_{\bar{t}}| - \mathbf{E}_{s_0}|S_{\bar{t}}|| \geq L/2) \leq \frac{16/(\delta n)}{L^2/4} = c\mathrm{e}^{-\gamma} ,$$

$$\mathbf{P}_{\mu_n}(|\tilde{S}_{\bar{t}}| \geq L/2) \leq \frac{16/(\delta n)}{L^2/4} = c\mathrm{e}^{-\gamma} .$$

Hence, letting π denote the stationary distribution of S_t, and A_L denote the set $\left[-\frac{L}{2}, \frac{L}{2}\right]$, we obtain that

$$\|\mathbf{P}_{s_0}(S_{\bar{t}} \in \cdot) - \pi\|_{\mathrm{TV}} \geq \pi(A_L) - \mathbf{P}_{s_0}(|S_{\bar{t}}| \in A_L) \geq 1 - 2c\mathrm{e}^{-\gamma} ,$$

which immediately gives (4.19). ∎

4.2. **Full Mixing of the Glauber dynamics.** In order to boost the mixing of the magnetization into the full mixing of the configurations, we will need the following result, which was implicitly proved in [11, Sections 3.3, 3.4] using a Two Coordinate Chain analysis. Although the authors of [11] were considering the case of $0 < \beta < 1$ fixed, one can follow the same arguments and extend this result to any $\beta < 1$. Following is this generalization of their result:

Theorem 4.6 ([11]). *Let (X_t) be an instance of the Glauber dynamics and μ_n the stationary distribution of the dynamics. Suppose X_0 is supported by*

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

For any $\sigma_0 \in \Omega_0$ and $\tilde{\sigma} \in \Omega$, we consider the dynamics (X_t) starting from σ_0 and an additional Glauber dynamics $(\tilde{X}_t)$ starting from $\tilde{\sigma}$, and define:

$$\tau_{\mathrm{mag}} := \min\{t : S(X_t) = S(\tilde{X}_t)\} \ ,$$

$$U(\sigma) := |\{i : \sigma(i) = \sigma_0(i) = 1\}| \ , \quad V(\sigma) := |\{i : \sigma(i) = \sigma_0(i) = -1\}| \ ,$$

$$\Xi := \{\sigma : \min\{U(\sigma), U(\sigma_0) - U(\sigma), V(\sigma), V(\sigma_0) - V(\sigma))\} \ge n/20\} \ ,$$

$$R(t) := \left| U(X_t) - U(\tilde{X}_t) \right| \ ,$$

$$H_1(t) := \{\tau_{\mathrm{mag}} \le t\} \ , \quad H_2(t_1, t_2) := \cap_{i=t_1}^{t_2} \{X_i \in \Xi \wedge \tilde{X}_i \in \Xi\} \ .$$

Then for any possible coupling of X_t and $\tilde{X}_t$, the following holds:

$$\max_{\sigma_0 \in \Omega_0} \|\mathbf{P}_{\sigma_0}(X_{r_2} \in \cdot) - \mu_n\|_{\mathrm{TV}} \le \max_{\substack{\sigma_0 \in \Omega_0 \\ \tilde{\sigma} \in \Omega}} \Big[\mathbf{P}_{\sigma_0, \tilde{\sigma}}(\overline{H_1(r_1)})$$

$$+ \mathbf{P}_{\sigma_0, \tilde{\sigma}}(R_{r_1} > \alpha\sqrt{n/\delta}) + \mathbf{P}_{\sigma_0, \tilde{\sigma}}(\overline{H_2(r_1, r_2)}) + \frac{\alpha c_1}{\sqrt{r_2 - r_1}} \cdot \sqrt{\frac{n}{\delta}} \Big] \ ,$$

(4.27)

where $r_1 < r_2$ and $\alpha > 0$.

The rest of this subsection will be devoted to establishing a series of properties satisfied by the magnetization throughout the mildly subcritical case, in order to ultimately apply the above theorem.

First, we shall show that any instance of the Glauber dynamics concentrates on Ω_0 once it performs an initial burn-in period of n/δ steps. It suffices to show this for the dynamics started from $s_0 = 1$: to see this, consider a monotone-coupling of the dynamics (X_t) starting from an arbitrary configuration, together with two additional instances of the dynamics, (X_t^+) starting from $s_0 = 1$ (from above) and (X_t^-) starting from $s_0 = -1$ (from below). By definition of the monotone-coupling, the chains (X_t^+) and (X_t^-) "trap" the chain (X_t), and by symmetry it indeed remains to show that

$$\mathbf{P}_1(|S_{t_0}| \le 1/2) = 1 - o(1) \ , \text{ where } t_0 = n/\delta \ .$$

Recalling (4.4), we have $s_{t+1} \leq (1 - \frac{\delta}{n})s_t$ where $s_t = \mathbf{E}\left[|S_t|\mathbf{1}_{\{\tau_0 > t\}}\right]$, thus

$$\mathbf{E}_1\left[|S_{t_0}|\mathbf{1}_{\{\tau_0 > t_0\}}\right] \leq e^{-1} .$$

Adding this to the fact that $\mathbf{E}_1 S_{t_0}\mathbf{1}\{\tau_0 \leq t_0\} = 0$, which follows immediately from symmetry, we conclude that $\mathbf{E}_1 S_{t_0} \leq e^{-1}$. Next, applying Lemma 3.4 to our case and noting that (3.7) holds for $\rho' = 1 - \frac{1}{n}\left(1 - n\tanh(\frac{\beta}{n})\right) \leq 1 - \frac{\delta}{n}$, we conclude that

$$\mathrm{Var}(S_t) \leq \nu_1 \frac{n}{\delta} \leq \left(\frac{4}{n}\right)^2 \frac{n}{\delta} = \frac{16}{\delta n} \quad \text{for all } t .$$

Hence, Chebyshev's inequality gives that $|S_{t_0}| \leq 1/2$ with high probability. We may therefore assume henceforth that our initial configuration already belongs to some good state $\sigma_0 \in \Omega_0$.

Next, set:

$$T := t_n(\gamma) \ , r_0 := t_n(2\gamma) \ , r_1 := t_n(3\gamma) \ , r_2 := t_n(4\gamma) .$$

We will next bound the terms in the righthand side of (4.27) in order. First, recall that Lemma 4.5 already provided us with a bound on the probability of $\overline{H_1(r_1)}$, by stating there for constant $c > 0$

$$\mathbf{P}(\tau_{\mathrm{mag}} > r_1) \leq \frac{c}{\sqrt{\gamma}} . \tag{4.28}$$

Our next task is to provide an upper bound on R_{r_1}, and namely, to show that it typically has order at most $\sqrt{n/\delta}$. In order to obtain such a bound, we will analyze the sum of the spins over the set $B := \{i : \sigma_0(i) = 1\}$. Define

$$M_t(B) := \frac{1}{2} \sum_{i \in B} X_t(i) ,$$

and consider the monotone-coupling of (X_t) with the chains (X_t^+) and (X_t^-) starting from the all-plus and all-minus positions respectively, such that $X_t^- \leq X_t \leq X_t^+$. By defining M_t^+ and M_t^- accordingly, we get that

$$\mathbf{E}(M_t(B))^2 \leq \mathbf{E}(M_t^+(B))^2 + \mathbf{E}(M_t^-(B))^2 = 2\mathbf{E}(M_t^+(B))^2 .$$

By (4.14), we immediately get that for $t \geq T$, $|\mathbf{E}M_t^+(B)| \leq \sqrt{\frac{n}{\delta}}$. We will next bound the variance of $M_t^+(B)$, by considering the following two cases:

(i) If every pair of spins of X_t^+ is positively correlated (since X_0^+ is the all-plus configuration, by symmetry, the covariances of each pair of spins is the same), then we can infer that

$$\mathrm{Var}(M_t^+(B)) \leq \mathrm{Var}\left(\frac{1}{2}\sum_{i \in [n]} X_t^+(i)\right) = \frac{n^2}{4}\mathrm{Var}\left(S(X_t^+)\right) \leq \frac{4n}{\delta} .$$

(ii) Otherwise, every pair of spins of X_t^+ is negatively correlated, and it follows that

$$\mathrm{Var}(M_t^+(B)) \leq \sum_{i \in B} \mathrm{Var}\left(\frac{1}{2}X_t^+(i)\right) \leq \frac{n}{4} \ .$$

Altogether, we conclude that for all $t \geq T$,

$$\mathbf{E}|M_t(B)| \leq \sqrt{\mathbf{E}\left(M_t(B)\right)^2} \leq \sqrt{2\,\mathrm{Var}(M_t^+(B) + 2\left(\mathbf{E}M_t(B)\right)^2}$$

$$\leq \sqrt{\frac{8n}{\delta} + \frac{2n}{\delta}} \leq 8\sqrt{\frac{n}{\delta}} \ . \tag{4.29}$$

This immediately implies that

$$\mathbf{E}R_{r_1} = \mathbf{E}|M_{r_1}(B) - \tilde{M}_{r_1}(B)| \leq \mathbf{E}|M_{r_1}(B)| + \mathbf{E}|\tilde{M}_{r_1}(B)| \leq 16\sqrt{\frac{n}{\delta}} \ ,$$

and an application of Markov's inequality now gives

$$\mathbf{P}(R_{r_1} \geq \alpha\sqrt{\frac{n}{\delta}}) \leq \frac{16}{\alpha} \ . \tag{4.30}$$

It remains to bound the probability of $\overline{H_2(r_1, r_2)}$. Define:

$$Y := \sum_{r_1 \leq t \leq r_2} \mathbf{1}\{|M_t(B)| > n/64\} \ ,$$

and notice that

$$\mathbf{P}\left(\bigcup_{t=r_1}^{r_2} \{|M_t(B)| \geq n/32\}\right) \leq \mathbf{P}(Y > n/64) \leq \frac{c_0 \mathbf{E}[Y]}{n} \ .$$

Recall that the second inequality of (4.29) actually gives $\mathbf{E}|M_t(B)|^2 \leq \frac{5n}{\delta}$. Hence, a standard second moment argument gives

$$\mathbf{P}(|M_t(B)| > n/64) = O\left(\frac{1}{\delta n}\right) \ .$$

Altogether, $\mathbf{E}_{\sigma_0} Y = O(1/\delta^2)$ and

$$\mathbf{P}_{\sigma_0}\left(\bigcup_{t=r_1}^{r_2} \{|M_t(B)| \geq n/32\}\right) = O\left(\frac{1}{\delta^2 n}\right) \ .$$

Applying an analogous argument to the chain $(\tilde{X}_t)$, we obtain that

$$\mathbf{P}_{\tilde{\sigma}}\left(\bigcup_{t=r_1}^{r_2} \left\{|\tilde{M}_t(B)| \geq n/32\right\}\right) = O\left(\frac{1}{\delta^2 n}\right) \ ,$$

and combining the last two inequalities, we conclude that

$$\mathbf{P}_{\sigma_0, \tilde{\sigma}}\left(\overline{H_2(r_1, r_2)}\right) = O\left(\frac{1}{\delta^2 n}\right) \ . \tag{4.31}$$

Finally, we have established all the properties needed in order to apply Theorem 4.6. At the cost of a negligible number of burn-in steps, the state of (X_t) with high probability belongs to Ω_0. We may thus plug in (4.28), (4.30) and (4.31) into Theorem 4.6, choosing $\alpha = \sqrt{\gamma}$, to obtain (4.1).

4.3. **Spectral gap Analysis.** By Proposition 3.9, it suffices to determine the spectral gap of the magnetization chain. The lower bound will follow from the next lemma of [4] (see also [12, Theorem 13.1]) along with the contraction properties of the magnetization chain.

Lemma 4.7 ([4]). *Suppose Ω is a metric space with distance ρ. Let P be a transition matrix for a Markov chain, not necessarily reversible. Suppose there exists a constant $\theta < 1$ and for each $x, y \in \Omega$, there is a coupling (X_1, Y_1) of $P(x, \cdot)$ and $P(y, \cdot)$ satisfying*

$$\mathbf{E}_{x,y}(\rho(X_1, Y_1)) \leq \theta\rho(x, y) \ .$$

If λ is an eigenvalue of P different from 1, then $|\lambda| \leq \theta$. In particular, the spectral gap satisfies $\mathtt{gap} \geq 1 - \theta$.

Recalling (4.12), the monotone coupling of S_t and $\tilde{S}_t$ implies that

$$\mathbf{E}_{s,\tilde{s}}|S_1 - \tilde{S}_1| \leq \left(1 - \frac{\delta}{n} + o\left(\frac{\delta}{n}\right)\right)|s - \tilde{s}| \ .$$

Therefore, Lemma 4.7 ensures that $\mathtt{gap} \geq (1 + o(1))\frac{\delta}{n}$.

It remains to show a matching upper bound on $\mathtt{gap}$, the spectral gap of the magnetization chain. Let M be the transition kernel of this chain, and π be its stationary distribution. Similar to our final argument in Proposition 3.9 (recall (3.13)), we apply the Dirichlet representation for the spectral gap (as given in [12, Lemma 13.7]) with respect to the function f being the identity map $\mathbb{1}$ on the space of normalized magnetization, we obtain that

$$\mathtt{gap} \leq \frac{\mathcal{E}(\mathbb{1})}{\langle \mathbb{1}, \mathbb{1} \rangle_\pi} = \frac{\langle (I - M)\mathbb{1}, \mathbb{1} \rangle_\pi}{\langle \mathbb{1}, \mathbb{1} \rangle_\pi} = 1 - \frac{\mathbf{E}_\pi\left[\mathbf{E}\left[S_t S_{t+1} \mid S_t\right]\right]}{\mathbf{E}_\pi S_t^2} \ , \qquad (4.32)$$

where $\mathbf{E}_\pi S_t^k$ is the k-th moment of the stationary magnetization chain (S_t). Recall (4.24) (where $\eta = 1 - \frac{\delta}{n}$), and notice that the following slightly stronger inequality in fact holds:

$$\mathbf{E}\left[\mathrm{sign}(S_t)S_{t+1} \mid S_t\right] \geq \eta|S_t| - \frac{|S_t|^3}{2n} - \frac{|S_t|}{n^2} \ .$$

(to see this, one needs to apply the same argument that led to (4.24), then verify the special cases $S_t \in \{0, \frac{1}{n}\}$). It thus follows that

$$\mathbf{E}\left[S_t S_{t+1} \mid S_t\right] \geq \eta S_t^2 - \frac{S_t^4}{2n} - \frac{S_t^2}{n^2} \ ,$$

and plugging the above into (4.32) we get

$$\texttt{gap} \leq \frac{\delta}{n} + \frac{1}{2n} \cdot \frac{\mathbf{E}_\pi S_t^4}{\mathbf{E}_\pi S_t^2} + \frac{1}{n^2} \ . \tag{4.33}$$

In order to bound the second term in (4.33), we need to give an upper bound for the fourth moment in terms of the second moment. The next argument is similar to the one used earlier to bound the third moment of the magnetization chain (see (4.20)), and hence will be described in a more concise manner.

For convenience, we use the abbreviations $h^+ := \tanh\big(\beta(s + n^{-1})\big)$ and $h^- := \tanh\big(\beta(s - n^{-1})\big)$. By definition (see (3.3)) the following then holds:

$$\begin{aligned}
\mathbf{E}[S_{t+1}^4 \mid S_t = s] &= s^4 + \frac{2}{n}s^3\left(-2s + h^- + h^+ + sh^- - h^+\right) \\
&\quad + \frac{6}{n^2}s^2\left(2 + h^+ - h^- - sh^- + h^+\right) \\
&\quad + \frac{8}{n^3}s^3\left(-2s + h^- + h^+ + sh^- - h^+\right) \\
&\quad + \frac{4}{n^4}\left(2 + h^+ - h^- - sh^- + h^+\right) \\
&\leq \left(1 - \frac{4\delta}{n}\right)s^4 + \frac{12}{n^2}s^2 + \frac{16}{n^4} \ .
\end{aligned}$$

Now, taking expectation and letting the S_t be distributed according to π, we obtain that

$$\mathbf{E}_\pi S_t^4 \leq \frac{3}{\delta n}\mathbf{E}_\pi S_t^2 + \frac{4}{\delta n^3} \ .$$

Recalling that, as the spins are positively correlated, $\mathrm{Var}_\pi(S_t) \geq \frac{1}{n}$, we get

$$\mathbf{E}_\pi S_t^4 \leq \left(3 + \frac{4}{n}\right)\frac{\mathbf{E}_\pi S_t^2}{n\delta} \ . \tag{4.34}$$

Plugging (4.34) into (4.33), we conclude that

$$\texttt{gap} \leq \frac{\delta}{n}\left(1 + O\Big(\frac{1}{\delta^2 n}\Big)\right) = (1 + o(1))\frac{\delta}{n} \ .$$

5. The critical window

In this section we prove Theorem 2, which establishes that the critical window has a mixing-time of order $n^{3/2}$ without a cutoff, as well as a spectral-gap of order $n^{-3/2}$.

5.1. **Upper bound.** Let (X_t) denote the Glauber dynamics, started from an arbitrary configuration σ, and let $(\tilde{X}_t)$ denote the dynamics started from the stationary distribution μ_n. As usual, let (S_t) and $(\tilde{S}_t)$ denote the (normalized) magnetization chains of (X_t) and $(\tilde{X}_t)$ respectively.

Let $\varepsilon > 0$. The case $\delta^2 n = O(1)$ of Theorem 4.3 implies that, for a sufficiently large $\gamma > 0$, $\mathbf{P}_\sigma\left(\tau_0 \geq \gamma n^{3/2}\right) < \varepsilon$. Plugging this into Lemma 3.1, we deduce that there exists some $c_\varepsilon > 0$, such that the chains S_t and $\tilde{S}_t$ coalesce after at most $c_\varepsilon n^{3/2}$ steps with probability at least $1 - \varepsilon$.

At this point, Lemma 3.3 implies that (X_t) and $(\tilde{X}_t)$ coalesce after at most $O(n^{3/2}) + O(n \log n) = O(n^{3/2})$ additional steps with probability arbitrarily close to 1, as required.

5.2. **Lower bound.** Throughout this argument, recall that δ is possibly negative, yet satisfies $\delta^2 n = O(1)$. By (4.22),

$$\mathbf{E}|S_{t+1}|^3 \leq \mathbf{E}\left(|S_t|^3 - 3\frac{\delta}{n}|S_t|^3 + \frac{c_1}{n^2}|S_t| + \frac{c_2}{n^3}\right)$$
$$\leq \left(1 - \frac{3\delta}{n}\right)\mathbf{E}|S_t|^3 + \frac{c_1}{n^2}\mathbf{E}|S_t| + \frac{c_2}{n^3} .$$

Recalling Lemma 4.4, and plugging the fact that $\delta = O(n^{-1/2})$ in (4.11), the following holds. If S_t and $\tilde{S}_t$ are the magnetization chains corresponding to two instances of the Glauber dynamics, then for some constant $c > 0$ and any sufficiently large n,

$$\mathbf{E}_{s,\tilde{s}}|S_1 - \tilde{S}_1| \leq (1 + cn^{-3/2})|s - \tilde{s}| . \tag{5.1}$$

Combining this with the extended form of Lemma 3.4, as given in (3.8), we deduce that if $t \leq \varepsilon n^{3/2}$ for some small fixed $\varepsilon > 0$, then $\mathrm{Var}_{s_0} S_t \leq 4t/n^2$. Therefore,

$$\mathbf{E}_{s_0}\left[|S_t|\right] \leq \sqrt{|\mathbf{E}_{s_0}S_t|^2 + \mathrm{Var}_{s_0} S_t} \leq \left(1 - \frac{\delta}{n}\right)^t |s_0| + \frac{2\sqrt{t}}{n} .$$

Therefore,

$$\mathbf{E}_{s_0}|S_{t+1}|^3 \leq \left(1 - \frac{3\delta}{n}\right)\mathbf{E}_{s_0}|S_t|^3 + \frac{c_1}{n^2}\left(1 - \frac{\delta}{n}\right)^t |s_0| + \frac{c_1'\sqrt{t}}{n^3}$$
$$\leq \eta^3 \mathbf{E}_{s_0}|S_t|^3 + \eta^t \frac{c_1}{n^2}|s_0| + \frac{c_1'\sqrt{t}}{n^3} ,$$

where again $\eta = 1 - \delta/n$. Iterating, we obtain

$$\mathbf{E}_{s_0}|S_{t+1}|^3 \leq \eta^{3t}|s_0|^3 + \eta^t \frac{c_1}{n^2}|s_0|\sum_{j=0}^t \eta^{2j} + \frac{c_1'\sqrt{t}}{n^3}\sum_{j=0}^t \eta^{3j}$$

$$\leq \eta^{3t}|s_0|^3 + \eta^t \frac{c_1}{n^2}\cdot\frac{\eta^{2t-1}-1}{\eta^2-1}|s_0| + \frac{c_1'\sqrt{t}}{n^3}\cdot\frac{\eta^{3t}-1}{\eta^3-1}$$

$$\leq \eta^{3t}|s_0|^3 + \eta^t \frac{c_1}{n^2}\cdot 2t|s_0| + \frac{c_1'\sqrt{t}}{n^3}\cdot 3t \ , \tag{5.2}$$

where the last inequality holds for sufficiently large n and $t \leq \varepsilon n^{3/2}$ with $\varepsilon > 0$ small enough (such a choice ensures that η^t will be suitably small). Define $Z_t := |S_t|\eta^{-t}$, whence $Z_0 = |S_0| = |s_0|$. Applying (4.24) (recall that it holds for any temperature) and using the fact that $\eta^{-1} \leq 2$, we get

$$\mathbf{E}[Z_{t+1} \mid S_t] \geq Z_t - \frac{1}{n}\left(\eta^{-t}|S_t|^3 + O(1/n)\right),$$

for n large enough, hence

$$\mathbf{E}[Z_t - Z_{t+1} \mid S_t] \leq \frac{1}{n}\left(\eta^{-t}|S_t|^3 + O(1/n)\right) \ .$$

Taking expectation and plugging in (4.23),

$$\mathbf{E}_{s_0}[Z_t - Z_{t+1}] \leq \frac{1}{n}\left(\eta^{2t}|s_0|^3 + \frac{c_2 t}{n^2}|s_0| + \eta^{-t}\frac{c_2' t^{3/2}}{n^3} + O(1/n)\right) \ . \tag{5.3}$$

Set $\bar{t} = n^{3/2}/A^4$ for some large constant A such that $\frac{1}{2} \leq \eta^{\bar{t}} \leq 2$. Summing over (5.3) we obtain that

$$|s_0| - \mathbf{E}_{s_0}Z_{\bar{t}} \leq \frac{1-\eta^{2\bar{t}}}{n(1-\eta^2)}|s_0|^3 + \bar{t}^2\frac{c_2}{n^3}|s_0| + 2\eta^{-\bar{t}}\cdot\bar{t}^{5/2}/n^4 + O(\bar{t}/n^2)$$

$$\leq \frac{2}{A^4}\sqrt{n}|s_0|^3 + \frac{c_2}{A^8}|s_0| + \frac{2}{A^{10}}e^{\sqrt{\delta^2 n}/A^4}n^{-1/4} + O(n^{-1/2}) \ .$$

We now select $s_0 = An^{-1/4}$ for some large constant A; this gives

$$An^{-1/4} - \mathbf{E}_{s_0}Z_{\bar{t}} \leq \left(\frac{2}{A} + \frac{c_2}{A^7} + \frac{2}{A^{10}}e^{\sqrt{\delta^2 n}/A^4}\right)n^{-1/4} + O(n^{-1/2}) \ .$$

Choosing A large enough to swallow the constant c_2 as well as the term $\delta^2 n$ (using the fact that $\delta^2 n$ is bounded), we obtain that

$$\mathbf{E}_{s_0}Z_{\bar{t}} \geq \frac{1}{2}An^{-1/4} \ .$$

Translating Z_t back to $|S_t|$, we obtain

$$\mathbf{E}_{s_0}|S_{\bar{t}}| \geq \eta^{\bar{t}}\cdot\frac{1}{2}An^{-1/4} \geq \sqrt{A}n^{-1/4} =: B \ , \tag{5.4}$$

provided that A is sufficiently large (once again, using the fact that $\eta^{\bar{t}}$ is bounded, this time from below). Since

$$\operatorname{Var}_{s_0}(S_{\bar{t}}) \leq 16\bar{t}/n^2 = \frac{16}{A^4}n^{-1/2} \ , \tag{5.5}$$

the following concentration result on the stationary chain $(\tilde{S}_t)$ will complete the proof:

$$\mathbf{P}_{\mu_n}(|\tilde{S}_t| \geq An^{-1/4})\} \leq \ \varepsilon(A) \ , \ \text{and} \ \lim_{A \to \infty} \varepsilon(A) = 0 \ . \tag{5.6}$$

Indeed, combining the above two statements, Chebyshev's inequality implies that

$$\|\mathbf{P}_{s_0}(S_{\bar{t}} \in \cdot) - \pi\|_{\mathrm{TV}} \geq \pi([-B/2, B/2]) - \mathbf{P}_{s_0}(|S_{\bar{t}}| \leq B/2)$$

$$\geq 1 - \frac{64}{A^5} - \varepsilon(\sqrt{A}) \ . \tag{5.7}$$

It remains to prove (5.6). Since we are proving a lower bound for the mixing-time, it suffices to consider a sub-sequence of the δ_n-s such that $\delta_n\sqrt{n}$ converges to some constant (possibly 0). The following result establishes the limiting stationary distribution of the magnetization chain in this case.

Theorem 5.1. *Suppose that* $\lim_{n \to \infty} \delta_n\sqrt{n} = \alpha \in \mathbb{R}$. *The following holds:*

$$\frac{S_{\mu_n}}{n^{-1/4}} \to \exp\left(-\frac{s^4}{12} - \alpha\frac{s^2}{2}\right) \ . \tag{5.8}$$

Proof. We need the following theorem:

Theorem 5.2 ([8, Theorem 3.9]). *Let ρ denote some probability measure, and let $S_n(\rho) = \frac{1}{n}\sum_{j=1}^{n} X_j(\rho)$, where the $\{X_j(\rho) : j \in [n]\}$ have joint distribution*

$$\frac{1}{Z_n}\exp\left[\frac{(x_1 + \ldots + x_n)^2}{2n}\right]\prod_{j=1}^{n} d\rho(x_j) \ ,$$

and Z_n is a normalization constant. Suppose that $\{\rho_n : n = 1, 2, \ldots\}$ are measures satisfying

$$\exp(x^2/2)d\rho_n \to \exp(x^2/2)d\rho \ . \tag{5.9}$$

Suppose further that ρ has the following properties:

(1) Pure: *the function*

$$G_\rho(s) := \frac{s^2}{2} - \log\int e^{sx} d\rho(x)$$

has a unique global minimum.

(2) Centered at m: *let m denote the location of the above global minimum.*

(3) Strength δ and type k: the parameters $k, \delta > 0$ are such that

$$G_\rho(s) = G_\rho(m) + \delta \frac{(s-m)^{2k}}{(2k)!} + o((s-m)^{2k}) ,$$

where the $o(\cdot)$-term tends to 0 as $s \to m$.

If, for some real numbers $\alpha_1, \ldots, \alpha_{2k-1}$ we have

$$G_{\rho_n}^{(j)}(m) = \frac{\alpha_j}{n^{1-j/2k}} + o(n^{-1+j/2k}) , \quad j = 1, 2, \ldots, 2k-1 , n \to \infty ,$$

then the following holds:

$$S_n(\rho_n) \to \mathbf{1}_{\{s \neq m\}}$$

and

$$\frac{S_n(\rho_n) - m}{n^{-1/2k}} \to \begin{cases} N\left(-\dfrac{\alpha_1}{\delta}, \dfrac{1}{\delta} - 1\right), & \text{if } k = 1 , \\[2ex] \exp\left(-\delta \dfrac{s^2 k}{(2k)!} - \displaystyle\sum_{j=1}^{2k-1} \alpha_j \dfrac{s^j}{j!}\right), & \text{if } k \geq 2 . \end{cases} ,$$

where $\delta^{-1} - 1 > 0$ for $k = 1$.

Let ρ denote the two-point uniform measure on $\{-1, 1\}$, and let ρ_n denote the two-point uniform measure on $\{-\beta_n, \beta_n\}$. As $|1 - \beta_n| = \delta_n = O(1/\sqrt{n})$, the convergence requirement (5.9) of the measures ρ_n is clearly satisfied. We proceed to verify the properties of ρ:

$$G_\rho(s) = \frac{s^2}{2} - \log \int e^{sx} d\rho(x) = \frac{s^2}{2} - \log \cosh(s) = \frac{s^4}{12} - \frac{s^6}{45} + O(s^8) .$$

This implies that G_ρ has a unique global minimum at $m = 0$, type $k = 2$ and strength $\delta = 2$. As $\delta_n \sqrt{n} \to \alpha$, we deduce that the G_{ρ_n}-s satisfy

$$G_{\rho_n}(s) = \frac{s^2}{2} - \log \cosh(\beta_n s) ,$$

$$G_{\rho_n}^{(1)}(0) = G_{\rho_n}^{(3)}(0) = 0 ,$$

$$G_{\rho_n}^{(2)}(0) = 1 - \beta_n^2 = \delta_n(2 - \delta_n) = \frac{2\alpha}{\sqrt{n}} + o(n^{-1/2}) .$$

This completes the verification of the conditions of the theorem, and we obtain that

$$\frac{S_n(\rho_n)}{n^{-1/4}} \to \exp\left(-\frac{s^4}{12} - \alpha \frac{s^2}{2}\right) . \tag{5.10}$$

Recalling that, if $x_i = \pm 1$ is the i-th spin,

$$\mu_n(x_1, \ldots, x_n) = \frac{1}{Z(\beta)} \exp\left(\frac{\beta}{n} \sum_{1 \leq i < j \leq n} x_i x_j\right), \tag{5.11}$$

clearly $S_n(\rho_n)$ has the same distribution as S_{μ_n} for any n. This completes the proof of the theorem. ∎

Remark. One can verify that the above analysis of the mixing time in the critical window holds also for the censored dynamics (where the magnetization is restricted to be non-negative, by flipping all spins whenever it becomes negative). Indeed, the upper bound immediately holds as the censored dynamics is a function of the original Glauber dynamics. For the lower bound, notice that our argument tracked the absolute value of the magnetization chain, and hence can readily be applied to the censored case as-well. Altogether, the censored dynamics has a mixing time of order $n^{3/2}$ in the critical window $1 \pm \delta$ for $\delta = O(1/\sqrt{n})$.

5.3. **Spectral gap analysis.** The spectral gap bound in the critical temperature regime is obtained by combining the above analysis with results of [5] on birth-and-death chains.

The lower bound on `gap` is a direct consequence of the fact that the mixing time has order $n^{3/2}$, and that the inequality $t_{\text{REL}} \le t_{\text{MIX}}\left(\frac{1}{4}\right)$ always holds. It remains to prove the matching bound $t_{\text{MIX}}\left(\frac{1}{4}\right) = O(t_{\text{REL}})$. Suppose that this is false, that is, $t_{\text{REL}} = o\left(t_{\text{MIX}}\left(\frac{1}{4}\right)\right)$.

Let A be some large constant, and let $s_0 = An^{-1/4}$. Notice that the case $\delta^2 n = O(1)$ in Theorem 4.3 implies that $\mathbf{E}_1 \tau_0 = O(n^{3/2})$. Furthermore, by Theorem 5.2, there exists a strictly positive function of A, $\varepsilon(A)$, such that $\lim_{A\to\infty} \varepsilon(A) = 0$ and

$$\frac{1}{2}\varepsilon(A) \le \pi(S \ge s_0) \le 2\varepsilon(A)$$

for sufficiently large n. Applying Lemma 3.7 with $\alpha = \pi(S \ge s_0)$ and $\beta = \frac{1}{2}$ gives $\mathbf{E}_{s_0}\tau_0 = o(n^{3/2})$. As in Subsection 5.2, set $\bar{t} = n^{3/2}/A^4$ for some large constant A. Combining Lemma 3.8 with Markov's inequality gives the following total variation bound for this birth-and-death chain:

$$\|\mathbf{P}_{s_0}(S_{\bar{t}} \in \cdot) - \pi\|_{\text{TV}} \le 4\varepsilon(A) + o(1) . \tag{5.12}$$

However, the lower bound (5.7) obtained in Subsection 5.2 implies that:

$$\|\mathbf{P}_{s_0}(S_{\bar{t}} \in \cdot) - \pi\|_{\text{TV}} \ge 1 - 4\varepsilon(\sqrt{A}/2) - 64/A^5 . \tag{5.13}$$

Choosing a sufficiently large constant A, (5.12) and (5.13) together lead to a contradiction for large n. We conclude that `gap` $= O(n^{-3/2})$, completing the proof.

Note that, as the condition `gap` $\cdot\, t_{\text{MIX}}(\frac{1}{4}) \to \infty$ is necessary for cutoff in any family of ergodic reversible finite Markov chains (see, e.g., [5]), we immediately deduce that there is no cutoff in this regime.

Remark. It is worth noting that the order of the spectral gap at $\beta_c = 1$ follows from a simpler argument. Indeed, in that case, the *upper bound* on `gap` can alternatively be derived from its Dirichlet representation, similar to the argument that appeared in the proof of Proposition 3.9 (where we substitute the identity function, i.e., the sum of spins in the Dirichlet form). For this argument, one needs a lower bound for the variance of the stationary magnetization. Such a bound is known for $\beta_c = 1$ (see [9]), rather than throughout the critical window.

6. Low temperature regime

In this section we prove Theorem 3, which establishes the order of the mixing time and the spectral gap in the super critical regime (where the mixing of the dynamics is exponentially slow and there is no cutoff).

6.1. **Exponential mixing.**

Recall that the normalized magnetization chain S_t is a birth-and-death chain on the space $\mathcal{X} = \{-1, -1+\frac{2}{n}, \ldots, 1-\frac{2}{n}, 1\}$, and for simplicity, assume throughout the proof that n is even (this is convenient since in this case we can refer to the 0 state. Whenever n is odd, the same proof holds by letting $\frac{1}{n}$ take the role of the 0 state).

The following notation will be useful. We define

$$\mathcal{X}[a,b] := \{x \in \mathcal{X} : a \leq x \leq b\} ,$$

and similarly define $\mathcal{X}(a,b)$, etc. accordingly. For all $x \in \mathcal{X}$, let p_x, q_x, h_x denote the transition probabilities of S_t to the right, to the left and to itself from the state x, that is:

$$p_x := P_M\left(x, x + \tfrac{2}{n}\right) = \frac{1-x}{2} \cdot \frac{1 + \tanh(\beta(x + n^{-1}))}{2} ,$$

$$q_x := P_M\left(x, x - \tfrac{2}{n}\right) = \frac{1+x}{2} \cdot \frac{1 - \tanh(\beta(x - n^{-1}))}{2} ,$$

$$h_x := P_M\left(x, x\right) = 1 - p_x - q_x .$$

By well known results on birth-and-death chains (see, e.g., [12]), the resistance r_x and conductance c_x of the edge $(x, x + 2/n)$, and the conductance c'_x of the self-loop of vertex x for $x \in \mathcal{X}[0,1]$ are (the negative parts can be obtained immediately by symmetry)

$$r_x = \prod_{y \in \mathcal{X}(0,x]} \frac{q_y}{p_y} , \; c_x = \prod_{y \in \mathcal{X}(0,x]} \frac{p_y}{q_y} , \; c'_x = \frac{h_x}{p_x + q_x}(c_{x-2/n} + c_x) , \quad (6.1)$$

and the commute-time between x and y, $C_{x,y}$ for $x < y$ (the minimal time it takes the chain, starting from x, to hit y then return to x) satisfies

$$\mathbf{E}C_{x,y} = 2c_S R(x \leftrightarrow y) , \quad (6.2)$$

where

$$c_S := \sum_{x \in \mathcal{X}} (c_x + c'_x) \quad \text{and} \quad R(x \leftrightarrow y) := \sum_{z \in \mathcal{X}[x,y)} r_z \ .$$

Our first goal is to estimate the expected commute time between 0 and ζ. This is incorporated in the next lemma.

Lemma 6.1. *The expected commute time between 0 and ζ has order*

$$t_{\exp} := \frac{n}{\delta} \exp\left(\frac{n}{2} \int_0^\zeta \log \frac{1 + g(x)}{1 - g(x)}\right) dx \ , \tag{6.3}$$

where $g(x) := (\tanh(\beta x) - x) / (1 - x \tanh(\beta x))$. In particular, in the special case $\delta \to 0$ we have $\mathbf{E}C_{0,\zeta} = \frac{n}{\delta} \exp\left((\frac{3}{4} + o(1))\delta^2 n\right)$, where the $o(1)$-term tends to 0 as $n \to \infty$.

Remark. If $\zeta \notin \mathcal{X}$, instead we simply choose a state in $\mathcal{X}$ which is the nearest possible to ζ. For a sufficiently large n, such a negligible adjustment would keep our calculations and arguments in tact. For the convenience of notation, let ζ denote the mentioned state in this case as well.

To prove Lemma 6.1, we need the following two lemmas, which establish the order of the total conductance and effective resistance respectively.

Lemma 6.2. *The total conductance satisfies*

$$c_S = \Theta\left(\sqrt{\frac{n}{\delta}} \exp\left(\frac{n}{2} \int_0^\zeta \log\left(\frac{1 + g(x)}{1 - g(x)}\right) dx\right)\right) \ .$$

Lemma 6.3. *The effective resistance between 0 and ζ satisfies*

$$R(0 \leftrightarrow \zeta) = \Theta(\sqrt{n/\delta}) \ .$$

Proof of Lemma 6.2. Notice that for any $x \in \mathcal{X}$, the holding probability h_x is uniformly bounded from below, and thus c'_x can be uniformly bounded from above by $(c_x + c_{x-2/n})$. It therefore follows that $c_S = \Theta(\tilde{c}_S)$ where $\tilde{c}_S := \sum_{x \in \mathcal{X}} c_x$, and it remains to determine $\tilde{c}_S$. We first locate the maximal edge conductance and determine its order, by means of classical analysis.

$$\log c_x = \sum_{y \in \mathcal{X}(0,x]} \log \frac{p_y}{q_y} = \sum_{y \in \mathcal{X}(0,x]} \log\left(\frac{1 - y}{1 + y} \cdot \frac{1 + \tanh(\beta(y + n^{-1}))}{1 - \tanh(\beta(y - n^{-1}))}\right)$$

$$= \sum_{y \in \mathcal{X}(0,x]} \log\left(\frac{1 + g(y)}{1 - g(y)} + O(1/n)\right) = \sum_{y \in \mathcal{X}(0,x]} \log\left(\frac{1 + g(y)}{1 - g(y)}\right) + O(x) \tag{6.4}$$

Note that $g(x)$ has a unique positive root at $x = \zeta$, and satisfies $g(x) > 0$ for $x \in (0, \zeta)$ and $g(x) < 0$ for $x > \zeta$. Therefore,

$$\log c_x \le \log c_\zeta + O(x) \le \log c_\zeta + O(1) \ ,$$

thus we move to estimate c_ζ. As $\log c_\zeta$ is simply a Riemann sum (after an appropriate rescaling), we deduce that

$$\log c_\zeta = \sum_{x \in \mathcal{X}(0,\zeta]} \log\left(\frac{1+g(x)}{1-g(x)}\right) + O(1) = \frac{n}{2}\int_0^\zeta \log\left(\frac{1+g(x)}{1-g(x)}\right) dx + O(1) \ ,$$

and therefore

$$c_\zeta = \Theta\left(\exp\left(\frac{n}{2}\int_0^\zeta \log\left(\frac{1+g(x)}{1-g(x)}\right) dx\right)\right) \ , \qquad (6.5)$$

$$c_x = O(c_\zeta) \ . \qquad (6.6)$$

Next, consider the ratio $c_{x+2/n}/c_x$; whenever $x \leq \zeta$, $g(x) \geq 0$, hence we have

$$\frac{c_{x+2/n}}{c_x} = \frac{p_{x+2/n}}{q_{x+2/n}} \geq \frac{1+g(x)}{1-g(x)} - O(1/n) \geq 1 + 2g(x) - O(1/n) \ .$$

Whenever $\frac{1}{\sqrt{\delta n}} \leq x \leq \zeta - \frac{1}{\sqrt{\delta n}}$ (using the Taylor expansions around 0 and around ζ) we obtain that $\tanh(\beta x) - x \geq \frac{1}{2}\sqrt{\delta/n}$. Combining this with the fact that $x\tanh(\beta x)$ is always non-negative, we obtain that for any such x, $2g(x) \geq \sqrt{\delta/n}$. Therefore, setting

$$\xi_1 := \sqrt{\frac{1}{\delta n}} \ , \ \ \xi_2 := \zeta - \sqrt{\frac{1}{\delta n}} \ , \ \ \xi_3 := \zeta + \sqrt{\frac{1}{\delta n}} \ , \qquad (6.7)$$

we get

$$\frac{c_{x+2/n}}{c_x} \geq 1 + \sqrt{\frac{\delta}{n}} - O(1/n) \ \ \text{for any } x \in \mathcal{X}[\xi_1, \xi_2] \ . \qquad (6.8)$$

Using the fact that $\delta^2 n \to \infty$, the sum of c_x-s in the above range is at most the sum of a geometric series with a quotient $1/(1 + \frac{1}{2}\sqrt{\delta/n})$ and an initial position c_ζ:

$$\sum_{x \in \mathcal{X}[\xi_1, \xi_2]} c_x \leq 3\sqrt{\frac{n}{\delta}} \cdot c_\zeta \ . \qquad (6.9)$$

We now treat $x \geq \xi_3$; since $g(\zeta) = 0$ and $g(x)$ is decreasing for any $x \geq \zeta$, then in particular whenever $\zeta + \sqrt{\delta/n} \leq x \leq 1$ we have $-1 = g(1) \leq g(x) \leq 0$, and therefore

$$\frac{c_{x+2/n}}{c_x} = \frac{p_{x+2/n}}{q_{x+2/n}} \leq 1 + g(x) + O(1/n) \ .$$

Furthermore, for any $\zeta + \sqrt{\delta/n} \leq x \leq 1$ (using Taylor expansion around ζ) we have $\tanh(\beta x) - x \leq -\sqrt{\delta/n}$, and hence $g(x) \leq -\sqrt{\delta/n}$. We deduce that

$$\frac{c_{x+2/n}}{c_x} \leq 1 - \sqrt{\frac{\delta}{n}} + O(1/n) \ \ \text{for any } x \in \mathcal{X}[\xi_3, 1] \ ,$$

and therefore

$$\sum_{x \in \mathcal{X}[\xi_3, 1]} c_x \leq 2\sqrt{\frac{n}{\delta}} \cdot c_\zeta \ . \tag{6.10}$$

Combining (6.9) and (6.10) together, and recalling (6.6), we obtain that

$$\tilde{c}_S = \sum_{x \in \mathcal{X}} c_x \leq 2\left(|\mathcal{X}[0, \xi_1]| + 5\sqrt{n/\delta} + |\mathcal{X}[\xi_2, \xi_3]|\right) c_\zeta = O\left(\sqrt{\frac{n}{\delta}} c_\zeta\right) \ .$$

Finally, consider $x \in \mathcal{X}[\xi_2, \xi_3]$; an argument similar to the ones above (i.e., perform Taylor expansion around ζ and bound the ratio of $c_{x+2/n}/c_x$) shows that c_x is of order c_ζ in this region. This implies that for some constant $b > 0$

$$\tilde{c}_S \geq \sum_{x \in \mathcal{X}[\xi_2, \xi_2]} c_x \geq b|\mathcal{X}[\xi_2, \xi_3]|c_\zeta \geq b\sqrt{\frac{n}{\delta}} c_\zeta \ , \tag{6.11}$$

and altogether, plugging in (6.5), we get

$$\tilde{c}_S = \Theta\left(\sqrt{\frac{n}{\delta}} \exp\left(\frac{n}{2} \int_0^\zeta \log\left(\frac{1 + g(x)}{1 - g(x)}\right) dx\right)\right) \ . \tag{6.12}$$

∎

Proof of Lemma 6.3. Translating the conductances, as given in (6.8), to resistances, we get

$$\frac{r_{x+2/n}}{r_x} \leq 1 - \sqrt{\frac{\delta}{n}} - O(1/n) \ \text{ for any } x \in \mathcal{X}[\xi_1, \xi_2] \ ,$$

and hence

$$\sum_{x \in \mathcal{X}[\xi_1, \xi_2]} r_x \leq r_{\xi_1} 2\sqrt{n/\delta} \leq 2\sqrt{n/\delta} \ ,$$

where in the last inequality we used the fact that $r_x \leq r_{x-2/n} (\leq r_0 = 1)$ for all $x \in \mathcal{X}[0, \zeta]$, which holds since $q_x \leq p_x$ for such x. Altogether, we have the following upper bound:

$$R(0 \leftrightarrow \zeta) = \sum_{x \in \mathcal{X}[0, \xi_1]} r_x + \sum_{x \in \mathcal{X}[\xi_1, \xi_2]} r_x + \sum_{x \in \mathcal{X}[\xi_2, \zeta]} r_x$$

$$\leq |\mathcal{X}[0, \xi_1]| + 2\sqrt{\frac{n}{\delta}} + |\mathcal{X}[\xi_2, \zeta]| \leq 4\sqrt{n/\delta} \ . \tag{6.13}$$

For a lower bound, consider $x \in \mathcal{X}[0, \xi_1]$. Clearly, for any $x \leq \frac{1}{\sqrt{\delta n}}$ we have $g(x) = \frac{\tanh(\beta x - x)}{1 - x \tanh(\beta x)} \leq 2\delta x$, and hence

$$\frac{r_{x+2/n}}{r_x} = \frac{1 - g(x)}{1 + g(x)} + O(1/n) \geq 1 - 5x\delta \geq \exp(-6x\delta) \ ,$$

yielding that

$$r_{\xi_1} \geq \exp\left(-3\delta n \cdot \xi_1^2\right) \geq e^{-3} \ .$$

Altogether,
$$R(0 \leftrightarrow \zeta) \geq e^{-3}|\mathcal{X}[0,\xi_1]| \geq e^{-4}\sqrt{n/\delta} \;,$$
and combining this with (6.13) we deduce that $R(0 \leftrightarrow \zeta) = \Theta(\sqrt{n/\delta})$. ∎

Proof of Lemma 6.1. Plugging in the estimates for $\tilde{c}_S$ and $R(0 \leftrightarrow \zeta)$ in (6.2), we get

$$\mathbf{E}C_{0,\zeta} = \Theta\left(\frac{n}{\delta}\exp\left(\frac{n}{2}\int_0^\zeta \log\left(\frac{1+g(x)}{1-g(x)}\right) dx\right)\right) \;. \tag{6.14}$$

This completes the proof of the lemma 6.1. ∎

Note that by symmetry, the expected hitting time from ζ to $-\zeta$ is exactly the expected commute time between 0 and ζ. Hence,

$$\mathbf{E}_\zeta[\tau_{-\zeta}] = \Theta(t_{\exp}) \;. \tag{6.15}$$

In order to show that the above hitting time is the leading order term in the mixing-time at low temperatures, we need the following lemma, which addresses the order of the hitting time from 1 to ζ.

Lemma 6.4. *The normalized magnetization chain S_t in the low temperature regimes satisfies $\mathbf{E}_1\tau_\zeta = o(t_{\exp})$.*

Proof. First consider the case where δ is bounded below by some constant. Notice that, as $p_x \leq q_x$ for all $x \geq \zeta$, in this region S_t is a supermartingale. Therefore, Lemma 3.5 (or simply standard results on the simple random walk, which dominates our chain in this case) implies that $\mathbf{E}_1\tau_\zeta = O(n^2)$. Combining this with the fact that $t_{\exp} \geq \exp(cn)$ for some constant c in this case, we immediately obtain that $\mathbf{E}_1\tau_\zeta = o(t_{\exp})$.

Next, assume that $\delta = o(1)$. Note that in this case, the Taylor expansion $\tanh(\beta x) = \beta x - \frac{1}{3}(\beta x)^3 + O((\beta x)^5)$ implies that

$$\zeta = \sqrt{3\delta/\beta^3 - O((\beta\zeta))^5} = \sqrt{3\delta} + O(\delta^{3/2}) \;. \tag{6.16}$$

Recalling that $\mathbf{E}[S_{t+1} \mid S_t = s] \leq s + \frac{1}{n}(\tanh(\beta s) - s)$ (as $s \geq 0$), Jensen's inequality (using the concavity of the Hyperbolic tangent) gives

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

$$\leq \frac{1}{n}\left(\tanh(\beta\mathbf{E}S_t) - \mathbf{E}S_t\right) \;. \tag{6.17}$$

Further note that the function $\tanh(\beta s)$ has the following Taylor expansion around ζ (for some ξ between s and ζ):

$$\tanh(\beta s) = \zeta + \beta(1-\zeta^2)(s-\zeta) + \beta^2(-1+\zeta^2)\zeta(s-\zeta)^2$$

$$+ \frac{\beta^3}{3}(-1+4\zeta^2-\zeta^4)(s-\zeta)^3 + \frac{\tanh^{(4)}(\xi)}{4!}(s-\zeta)^4 \;. \tag{6.18}$$

Since $\tanh^{(4)}(x) < 5$ for any $x \geq 0$, (6.18) implies that for a sufficiently large n the term $-\frac{1}{3}(s - \zeta)^3$ absorbs the last term in the expansion (6.18). Together with (6.16), we obtain that

$$\tanh(\beta s) \leq \zeta + \beta(1 - \zeta^2)(s - \zeta) + \beta^2(-1 + \zeta^2)\sqrt{\delta}(s - \zeta)^2 \ .$$

Therefore, (6.17) follows:

$$\mathbf{E}[S_{t+1} - S_t] \leq -\frac{\sqrt{\delta}}{2n}(\mathbf{E}S_t - \zeta)^2 \ . \tag{6.19}$$

Set

$$b_i = 2^{-i} \ , i_2 = \min\{i : b_i < \sqrt{\delta}\} \text{ and } u_i = \min\{t : \mathbf{E}S_t - \zeta < b_i\} \ ,$$

noting that this gives $b_i/2 \leq \mathbf{E}S_t - \zeta \leq b_i$ for any $t \in [u_i, u_{i+1}]$. It follows that

$$u_{i+1} - u_i \leq \frac{b_i/2}{\frac{\sqrt{\delta}}{2n}(\frac{b_i}{2})^2} = \frac{4n}{\sqrt{\delta}b_i} \ ,$$

and hence

$$\sum_{i=1}^{i_2} u_{i+1} - u_i \leq \sum_{i:b_i^2 > \delta} \frac{4n}{\sqrt{\delta}b_i} = O(n/\delta) \ ,$$

where we used the fact that the series $\{b_i^{-1}\}$ is geometric with ratio 2. We claim that this implies the required bound on $\mathbf{E}_1\tau_\zeta$. To see this, recall (6.19), according to which $W_t := n(S_t - \zeta)\mathbf{1}_{\{\tau_\zeta > t\}}$ is a supermartingale with bounded increments, whose variance is uniformly bounded from below on the event $\tau_\zeta > t$ (as the holding probabilities of (S_t) are uniformly bounded from above, see (3.4)). Moreover, the above argument gives $\mathbf{E}W_t \leq n\sqrt{\delta}$ for some $t = O(n/\delta)$. Thus, applying Lemma 3.5 and taking expectation, we deduce that $\mathbf{E}_1\tau_\zeta = O(n/\delta + \delta n^2) = O(\delta n^2)$, which in turns gives $\mathbf{E}_1\tau_\zeta = o(t_{\exp})$. ∎

Remark. With additional effort, we can establish that $\mathbf{E}_0\tau_{\pm\zeta} = o(t_{\exp})$ (for more details, see the companion paper [6]), where $\tau_{\pm\zeta} = \min\{t : |S_t| \geq \zeta\}$. By combining this with the of S_t symmetry and applying the geometric trial method, we can obtain the expected commute time between 0 and ζ:

$$\mathbf{E}_\zeta\tau_0 = (\tfrac{1}{2} + o(1))\mathbf{E}C_{0,\zeta} = \Theta(t_{\exp}) \ ,$$

and therefore conclude that $\mathbf{E}_1\tau_0 = \Theta(t_{\exp})$.

6.1.1. *Upper bound for mixing.* Combining Lemma 6.4 and (6.15), we conclude that $\mathbf{E}_1 \tau_{-\zeta} = \Theta(t_{\exp})$ and hence $\mathbf{E}_1 \tau_0 = O(t_{\exp})$. Together with Lemma 3.1, this implies that the magnetization chain will coalescence in $O(t_{\exp})$ steps with probability arbitrarily close to 1. At this point, Lemma 3.3 immediately gives that the Glauber dynamics achieves full mixing within $O(n \log n)$ additional steps. The following simple lemma thus completes the proof of the upper bound for the mixing time.

Lemma 6.5. *Let $t_{\exp}$ be as defined in Lemma 6.1. Then $n \log n = o(t_{\exp})$.*

Proof. In case $\delta \geq c > 0$ for some constant c, we have $t_{\exp} \geq n \exp(c'n)$ for some constant $c' > 0$ and hence $n \log n = o(t_{\exp})$. It remains to treat the case $\delta = o(1)$.

Suppose first that $\delta = o(1)$ and $\delta \geq cn^{-1/3}$ for some constant $c > 0$. In this case, we have $t_{\exp} = \frac{n}{\delta} \exp\left((\frac{3}{4} + o(1))\delta^2 n\right)$ and thus $n \exp(\frac{1}{2}n^{1/3}) = O(t_{\exp})$, giving $n \log n = o(t_{\exp})$. Finally, if $\delta = o(n^{-1/3})$, we can simply conclude that $n^{4/3} = O(t_{\exp})$ and hence $n \log n = o(t_{\exp})$. ∎

6.1.2. *Lower bound for mixing.* The lower bound will follow from showing that the probability of hitting $-\zeta$ within $\varepsilon t_{\exp}$ steps is small, for some small $\varepsilon > 0$ to be chosen later. To this end, we need the following simple lemma:

Lemma 6.6. *Let X denote a Markov chain over some finite state space Ω, $y \in \Omega$ denote a target state, and T be an integer. Further let $x \in \Omega$ denote the state with the smallest probability of hitting y after at most T steps, i.e., x minimizes $\mathbf{P}_x(\tau_y \leq T)$. The following holds:*

$$\mathbf{P}_x(\tau_y \leq T) \leq \frac{T}{\mathbf{E}_x \tau_y} \ .$$

Proof. Set $p = \mathbf{P}_x(\tau_y \leq T)$. By definition, $\mathbf{P}_z(\tau_y \leq T) \geq p$ for all $z \in \Omega$, hence the hitting time from x to y is stochastically dominated by a geometric random variable with success probability p, multiplied by T. That is, we have $\mathbf{E}_x \tau_y \leq T/p$, completing the proof. ∎

The final fact we would require is that the stationary probability of $\mathcal{X}[-1, -\zeta]$ is strictly positive. This is stated by the following lemma.

Lemma 6.7. *There exists some absolute constant $0 < C_\pi < 1$ such that*

$$C_\pi \leq \pi(\mathcal{X}[\zeta, 1]) \ (= \pi(\mathcal{X}[-1, -\zeta])) \ .$$

Proof. Repeating the derivation of (6.11), we can easily get

$$c_{\mathcal{X}[\zeta,1]} := \sum_{x \in \mathcal{X}[\zeta,1]} (c_x + c'_x) \geq \Theta\left(\sqrt{\frac{n}{\delta}} \exp\left(\frac{n}{2} \int_0^\zeta \log \frac{1 + g(x)}{1 - g(x)} \, dx\right)\right) \ .$$

Combining the above bound with (6.12), we conclude that there exists some $C_\pi > 0$, such that $\pi(\mathcal{X}[\zeta, 1]) \geq C_\pi$. ∎

Plugging in the target state $-\zeta$ into Lemma 6.6, and recalling that the monotone-coupling implies that, for any T, the initial state $s_0 = 1$ has the smallest probability (among all initial states) of hitting $-\zeta$ within T steps, we deduce that, for a sufficiently small $\varepsilon > 0$,

$$\mathbf{P}_1(\tau_{-\zeta} \leq \varepsilon t_{\exp}) \leq \frac{1}{2} C_\pi \ .$$

This implies that

$$t_{\exp} = O\Big(t_{\mathrm{MIX}}\Big(\frac{1}{2} C_\pi\Big)\Big) \ ,$$

which in turn gives

$$t_{\exp} = O\left(t_{\mathrm{MIX}}\big(\tfrac{1}{4}\big)\right) \ .$$

6.2. **Spectral gap analysis.** The lower bound is straightforward (as the relaxation time is always at most the mixing time) and we turn to prove the upper bound. Note that, by Lemma 6.7, we have $\pi(\mathcal{X}[\zeta, 1]) \geq C_\pi > 0$. Suppose first that $\mathtt{gap} \cdot t_{\mathrm{MIX}}(\frac{1}{4}) \to \infty$. In this case, one can apply Lemma 3.7 onto the birth-and-death chain (S_t), with a choice of $\alpha = \pi(\mathcal{X}[\zeta, 1])$ and $\beta = 1 - \pi(\mathcal{X}[\zeta, 1])$ (recall that $t_{\mathrm{MIX}}(\frac{1}{4}) = \Theta(\mathbf{E}_1 \tau_{-\zeta})$). It follows that

$$\mathbf{E}_\zeta \tau_{-\zeta} = o\left(\mathbf{E}_1 \tau_{-\zeta}\right) \ .$$

However, as both quantities above should have the same order as $t_{\mathrm{MIX}}(\frac{1}{4})$, this leads to a contradiction. We therefore have $\mathtt{gap} \cdot t_{\mathrm{MIX}}(\frac{1}{4}) = O(1)$, completing the proof of the upper bound.

References

[1] D. Aldous and J. A. Fill, *Reversible Markov Chains and Random Walks on Graphs.* In preparation, `http://www.stat.berkeley.edu/~aldous/RWG/book.html`.

[2] M. Aizenman and R. Holley, *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, 1987, pp. 1–11.

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

[4] M.-F. Chen, *Trilogy of couplings and general formulas for lower bound of spectral gap*, Probability towards 2000 (New York, 1995), Lecture Notes in Statist., vol. 128, Springer, New York, 1998, pp. 123–136 ↑157,230,231.

[5] J. Ding, E. Lubetzky, and Y. Peres, *Total-variation cutoff in birth-and-death chains*, preprint.

[6] J. Ding, E. Lubetzky, and Y. Peres, *Censored Glauber dynamics for the mean-field Ising Model*, preprint.

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

[8]   R. S. Ellis and C. M. Newman, *Limit theorems for sums of dependent random variables occurring in statistical mechanics*, Z. Wahrsch. Verw. Gebiete **44** (1978), no. 2, 117–139.

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

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

[11] D. A. Levin, M. Luczak, and Y. Peres, *Glauber dynamics for the Mean-field Ising Model: cut-off, critical power law, and metastability*, Probability Theory and Related Fields, to appear.

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

[13] Ş. Nacu, *Glauber Dynamics on the Cycle is Monotone*, Probability Theory and Related Fields **127** (2003), 177-185.

Jian Ding

Department of Statistics, UC Berkeley, Berkeley, CA 94720, USA.
  E-mail address: `jding@stat.berkeley.edu`

Eyal Lubetzky

Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399, USA.
  E-mail address: `eyal@microsoft.com`

Yuval Peres

Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399, USA.
  E-mail address: `peres@microsoft.com`