Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2011, Vol. 47, No. 4, 1020–1028
DOI: 10.1214/10-AIHP402

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Mixing time for the Ising model: A uniform lower bound for all graphs

Jian Ding[a] and Yuval Peres[b]

[a]*Department of Statistics, UC Berkeley, Berkeley, CA 94720, USA. E-mail: jding@stat.berkeley.edu*
[b]*Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399, USA. E-mail: peres@microsoft.com*

Received 16 November 2009; revised 3 November 2010; accepted 10 November 2010

Abstract. Consider Glauber dynamics for the Ising model on a graph of n vertices. Hayes and Sinclair showed that the mixing time for this dynamics is at least $n\log n/f(\Delta)$, where Δ is the maximum degree and $f(\Delta)=\Theta(\Delta\log^2\Delta)$. Their result applies to more general spin systems, and in that generality, they showed that some dependence on Δ is necessary. In this paper, we focus on the ferromagnetic Ising model and prove that the mixing time of Glauber dynamics on any n-vertex graph is at least $(1/4+\mathrm{o}(1))n\log n$.

Résumé. Dans cet article nous étudions la dynamique de Glauber du modèle d'Ising sur un graphe fini à n sommets. Hayes et Sinclair ont montré que le temps de mélange de cette dynamique est au moins de $n\log(n)f(\Delta)$, où Δ est le degré maximum d'un sommet du graphe et $f(\Delta)=\Theta(\Delta\log^2(\Delta))$. Leur résultat s'applique également à des modèles de spins généraux où la dépendance en Δ est nécessaire. Dans ce travail nous nous concentrons sur le modèle d'Ising ferromagnétique et montrons que le temps de mélange de la dynamique de Glauber est au moins de $(1/4+\mathrm{o}(1))n\log(n)$ sur n'importe quel graphe à n sommets.

MSC: Primary 60J10; secondary 60K35; 68W20

Keywords: Glauber dynamics; Mixing time; Ising model

1. Introduction

Consider a finite graph $G=(V,E)$ and a finite alphabet Q. A general *spin system* on G is a probability measure μ on Q^V; well studied examples in computer science and statistical physics include the uniform measure on proper colorings and the Ising model. Glauber (heat-bath) dynamics are often used to sample from μ (see, e.g., [9,11,15]). In discrete-time Glauber dynamics, at each step a vertex v is chosen uniformly at random and the label at v is replaced by a new label chosen from the μ-conditional distribution given the labels on the other vertices. This Markov chain has stationary distribution μ, and the key quantity to analyze is the mixing time t_{mix}, at which the distribution of the chain is close in total variation to μ (precise definitions are given below).

If $|V|=n$, it takes $(1+\mathrm{o}(1))n\log n$ steps to update all vertices (coupon collecting), and it is natural to guess that this is a lower bound for the mixing time. However, for the Ising model at infinite temperature or equivalently, for the 2-colorings of the graph $(V,\varnothing)$, the mixing time of Glauber dynamics is asymptotic to $n\log n/2$, since these models reduce to the lazy random walk on the hypercube, first analyzed in [1]. Thus mixing can occur before all sites are updated, so the coupon collecting argument does not suffice to obtain a lower bound for the mixing time. The first general bound of the right order was obtained by Hayes and Sinclair [6], who showed that the mixing time for Glauber dynamics is at least $n\log n/f(\Delta)$, where Δ is the maximum degree and $f(\Delta)=\Theta(\Delta\log^2\Delta)$. Their result applies for quite general spin systems, and they gave examples of spin systems μ where some dependence on Δ is necessary.

After the work of [6], it remained unclear whether a uniform lower bound of order $n\log n$, *that does not depend on* Δ, holds for the most extensively studied spin systems, such as proper colorings and the Ising model.

In this paper, we focus on the ferromagnetic Ising model, and obtain a lower bound of $(1/4+\mathrm{o}(1))n\log n$ on any graph with general (non-negative) interaction strengths.

Definitions. *The Ising model on a finite graph $G=(V,E)$ with interaction strengths $J=\{J_{uv}\geq 0: uv\in E\}$ is a probability measure μ_G on the configuration space $\Omega=\{\pm 1\}^V$, defined as follows. For each $\sigma\in\Omega$,*

$$\mu_G(\sigma)=\frac{1}{Z(J)}\exp\biggl(\sum_{uv\in E}J_{uv}\sigma(u)\sigma(v)\biggr), \tag{1.1}$$

where $Z(J)$ is a normalizing constant called the partition function. The measure μ_G is also called the Gibbs measure corresponding to the interaction matrix J. When there is no ambiguity regarding the base graph, we sometimes write μ for μ_G.

Recall the definition of the Glauber dynamics: At each step, a vertex is chosen uniformly at random, and its spin is updated according to the conditional Gibbs measure given the spins of all the other vertices. It is easy to verify that this chain is reversible with respect to μ_G.

Next we define the *mixing time*. Let (X_t) denote an aperiodic irreducible Markov chain on a finite state space Ω with transition kernel P and stationary measure π. For any two distributions μ,ν on Ω, their *total-variation distance* is defined to be

$$\|\mu-\nu\|_{\mathrm{TV}}\stackrel{\triangle}{=}\sup_{A\subset\Omega}\bigl|\mu(A)-\nu(A)\bigr|=\frac{1}{2}\sum_{x\in\Omega}\bigl|\mu(x)-\nu(x)\bigr|.$$

For $x\in\Omega$ let $\mathbb{P}_x$ denotes the probability given $X_0=x$ and let

$$t_{\mathrm{mix}}^x=\min\biggl\{t:\bigl\|\mathbb{P}_x(X_t\in\cdot)-\pi\bigr\|_{\mathrm{TV}}\leq\frac{1}{4}\biggr\}$$

be the mixing time with initial state x. (The choice of $1/4$ here is by convention, and can be replaced by any constant in $(0,1/2)$, without affecting the $(1/4+\mathrm{o}(1))n\log n$ lower bound in the next theorem.) The *mixing time* t_{mix} is then defined to be $\max_{x\in\Omega}t_{\mathrm{mix}}^x$.

We now state our main result.

Theorem 1. *Consider the Ising model* (1.1) *on the graph G with interaction matrix J, and let $t_{\mathrm{mix}}^+(G,J)$ denote the mixing time of the corresponding Glauber dynamics, started from the all-plus configuration. Then*

$$\inf_{G,J}t_{\mathrm{mix}}^+(G,J)\geq\bigl(1/4+\mathrm{o}(1)\bigr)n\log n,$$

where the infimum is over all n-vertex graphs G and all non-negative interaction matrices J.

Remark. *Theorem* 1 *is sharp up to a factor of* 2. *We conjecture that* $(1/4+\mathrm{o}(1))$ *in the theorem could be replaced by* $(1/2+\mathrm{o}(1))$, *i.e., the mixing time is minimized* (*at least asymptotically*) *by taking* $J\equiv 0$.

Hayes and Sinclair [6] constructed spin systems where the mixing time of the Glauber dynamics has an upper bound $\mathrm{O}(n\log n/\log\Delta)$. This, in turn, implies that in order to establish a lower bound of order $n\log n$ for the Ising model on a general graph, we have to employ some specific properties of the model. In our proof of Theorem 1, given in the next section, we use the GHS inequality [5] (see also [7] and [3]) and a recent censoring inequality [14] due to Peter Winkler and the second author.

2. Proof of Theorem 1

The intuition for the proof is the following: In the case of strong interactions, the spins are highly correlated and the mixing should be quite slow; In the case of weak interaction strengths, the spins should be weakly dependent and close to the case of the graph with no edges, therefore one may extend the arguments for the lazy walk on the hypercube.

We separate the two cases by considering the *spectral gap*. Recall that the spectral gap of a reversible discrete-time Markov chain, denoted by $\mathtt{gap}$, is $1-\lambda$, where λ is the second largest eigenvalue of the transition kernel. The following simple lemma gives a lower bound on t_{mix}^{+} in terms of the spectral gap.

Lemma 2.1. *The Glauber dynamics for the ferromagnetic Ising model* (1.1) *satisfies* $t_{\mathrm{mix}}^{+} \ge \log 2 \cdot (\mathtt{gap}^{-1}-1)$.

Proof. It is well known that $t_{\mathrm{mix}} \ge \log 2 \cdot (\mathtt{gap}^{-1}-1)$ (see, e.g., Theorem 12.4 in [9]). Actually, it is shown in the proof of [9], Theorem 12.4, that $t_{\mathrm{mix}}^{x} \ge \log 2 \cdot (\mathtt{gap}^{-1}-1)$ for any state x satisfying $f(x) = \|f\|_\infty$, where f is an eigenfunction corresponding to the second largest eigenvalue. Since the second eigenvalue of the Glauber dynamics for the ferromagnetic Ising model has an increasing eigenfunction f (see [12], Lemma 3), we infer that either $\|f\|_\infty = f(+)$ or $\|f\|_\infty = f(-)$. By symmetry of the all-plus and the all-minus configurations in the Ising model (1.1), we have $t_{\mathrm{mix}}^{+} = t_{\mathrm{mix}}^{-}$, and this concludes the proof. □

Lemma 2.1 implies that Theorem 1 holds if $\mathtt{gap}^{-1} \ge n\log n$. It remains to consider the case $\mathtt{gap}^{-1} \le n \log n$.

Lemma 2.2. *Suppose that the Glauber dynamics for the Ising model on a graph $G=(V,E)$ with n vertices satisfies $\mathtt{gap}^{-1} \le n\log n$. Then there exists a subset $F \subset V$ of size $\lfloor \sqrt{n}/\log n \rfloor$ such that*

$$\sum_{u,v\in F, u\neq v} \mathrm{Cov}_\mu\big(\sigma(u),\sigma(v)\big) \le \frac{2}{\log n}.$$

Proof. We first establish an upper bound on the variance of the sum of spins $S = S(\sigma) = \sum_{v\in V}\sigma(v)$. The variational principle for the spectral gap of a reversible Markov chain with stationary measure π gives (see, e.g., [2], Chapter 3, or [9], Lemma 13.12):

$$\mathtt{gap} = \inf_f \frac{\mathcal{E}(f)}{\mathrm{Var}_\pi(f)},$$

where $\mathcal{E}(f)$ is the Dirichlet form defined by

$$\mathcal{E}(f) = \big\langle (I-P)f, f\big\rangle_\pi = \frac{1}{2}\sum_{x,y\in\Omega}\big[f(x)-f(y)\big]^2\pi(x)P(x,y).$$

Applying the variational principle with the test function S, we deduce that

$$\mathtt{gap} \le \frac{\mathcal{E}(S)}{\mathrm{Var}_\mu(S)}.$$

Since the Glauber dynamics updates a single spin at each step, $\mathcal{E}(S) \le 2$, whence

$$\mathrm{Var}_\mu(S) \le \mathcal{E}(S)\mathtt{gap}^{-1} \le 2n\log n. \tag{2.1}$$

The covariance of the spins for the ferromagnetic Ising model is non-negative by the FKG inequality (see, e.g., [4]). Applying Claim 2.3 below with $k = \lfloor \frac{\sqrt{n}}{\log n} \rfloor$ to the covariance matrix of σ concludes the proof of the lemma. □

Claim 2.3. *Let A be an $n \times n$ matrix with non-negative entries. Then for any $k \le n$ there exists $F \subset \{1, \dots, n\}$ such that $|F| = k$ and*

$$\sum_{i,j \in F} A_{i,j} \mathbf{1}_{\{i \neq j\}} \le \frac{k^2}{n^2} \sum_{i \neq j} A_{i,j}.$$

Proof. Let R be a uniform random subset of $\{1, \dots, n\}$ with $|R| = k$. Then,

$$\mathbb{E}\left[\sum_{i,j \in R} A_{i,j} \mathbf{1}_{\{i \neq j\}}\right] = \sum_{1 \le i,j \le n} A_{i,j} \mathbf{1}_{\{i \neq j\}} \mathbb{P}(i, j \in R) = \frac{k(k-1)}{n(n-1)} \sum_{1 \le i,j \le n} A_{i,j} \mathbf{1}_{\{i \neq j\}} \le \frac{k^2}{n^2} \sum_{i \neq j} A_{i,j}.$$

Existence of the desired subset F follows immediately. □

We now consider a version of accelerated dynamics (X_t) with respect to the subset F as in Lemma 2.2. The accelerated dynamics selects a vertex $v \in V$ uniformly at random at each time and updates in the following way:

- If $v \notin F$, we update $\sigma(v)$ as in the usual Glauber dynamics.
- If $v \in F$, we update the spins on $\{v\} \cup F^c$ all together as a block, according to the conditional Gibbs measure given the spins on $F \setminus \{v\}$.

The next censoring inequality for monotone systems of [14] guarantees that, starting from the all-plus configuration, the accelerated dynamics indeed mixes faster than the original one. A *monotone system* is a Markov chain on a partially ordered set with the property that for any pair of states $x \le y$ there exist random variables $X_1 \le Y_1$ such that for every state z

$$\mathbb{P}(X_1 = z) = p(x, z), \qquad \mathbb{P}(Y_1 = z) = p(y, z).$$

In what follows, write $\mu \preceq \nu$ if ν stochastically dominates μ.

Theorem 2.4 ([14] and also see [13], Theorem 16.5). *Let (Ω, S, V, π) be a monotone system and let μ be the distribution on Ω which results from successive updates at sites $v_1, \dots, v_m$, beginning at the top configuration. Define ν similarly but with updates only at a subsequence $v_{i_1}, \dots, v_{i_k}$. Then $\mu \preceq \nu$, and $\|\mu - \pi\|_{\mathrm{TV}} \le \|\nu - \pi\|_{\mathrm{TV}}$. Moreover, this also holds if the sequence $v_1, \dots, v_m$ and the subsequence $i_1, \dots, i_k$ are chosen at random according to any prescribed distribution.*

In order to see how the above theorem indeed implies that the accelerated dynamics (X_t) mixes at least as fast as the usual dynamics, first note that any vertex $u \notin F$ is updated according to the original rule of the Glauber dynamics. Second, for $u \in F$, instead of updating the block $\{u\} \cup F^c$, we can simulate this procedure by performing sufficiently many single-site updates in $\{u\} \cup F^c$. This approximates the accelerated dynamics arbitrarily well, and contains a superset of the single-site updates of the usual Glauber dynamics. In other words, the single-site Glauber dynamics can be considered as a "censored" version of our accelerated dynamics. Theorem 2.4 thus completes this argument.

Let (Y_t) be the projection of the chain (X_t) onto the subgraph F. Recalling the definition of the accelerated dynamics, we see that (Y_t) is also a Markov chain, and the stationary measure ν_F for (Y_t) is the projection of μ_G to F. Furthermore, consider the subsequence (Z_t) of the chain (Y_t) obtained by skipping those times when updates occurred outside of F in (X_t). Namely, let $Z_t = Y_{K_t}$ where K_t is the tth time that a block $\{v\} \cup F^c$ is updated in the chain (X_t). Clearly, (Z_t) is a Markov chain on the space $\{-1, 1\}^F$, where at each time a uniform vertex v from F is selected and updated according to the conditional Gibbs measure μ_G given the spins on $F \setminus \{v\}$. The stationary measure for (Z_t) is also ν_F.

Let $\mathcal{S}_t = \sum_{v \in F} Z_t(v)$ be the sum of spins over F in the chain (Z_t). It turns out that $\mathcal{S}_t$ is a distinguishing statistic and its analysis yields a lower bound on the mixing time for chain (Z_t). To this end, we need to estimate the first two moments of S_t.

Lemma 2.5. *Started from all-plus configuration, the sum of spins satisfies that*

$$\mathbb{E}_+(\mathcal{S}_t) \ge |F|\left(1-\frac{1}{|F|}\right)^t.$$

Proof. The proof follows essentially from a coupon collecting argument. Let $(Z_t^{(+)})$ be an instance of the chain (Z_t) started at the all-plus configuration, and let (Z_t^*) be another instance of the chain (Z_t) started from ν_F. It is obvious that we can construct a monotone coupling between $(Z_t^{(+)})$ and (Z_t^*) (namely, $Z_t^{(+)} \ge Z_t^*$ for all $t \in \mathbb{N}$) such that the vertices selected for updating in both chains are always the same. Denote by $U[t]$ this (random) sequence of vertices updated up to time t. Note that Z_t^* has law ν_F, even if conditioned on the sequence $U[t]$. Recalling that $Z_t^{(+)} \ge Z_t^*$ and $\mathbb{E}_\mu\sigma(v) = 0$, we obtain that

$$\mathbb{E}_+\left[Z_t^{(+)}(v) \mid v \in U[t]\right] \ge 0.$$

It is clear that $Z_t^{(+)}(v) = 1$ if $v \notin U[t]$. Therefore,

$$\mathbb{E}_+\left[Z_t^{(+)}(v)\right] \ge \mathbb{P}\left(v \notin U[t]\right) = \left(1-\frac{1}{|F|}\right)^t.$$

Summing over $v \in F$ concludes the proof. □

We next establish a contraction result for the chain (Z_t). We need the GHS inequality of [5] (see also [7] and [3]). To state this inequality, we recall the definition of the Ising model *with an external field*. Given a finite graph $G = (V, E)$ with interaction strengths $J = \{J_{uv} \ge 0 : uv \in E\}$ and external magnetic field $H = \{H_v : v \in V\}$, the probability for a configuration $\sigma \in \Omega = \{\pm 1\}^V$ is given by

$$\mu_G^H(\sigma) = \frac{1}{Z(J,H)} \exp\left(\sum_{uv\in E} J_{uv}\sigma(u)\sigma(v) + \sum_{v\in V} H(v)\sigma(v)\right), \tag{2.2}$$

where $Z(J, H)$ is a normalizing constant. Note that this specializes to (1.1) if $H \equiv 0$. When there is no ambiguity for the base graph, we sometimes drop the subscript G. We can now state the following inequality.

GHS inequality [5]. For a graph $G = (V, E)$, let $\mu^H = \mu_G^H$ as above, and denote by $m_v(H) = \mathbb{E}_{\mu^H}[\sigma(v)]$ the local magnetization at vertex v. If $H_v \ge 0$ for all $v \in V$, then for any three vertices $u, v, w \in V$ (not necessarily distinct),

$$\frac{\partial^2 m_v(H)}{\partial H_u\, \partial H_w} \le 0.$$

The following is a consequence of the GHS inequality.

Corollary 2.6. *For the Ising measure μ with no external field, we have*

$$\mathbb{E}_\mu\left[\sigma(u) \mid v_i = 1 \textit{ for all } 1 \le i \le k\right] \le \sum_{i=1}^k \mathbb{E}_\mu\left[\sigma(u) | v_i = 1\right].$$

Proof. The function $f(H) = m_u(H)$ satisfies $f(0) = 0$. By the GHS inequality and Claim 2.7 below, we obtain that for all $H, H' \in \mathbb{R}_+^n$:

$$m_u\left(H + H'\right) \le m_u(H) + m_u\left(H'\right). \tag{2.3}$$

For $1 \le i \le k$ and $h \ge 0$, let H_i^h be the external field taking value h on v_i and vanishing on $V \setminus \{v_i\}$. Applying the inequality (2.3) inductively, we deduce that

$$m_u\left(\sum_i H_i^h\right) \le \sum_i m_u\left(H_i^h\right).$$

Finally, let $h \to \infty$ and observe that $m_u(H_i^h) \to \mathbb{E}_\mu[\sigma(u)|\sigma(v_i) = 1]$ and $m_u(\sum_i H_i^h) \to \mathbb{E}_\mu[\sigma(u)|\sigma(v_i) = 1$ for all $1 \le i \le k]$. □

Claim 2.7. *Write $\mathbb{R}_+ = [0, \infty)$ and let $f : \mathbb{R}_+^n \mapsto \mathbb{R}$ be a C^2-function such that $\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \le 0$ for all $x \in \mathbb{R}_+^n$ and $1 \le i, j \le n$. Then for all $x, y \in \mathbb{R}_+^n$,*

$$f(x + y) - f(x) \le f(y) - f(0).$$

Proof. Since all the second derivatives are non-positive, $\frac{\partial f(x)}{\partial x_i}$ is decreasing in every coordinate with x for all $x \in \mathbb{R}_+^n$ and $i \le n$. Hence, $\frac{\partial f(x)}{\partial x_i}$ is decreasing in $\mathbb{R}_+^n$. Let

$$g_x(t) = \frac{\mathrm{d}f(x + ty)}{\mathrm{d}t} = \sum_i y_i \frac{\partial f(x)}{\partial x_i}(x + ty).$$

It follows that $g_x(t) \le g_0(t)$ for all $x, y \in \mathbb{R}_+^n$. Integrating over $t \in [0, 1]$ yields the claim. □

Lemma 2.8. *Suppose that $n \ge \mathrm{e}^4$. Let $(\tilde{Z}_t)$ be another instance of the chain (Z_t). Then for all starting states z_0 and $\tilde{z}_0$, there exists a coupling such that*

$$\mathbb{E}_{z_0,\tilde{z}_0}\biggl[\sum_{v\in F}\bigl|Z_t(v) - \tilde{Z}_t(v)\bigr|\biggr] \le \biggl(1 - \frac{1}{2|F|}\biggr)^t \sum_{v\in F}\bigl|z_0(v) - \tilde{z}_0(v)\bigr|.$$

Proof. Fix $\eta, \tilde{\eta} \in \{-1, 1\}^F$ such that η and $\tilde{\eta}$ differ only at the vertex v and $\eta(v) = 1$. We consider two chains (Z_t) and $(\tilde{Z}_t)$ under monotone coupling, started from η and $\tilde{\eta}$, respectively. Let η_A be the restriction of η to A for $A \subset F$ (namely, $\eta_A \in \{-1, 1\}^A$ and $\eta_A(v) = \eta(v)$ for all $v \in A$), and write

$$\psi(u, \eta, \tilde{\eta}) = \mathbb{E}_\mu\bigl[\sigma(u)|\sigma_{F\setminus\{u\}} = \eta_{F\setminus\{u\}}\bigr] - \mathbb{E}_\mu\bigl[\sigma(u)|\sigma_{F\setminus\{u\}} = \tilde{\eta}_{F\setminus\{u\}}\bigr].$$

By the monotone property and symmetry of the Ising model,

$$\psi(u, \eta, \tilde{\eta}) \le \mathbb{E}_\mu\bigl[\sigma(u)|\sigma_{F\setminus\{u\}} = +\bigr] - \mathbb{E}_\mu\bigl[\sigma(u)|\sigma_{F\setminus\{u\}} = -\bigr] = 2\mathbb{E}_\mu\bigl[\sigma(u)|\sigma_{F\setminus\{u\}} = +\bigr].$$

By symmetry, we see that $\mathbb{E}(\sigma(u)|\sigma(w) = 1) = -\mathbb{E}(\sigma(u)|\sigma(w) = -1)$ and $\mathbb{E}(\sigma(u)) = 0$. Thus, $\mathrm{Cov}(\sigma(u), \sigma(w)) = \mathbb{E}(\sigma(u)|\sigma(w) = 1)$. Combined with Corollary 2.6, it yields that

$$\psi(u, \eta, \tilde{\eta}) \le 2 \sum_{w\in F\setminus\{u\}} \mathbb{E}_\mu\bigl[\sigma(u)|\sigma(w) = 1\bigr] = 2 \sum_{w\in F\setminus\{u\}} \mathrm{Cov}\bigl(\sigma(u), \sigma(w)\bigr).$$

Recalling the non-negative correlations between the spins, we deduce that under the monotone coupling

$$\begin{aligned}\mathbb{E}_{\eta,\tilde{\eta}}\biggl[\frac{1}{2}\sum_{v'\in F}\bigl|Z_1(v') - \tilde{Z}_1(v')\bigr|\biggr] &= 1 - \frac{1}{|F|} + \frac{1}{2|F|}\sum_{u\in F\setminus\{v\}} \psi(u, \eta, \tilde{\eta}) \\ &\le 1 - \frac{1}{|F|} + \frac{1}{|F|}\sum_{u\in F\setminus\{v\}}\sum_{w\in F\setminus\{u\}} \mathrm{Cov}\bigl(\sigma(u), \sigma(w)\bigr).\end{aligned}$$

By Lemma 2.2, we get that for $n \ge \mathrm{e}^4$,

$$\mathbb{E}_{\eta,\tilde{\eta}}\biggl[\frac{1}{2}\sum_{v'\in F}\bigl|Z_1(v') - \tilde{Z}_1(v')\bigr|\biggr] \le 1 - \frac{1}{|F|} + \frac{2}{|F|\log n} \le 1 - \frac{1}{2|F|}.$$

Using the triangle inequality and recursion, we conclude the proof. □

From the contraction result, we can derive the uniform variance bound on $\mathcal{S}_t$. This type of argument appeared in [8] (see Lemma 2.4) when (Z_t) is a one-dimensional chain. The argument naturally extends to multi-dimensional case and we include the proof for completeness.

Lemma 2.9. *Let (Z_t) and $(\tilde{Z}_t)$ be two instances of a Markov chain taking values in $\mathbb{R}^n$. Assume that for some $\rho < 1$ and all initial states z_0 and $\tilde{z}_0$, there exists a coupling satisfying*

$$\mathbb{E}_{z_0,\tilde{z}_0}\left[\sum_i \left|Z_t(i) - \tilde{Z}_t(i)\right|\right] \le \rho^t \sum_i \left|z_0(i) - \tilde{z}_0(i)\right|,$$

where we used the convention that $z(i)$ stands for the ith coordinate of z for $z \in \mathbb{R}^n$. Furthermore, suppose that $\sum_i |Z_t(i) - Z_{t-1}(i)| \le R$ for all t. Then for any $t \in \mathbb{N}$ and starting state $z \in \mathbb{R}^n$,

$$\mathrm{Var}_z\left(\sum_i Z_t(i)\right) \le \frac{2}{1-\rho^2} R^2.$$

Proof. Let Z_t and Z'_t be two independent instances of the chain both started from z. Defining $Q_t = \sum_i Z_t(i)$ and $Q'_t = \sum_i Z'_t(i)$, we obtain that

$$\begin{aligned}\left|\mathbb{E}_z[Q_t|Z_1 = z_1] - \mathbb{E}_z\left[Q'_t|Z'_1 = z'_1\right]\right| &= \left|\mathbb{E}_{z_1}[Q_{t-1}] - \mathbb{E}_{z'_1}\left[Q'_{t-1}\right]\right| \\ &\le \rho^{t-1} \sum_i \left|z_1(i) - z'_1(i)\right| \le 2\rho^{t-1} R\end{aligned}$$

for all possible choices of z_1 and z'_1. It follows that for any starting state z

$$\mathrm{Var}_z\left(\mathbb{E}_z[Q_t|Z_1]\right) = \frac{1}{2}\mathbb{E}_z\left[\left(\mathbb{E}_{Z_1}[Q_{t-1}] - \mathbb{E}_{Z'_1}\left[Q'_{t-1}\right]\right)^2\right] \le 2\left(\rho^{t-1} R\right)^2.$$

Therefore, by the total variance formula, we obtain that for all z

$$\mathrm{Var}_z(Q_t) = \mathrm{Var}_z\left(\mathbb{E}_z[Q_t|Z_1]\right) + \mathbb{E}_z\left[\mathrm{Var}_z(Q_t|Z_1)\right] \le 2\left(\rho^{t-1} R\right)^2 + \nu_{t-1},$$

where $\nu_t \stackrel{\triangle}{=} \max_z \mathrm{Var}_z(Q_t)$. Thus $\nu_t \le 2(\rho^{t-1}R)^2 + \nu_{t-1}$, whence

$$\nu_t \le \sum_{i=1}^t (\nu_i - \nu_{i-1}) \le \sum_{i=1}^t 2\rho^{2(t-1)} R^2 \le \frac{2R^2}{1-\rho^2},$$

completing the proof. □

Combining the above two lemmas gives the following variance bound (note that in our case $R = 2$ and $\rho = 1 - \frac{1}{2|F|}$, so $1 - \rho^2 \ge \frac{1}{2|F|}$).

Lemma 2.10. *For all t and starting position z, we have* $\mathrm{Var}_z(\mathcal{S}_t) \le 16|F|$.

We can now derive a lower bound on the mixing time for the chain (Z_t).

Lemma 2.11. *The chain (Z_t) has a mixing time $t^+_{\mathrm{mix}} \ge \frac{1}{2}|F|\log|F| - 20|F|$.*

Proof. Let $(Z_t^{(+)})$ be an instance of the dynamics (Z_t) started from the all-plus configuration and let $Z^* \in \{-1, 1\}^F$ be distributed as ν_F. Write

$$T_0 = \frac{1}{2}|F|\log|F| - 20|F|.$$

It suffices to prove that

$$d_{\mathrm{TV}}\big(\mathcal{S}_{T_0}^{(+)}, \mathcal{S}^*\big) \geq \frac{1}{4}, \tag{2.4}$$

where $\mathcal{S}_{T_0}^{(+)} = \sum_{v\in F} Z_{T_0}^{(+)}(v)$ as before and $\mathcal{S}^* = \sum_{v\in F} Z^*(v)$ be the sum of spins in stationary distribution. To this end, notice that by Lemmas 2.5 and 2.10:

$$\mathbb{E}_+\big(\mathcal{S}_{T_0}^{(+)}\big) \geq \mathrm{e}^{20+\mathrm{o}(1)}\sqrt{|F|} \quad \text{and} \quad \mathrm{Var}_+\big(\mathcal{S}_{T_0}^{(+)}\big) \leq 16|F|.$$

An application of Chebyshev's inequality gives that for large enough n

$$\mathbb{P}_+\big(\mathcal{S}_{T_0}^{(+)} \leq \mathrm{e}^{10}\sqrt{|F|}\big) \leq \frac{16|F|}{(\mathrm{e}^{20+\mathrm{o}(1)} - \mathrm{e}^{10})\sqrt{|F|})^2} \leq \frac{1}{4}. \tag{2.5}$$

On the other hand, it is clear by symmetry that $\mathbb{E}_{\nu_F}\mathcal{S}^* = 0$. Moreover, since Lemma 2.10 holds for all t, taking $t \to \infty$ gives that $\mathrm{Var}_{\nu_F}\mathcal{S}^* \leq 16|F|$. Applying Chebyshev's inequality again, we deduce that

$$\mathbb{P}_{\nu_F}\big(\mathcal{S}^* \geq \mathrm{e}^{10}\sqrt{|F|}\big) \leq \frac{16|F|}{(\mathrm{e}^{10}\sqrt{|F|})^2} \leq \frac{1}{4}.$$

Combining the above inequality with (2.5) and the fact that

$$d_{\mathrm{TV}}\big(\mathcal{S}_{T_0}^{(+)}, \mathcal{S}^*\big) \geq 1 - \mathbb{P}_+\big(\mathcal{S}_{T_0}^{(+)} \leq \mathrm{e}^{10}\sqrt{|F|}\big) - \mathbb{P}_\mu\big(\mathcal{S}^* \geq \mathrm{e}^{10}\sqrt{|F|}\big),$$

we conclude that (2.5) indeed holds (with room to spare), as required. □

We are now ready to derive Theorem 1. Observe that the dynamics (Y_t) is a lazy version of the dynamics (Z_t). Consider an instance (Y_t^+) of the dynamics (Y_t) started from the all-plus configuration and let $Y^* \in \{-1,1\}^F$ be distributed according to the stationary distribution ν_F. Let $\mathcal{S}_t^{(+)}$ and $\mathcal{S}^*$ again be the sum of spins over F, but with respect to the chain $(Y_t^{(+)})$ and the variable Y^* respectively. Write

$$T = \frac{n}{|F|}\bigg(\frac{1}{2}|F|\log|F| - 40|F|\bigg),$$

and let N_T be the number of steps in $[1, T]$ where a block of the form $\{v\} \cup F$ is selected to update in the chain $(Y_t^{(+)})$. By Chebyshev's inequality,

$$\mathbb{P}\bigg(N_T \geq \frac{1}{2}|F|\log|F| - 20|F|\bigg) \leq \frac{T|F|/n}{(20|F|)^2} = \mathrm{o}(1).$$

Repeating the arguments in the proof of Lemma 2.11, we deduce that for all $t \leq T_0 = \frac{1}{2}|F|\log|F| - 20|F|$, we have

$$\mathbb{P}_+\big(\mathcal{S}_t^{(+)} \leq \mathrm{e}^{10}\sqrt{|F|}\big) \leq \frac{1}{4}.$$

Therefore

$$\begin{aligned}\big\|\mathbb{P}_+\big(Y_T^{(+)} \in \cdot\big) - \nu_F\big\|_{\mathrm{TV}} \geq{}& 1 - \mathbb{P}(N_T \geq T_0) - \mathbb{P}_{\mu_Y}\big(\mathcal{S}^* \geq \mathrm{e}^{10}\sqrt{|F|}\big) \\ &- \mathbb{P}_+\big(\mathcal{S}_T^{(+)} \leq \mathrm{e}^{10}\sqrt{|F|} \mid N_T \leq T_0\big).\end{aligned}$$

Altogether, we have that

$$\big\|\mathbb{P}_+\big(Y_T^{(+)} \in \cdot\big) - \nu_F\big\|_{\mathrm{TV}} \geq \frac{1}{2} + \mathrm{o}(1) \geq \frac{1}{4},$$

and hence that

$$t_{\mathrm{mix}}^{+,Y} \geq T \geq \frac{1+\mathrm{o}(1)}{4} n \log n,$$

where $t_{\mathrm{mix}}^{+,Y}$ refers to the mixing time for chain $(Y_t^{(+)})$. Since the chain (Y_t) is a projection of the chain (X_t), it follows that the mixing time for the chain (X_t) satisfies $t_{\mathrm{mix}}^{+,X} \geq (1/4+\mathrm{o}(1))n \log n$. Combining this bound with Theorem 2.4 (see the discussion following the statement of the theorem), we conclude that the Glauber dynamics started with the all-plus configuration has mixing time $t_{\mathrm{mix}}^{+} \geq (1/4+\mathrm{o}(1))n \log n$.

Remark. *The analysis naturally extends to the continuous-time Glauber dynamics, where each site is associated with an independent Poisson clock of unit rate determining the update times of this site as above (note that the continuous dynamics is $|V|$ times faster than the discrete dynamics). We can use similar arguments to these used above to handel the laziness in the transition from the chain (Z_t) to the chain (Y_t). Namely, we could condition on the number of updates up to time t and then repeat the above arguments to establish that $t_{\mathrm{mix}}^{+} \geq (1/4+\mathrm{o}(1)) \log n$ in the continuous-time case.*

Remark. *We believe that Theorem 1 should have analogues (with t_{mix} in place of t_{mix}^{+}) for the Ising model with arbitrary magnetic field, as well as for the Potts model and proper colorings. The first of these may be accessible to the methods of this paper, but the other two models need new ideas.*

Remark. *For the Ising model in a box of $\mathbb{Z}^d$ at high temperature, it is well known that the mixing time is $\Theta(n \log n)$. Recently, the sharp asymptotics (the so-called cutoff phenomenon) was established by Lubetzky and Sly [10].*

Acknowledgments

We thank Allan Sly and Asaf Nachmias for helpful comments.

References

[1] D. Aldous. Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on Probability, XVII* 243–297. *Lecture Notes in Math.* **986**. Springer, Berlin, 1983. MR0770418

[2] D. Aldous and J. A. Fill. Reversible Markov chains and random walks on graphs. Available at http://www.stat.berkeley.edu/~aldous/RWG/book.html. To appear.

[3] R. S. Ellis and J. L. Monroe. A simple proof of the GHS and further inequalities. *Comm. Math. Phys.* **41** (1975) 33–38. MR0376053

[4] H.-O. Georgii, O. Häggström and C. Maes. The random geometry of equilibrium phases. In *Phase Transitions and Critical Phenomena* 1–142. *Phase Transit. Crit. Phenom.* **18**. Academic Press, San Diego, CA, 2001. MR2014387

[5] R. B. Griffiths, C. A. Hurst and S. Sherman. Concavity of magnetization of an Ising ferromagnet in a positive external field. *J. Math. Phys.* **11** (1970) 790–795. MR0266507

[6] T. P. Hayes and A. Sinclair. A general lower bound for mixing of single-site dynamics on graphs. *Ann. Appl. Probab.* **17** (2007) 931–952. Preliminary version appeared in *Proceedings of IEEE FOCS 2005* 511–520. MR2326236

[7] J. L. Lebowitz. GHS and other inequalities. *Comm. Math. Phys.* **35** (1974) 87–92. MR0339738

[8] D. A. Levin, M. Luczak and Y. Peres. Glauber dynamics for the mean-field Ising model: Cut-off, critical power law, and metastability. *Probab. Theory Related Fields* **146** (2009) 223–265. MR2550363

[9] D. A. Levin, Y. Peres and E. L. Wilmer. *Markov Chains and Mixing Times*. Amer. Math. Soc., Providence, RI, 2009. MR2466937

[10] E. Lubetzky and A. Sly. Cutoff for the Ising model on the lattice. Preprint, 2009.

[11] F. Martinelli. Lectures on Glauber dynamics for discrete spin models. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1997)* 93–191. *Lecture Notes in Math.* **1717**. Springer, Berlin, 1999. MR1746301

[12] Ş. Nacu. Glauber dynamics on the cycle is monotone. *Probab. Theory Related Fields* **127** (2003) 177–185. MR2013980

[13] Y. Peres. Lectures on "Mixing for Markov Chains and Spin Systems," Univ. British Columbia, August 2005. Available at http://www.stat.berkeley.edu/~peres/ubc.pdf.

[14] Y. Peres and P. Winkler. Can extra updates delay mixing? To appear.

[15] A. Sinclair. *Algorithms for Random Generation and Counting*. *Progress in Theoretical Computer Science*. Birkhäuser, Boston, MA, 1993. MR1201590