CUTOFF FOR THE ISING MODEL ON THE LATTICE

EYAL LUBETZKY AND ALLAN SLY

Abstract. Introduced in 1963, Glauber dynamics is one of the most practiced and extensively studied methods for sampling the Ising model on lattices. It is well known that at high temperatures, the time it takes this chain to mix in L^1 on a system of size n is $O(\log n)$. Whether in this regime there is *cutoff*, i.e. a sharp transition in the L^1-convergence to equilibrium, is a fundamental open problem: If so, as conjectured by Peres, it would imply that mixing occurs abruptly at $(c + o(1)) \log n$ for some fixed $c > 0$, thus providing a rigorous stopping rule for this MCMC sampler. However, obtaining the precise asymptotics of the mixing and proving cutoff can be extremely challenging even for fairly simple Markov chains. Already for the one-dimensional Ising model, showing cutoff is a longstanding open problem.

We settle the above by establishing cutoff and its location at the high temperature regime of the Ising model on the lattice with periodic boundary conditions. Our results hold for any dimension and at any temperature where there is strong spatial mixing: For $\mathbb{Z}^2$ this carries all the way to the critical temperature. Specifically, for fixed $d \geq 1$, the continuous-time Glauber dynamics for the Ising model on $(\mathbb{Z}/n\mathbb{Z})^d$ with periodic boundary conditions has cutoff at $(d/2\lambda_\infty) \log n$, where λ_∞ is the spectral gap of the dynamics on the infinite-volume lattice. To our knowledge, this is the first time where cutoff is shown for a Markov chain where even understanding its stationary distribution is limited.

The proof hinges on a new technique for translating L^1-mixing to L^2-mixing of projections of the chain, which enables the application of logarithmic-Sobolev inequalities. The technique is general and carries to other monotone and anti-monotone spin-systems, e.g. gas hard-core, Potts, anti-ferromagentic Ising, arbitrary boundary conditions, etc.

1. Introduction

The total-variation *cutoff phenomenon* describes a sharp transition in the L^1-mixing of a finite ergodic Markov chain: Over a negligible time period, the distance of the chain from equilibrium drops abruptly from near its maximum to near 0. Though believed to be widespread, including many important families of chains arising from statistical physics, cutoff has been rigorously shown only in relatively few cases (ones where the stationary distribution is completely understood and has many symmetries, e.g. uniform on the symmetric group). Here we establish cutoff for Glauber dynamics for the Ising model, one of the most studied models in mathematical physics.

Already establishing the order of the mixing time is in many cases challenging, with an entire industry devoted to the study of such problems. Proving cutoff and its location entails not only obtaining the order, but also deriving the precise asymptotics of the time it takes the chain to mix. In his 1995 survey of the cutoff phenomenon Diaconis [7] wrote "*At present writing, proof of a cutoff is a difficult, delicate affair, requiring detailed knowledge of the chain, such as all eigenvalues and eigenvectors. Most of the examples where this can be pushed through arise from random walk on groups, with the walk having a fair amount of symmetry*". To this date this essentially remained to be the case, and present technology (e.g., representation theory, spectral theory, techniques for analyzing 1-dimensional chains, etc.) does not suffice for proving cutoff in high-dimensional chains with limited understanding of their stationary distribution, such as Glauber dynamics for the Ising model (stochastic Ising model) on the 3-dimensional lattice.

Introduced in 1963 [16], Glauber dynamics for the Ising model on the lattice (see Section 1.1 for formal definitions) is one of the most practiced methods to sample the Gibbs distribution, and an extensively studied dynamical system in itself, having a rich interplay of properties with the static stationary distribution. For instance, as we describe in Section 1.1, it is known that on $(\mathbb{Z}/\mathbb{Z}_n)^2$, at the critical inverse-temperature β_c for uniqueness of the static Gibbs distribution, the spectral gap of the Markov semigroup generator of the Glauber dynamics exhibits a phase-transition from being uniformly bounded to tending to 0 exponentially fast in n. It is further known that at high temperatures on $(\mathbb{Z}/\mathbb{Z}_n)^d$ this dynamics mixes in time $O(\log n)$, yet the precise asymptotics of the L^1-mixing time were unknown even in the one-dimensional case: It is an open problem of Peres (cf. [23]) to determine cutoff for the Ising model on $\mathbb{Z}/\mathbb{Z}_n$.

The only underlying geometry for which cutoff for the Ising model had so far been established is the complete graph ([13, 22]), where the high symmetry reduces the analysis to a birth-and-death magnetization chain. However, this sheds no light on the existence of cutoff for lattices, where there is no such reduction. Peres ([22, 23]) conjectured that in any dimension d, Glauber dynamics for the Ising model on $(\mathbb{Z}/\mathbb{Z}_n)^d$ should exhibit cutoff.

Our main results confirm the above conjecture and moreover establish cutoff and its location in a wide range of spin system models and geometries. We first formulate this for the classical two-dimensional Ising model.

Theorem 1. *Let $\beta_c = \frac{1}{2}\log(1+\sqrt{2})$ be the critical inverse-temperature for the Ising model on $\mathbb{Z}^2$. Then the continuous-time Glauber dynamics for the Ising model on $(\mathbb{Z}/n\mathbb{Z})^2$ at inverse-temperature $0 \leq \beta < \beta_c$ with periodic boundary conditions has cutoff at $\lambda_\infty^{-1}\log n$ with a window of $O(\log\log n)$, where λ_∞ is the spectral gap of the dynamics on the infinite-volume lattice.*

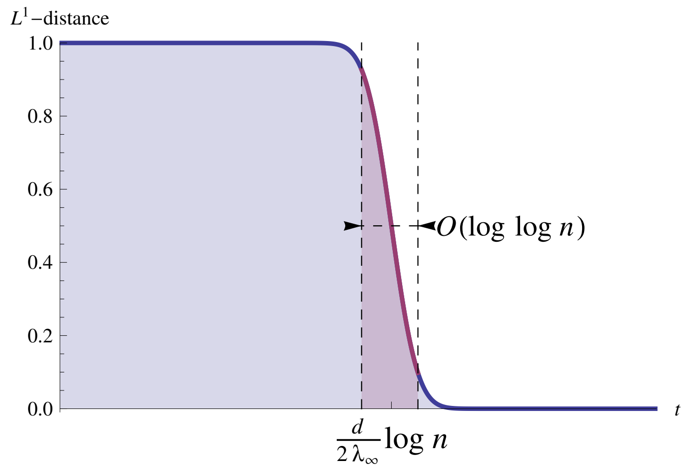

FIGURE 1. Cutoff phenomenon for the L^1 (total-variation) distance from stationarity along time in Glauber dynamics for the Ising model on $\mathbb{Z}_n^d$, as established by Theorem 2. Highlighted region denotes the cutoff window of $O(\log\log n)$.

In the above theorem, the term *cutoff window* refers to the rate at which the L^1-distance from stationarity drops from near 1 to near 0. More precisely, let $t_{\mathrm{MIX}}(\varepsilon)$ be the minimum $t \geq 0$ where the heat-kernel H_t associated with a Markov chain is within a total-variation distance of ε from stationarity. A family of chains is said to exhibit cutoff if for every fixed $0 < \varepsilon < \frac{1}{2}$ we have $t_{\mathrm{MIX}}(\varepsilon)/t_{\mathrm{MIX}}(1-\varepsilon) \to 1$ as the system size tends to ∞. A sequence w_n is said to be a cutoff window if $t_{\mathrm{MIX}}(\varepsilon) = t_{\mathrm{MIX}}(1-\varepsilon) + O(w_n)$ for every ε.

Our results hold for $(\mathbb{Z}/n\mathbb{Z})^d$ in any dimension d, inverse-temperature β and external field h so that the corresponding static Gibbs distribution has a spatial dependence property known as *strong spatial mixing* (and also as *regular complete analyticity*). This property, defined by Martinelli and Olivieri in their seminal paper [30], holds in all regimes where $O(\log n)$ mixing is known for Glauber dynamics for the Ising model. In particular (see [30–32]), on $\mathbb{Z}^2$ there is strong spatial mixing for any β with an external field $h \neq 0$, as well as for any $0 \leq \beta < \beta_c$ when $h = 0$.

The next result settles the conjecture of Peres for cutoff for the high temperature regime of the Ising model in any dimension $d \geq 1$.

Theorem 2. *Let $d \geq 1$ and consider the continuous-time Glauber dynamics for the ferromagnetic Ising model on $(\mathbb{Z}/n\mathbb{Z})^d$ with periodic boundary conditions, inverse-temperature β and external field h. Suppose that β, h are such that there is strong spatial mixing. Then the dynamics exhibits cutoff at $(d/2\lambda_\infty)\log n$ with a window of $O(\log\log n)$, where λ_∞ is the spectral gap of the dynamics on the infinite-volume lattice.*

In the special case of $d = 1$ and no external field, it is known that strong spatial mixing always holds and that the spectral gap at inverse-temperature β is $1 - \tanh(2\beta)$ independent of the system size (cf., e.g., [23]). We thus have the following corollary to establish the asymptotic mixing time of the one-dimensional Ising model, answering the aforementioned question of [23].

Corollary 3. *For any $\beta \geq 0$, the continuous-time Glauber dynamics for the Ising model on $\mathbb{Z}/n\mathbb{Z}$ with periodic boundary conditions, inverse-temperature β and no external field has cutoff at $\frac{1}{2}(1 - \tanh(2\beta))^{-1} \log n$.*

Our proofs determine the cutoff location in terms of $\lambda(r)$, the spectral gap of the dynamics on the d-dimensional lattice $(\mathbb{Z}/r\mathbb{Z})^d$ for a certain $r = r(n)$. As a biproduct, we are able to relate the spectral gap on tori of varying sizes and obtain that they converge polynomially fast to the spectral gap of the dynamics on the infinite-volume lattice.

Theorem 4. *For $d \geq 1$ let $\lambda(n)$ be the spectral gap of the continuous-time Glauber dynamics for the Ising model on $(\mathbb{Z}/n\mathbb{Z})^d$ with inverse-temperature $\beta \geq 0$ and external field h. If there is strong spatial mixing for β, h then*

$$|\lambda(n) - \lambda_\infty| \leq n^{-1/2+o(1)}\,,$$

where λ_∞ is the spectral gap of the dynamics on the infinite-volume lattice.

1.1. **Background and previous work.** While our results hold in greater generality, we will focus on single-site uniform interactions for the sake of the exposition: The *Ising model* on a finite graph with vertex-set V and edge-set E is defined as follows. Its set of possible configurations is $\Omega = \{\pm 1\}^V$, where each configuration corresponds to an assignment of plus/minus spins to the sites in V. The probability that the system is in a configuration $\sigma \in \Omega$ is given by the Gibbs distribution

$$\mu(\sigma) = \frac{1}{Z(\beta)} \exp\left(\beta \sum_{uv \in E} \sigma(u)\sigma(v) + h \sum_{u \in V} \sigma(u)\right)\,, \tag{1.1}$$

where the partition function $Z(\beta)$ is a normalizing constant. The parameters β and h are the inverse-temperature and external field respectively; for $\beta \geq 0$ we say that the model is ferromagnetic, otherwise it is anti-ferromagnetic. These definitions extend to infinite locally finite graphs (see e.g. [24, 28]).

We denote the boundary of a set $\Lambda \subset V$ as the neighboring sites of Λ in $V \setminus \Lambda$ and call $\tau \in \{\pm 1\}^{\partial\Lambda}$ a *boundary condition*. A periodic boundary condition on $(\mathbb{Z}/n\mathbb{Z})^d$ corresponds to a d-dimensional torus of side-length n.

The *Glauber dynamics* for the Ising model is a family of continuous-time Markov chains on the state space Ω, reversible with respect to the Gibbs distribution, given by the generator

$$(\mathcal{L}f)(\sigma) = \sum_{x \in V} c(x,\sigma)\,(f(\sigma^x) - f(\sigma)) \tag{1.2}$$

where σ^x is the configuration σ with the spin at x flipped. The transition rates $c(x,\sigma)$ are chosen to satisfy finite range interactions, detailed balance, positivity and boundedness and translation invariance (see Section 2). Two notable examples for the transition rates are

(i) *Metropolis*: $c(x,\sigma) = \exp\Big(2h\sigma(x) + 2\beta\sigma(x)\sum_{y\sim x}\sigma(y)\Big) \wedge 1$.

(ii) *Heat-bath*: $c(x,\sigma) = \Big[1 + \exp\Big(- 2h\sigma(x) - 2\beta\sigma(x)\sum_{y\sim x}\sigma(y)\Big)\Big]^{-1}$.

These chains have useful graphical interpretations: for instance, heat-bath Glauber dynamics is equivalent to updating the spins via i.i.d. rate-one Poisson clocks, each time resetting a spin according to the conditional distribution given its neighbors.

Ever since its introduction in 1925, the static properties of the Ising model and its Gibbs states, and more recently the Glauber dynamics for this model, have been the focus of intensive research. A series of breakthrough papers by Aizenman, Dobrushin, Holley, Shlosman, Stroock et al. (cf., e.g., [1, 15, 19–21, 24, 25, 30–32, 35–39]) starting from the late 1970's has developed the theory of the convergence rate of the Glauber dynamics to stationarity. It was shown by Aizenman and Holley [1] that the spectral gap of the dynamics on the infinite-volume lattice is uniformly bounded whenever the Dobrushin-Shlosman uniqueness condition holds. Stroock and Zegarliński [35, 37, 38] proved that the logartihmic-Sobolev constant is uniformly bounded provided given the Dobrushin-Shlosman mixing conditions (complete analyticity). Finally, Martinelli and Olivieri [30, 31] obtained this for cubes under the more general condition of strong spatial mixing. This in particular established $O(\log n)$ mixing throughout the uniqueness regime in two-dimensions. See the excellent surveys [28, 29] for further details.

To conclude this collection of seminal papers that altogether established $O(\log n)$ mixing throughout the regime of strong spatial mixing, it remains to pinpoint the asymptotics of the mixing time and determine whether or not there is cutoff in this regime.

The cutoff phenomenon was first identified for random transpositions on the symmetric group in [12], and for the riffle-shuffle and random walks on the hypercube in [2]. The term "cutoff" was coined by Aldous and Diaconis in [4], where cutoff was shown for the top-in-at-random card shuffling process. See [6, 7, 34] and the references therein for more on the cutoff phenomenon. In these examples, and most others where cutoff has been rigorously shown, the stationary distribution has many symmetries or is essentially one-dimensional (e.g. uniform on the symmetric group [12], uniform on the hypercube [2] and one-dimensional birth-and-death chains [14]). Even for random walks on random regular graphs (where the stationary distribution is uniform), cutoff was only recently verified [27].

In the context of spin systems, cutoff was conjectured by Peres (see [22, Conjecture 1]) for the Glauber dynamics on any sequence of transitive graphs on n vertices where its mixing time is $O(\log n)$. More specifically, it was conjectured in [22] that cutoff holds for the Ising model on any d-dimensional torus $(\mathbb{Z}/n\mathbb{Z})^d$ in the high temperature regime; see also [23, Question 8], where the special case of $d = 1$ (Ising model on the cycle) was emphasized.

However, so far the only spin-system where cutoff has been established at some inverse-temperature $\beta > 0$ is Glauber dynamics for the Ising model on the complete graph [13, 22]. There, the magnetization (sum-of-spins) is in fact a one-dimensional Markov chain whose mixing and cutoff govern that of the entire dynamics. While this result motivates the conjecture on cutoff for the Ising model on lattices, its proof fails to provide insight for the latter setting since the complete graph has no geometry to consider.

1.2. **Cutoff for spin systems on the lattice.** Our proof of cutoff for the Ising model on the d-dimensional lattice, as stated in Theorems 1,2, in fact applies to a broad class of spin systems (essentially any monotone or anti-monotone system), underlying geometries and boundary conditions. We demonstrate this by establishing cutoff for the anti-ferromagnetic Ising model and the gas hard-core model (see, e.g., [23] for definitions).

Theorem 5. *The cutoff result given in Theorem 2 for the ferromagnetic Ising model also holds for the anti-ferromagnetic Ising model.*

Theorem 6. *Let $d \geq 1$. The following holds for the Glauber dynamics for the gas hard-core model on $(\mathbb{Z}/n\mathbb{Z})^d$ with fugacity β and periodic boundary conditions. If β is such that there is strong spatial mixing, then the dynamics exhibits cutoff at $(d/2\lambda_\infty)\log n$ with window of $O(\log\log n)$, where λ_∞ is the spectral gap of the dynamics on the infinite volume lattice.*

Our methods also establish cutoff on more general spin-systems such as the Potts model, provided that the temperature is sufficiently high. This is discussed further in the companion paper [26].

Furthermore, our results are also not confined to the underlying geometry of the cubic lattice and in fact Theorem 2 holds for any non-amenable translation invariant lattice (e.g. triangular/hexagonal/ladder lattice etc.).

Similarly, the choice of periodic boundary conditions is not a prerequisite for establishing cutoff, though it does enables us to determine the mixing time in terms of the infinite-volume spectral gap. For *arbitrary* boundary

FIGURE 2. Sparse geometry of the update support along a short time interval (mininum subset of sites determining the final configuration, conditioned on the update sequence), simulated on a 500×500 square lattice at $\beta = 0.4$. Colormap highlights the last time a spin belonged to the support (black being earliest and white being latest).

conditions we can still establish the existence of cutoff and in various special cases (e.g. under the all-plus boundary conditions) we can pinpoint its location as stated by the following results of the companion paper [26].

Theorem 7 ([26]). *Let $d \geq 1$ and consider the Glauber dynamics for the Ising model on $(\mathbb{Z}/n\mathbb{Z})^d$ with arbitrary boundary conditions. If there is strong spatial mixing then the dynamics exhibits cutoff.*

Theorem 8 ([26]). *Consider Glauber dynamics for the Ising model on $(\mathbb{Z}/n\mathbb{Z})^2$ with all-plus boundary condition. If there is strong spatial mixing then the dynamics exhibits cutoff at $(\lambda_\infty \wedge 2\lambda_{\mathbb{H}})^{-1} \log n$, where $\lambda_\infty, \lambda_{\mathbb{H}}$ are the spectral gaps of the dynamics on the infinite-volume lattice and on the halfplane with all-plus boundary condition respectively.*

1.3. **Main techniques.** To prove total-variation cutoff on the cubic lattice of dimension d, we introduce a technique for bounding the L^1-mixing of the dynamics on $(\mathbb{Z}/n\mathbb{Z})^d$ from below and above in terms of a quantity that measures the L^2-mixing of the dynamics on smaller lattices $(\mathbb{Z}/r\mathbb{Z})^d$ with $r \asymp \log^3 n$ (see the definition of $\mathfrak{m}_t$ in (3.1)).

A key ingredient in this technique is an analysis of the geometry of the "update support": the minimum subset of sites whose initial spins, conditioned on the update sequence in a given time period, determine the final configuration. We show that even for relatively short update sequences, this minimum subset of sites is typically "sparse", i.e. comprises remote clusters of small diameter. Namely, the clusters have diameter $O(\log^3 n)$ and the pairwise distances between distinct clusters are all at least of order $\log^2 n$ (see Definition 3.3 of a sparse set). Fundamentally, this is ensured by the monotonicity of the system and the strong spatial mixing property, as the dynamics on cubes mixes in timescales much smaller than their diameter. This is illustrated in Figure 2 where the clusters are shown in white.

We reduce the analysis of the L^1-mixing of the dynamics on $(\mathbb{Z}/n\mathbb{Z})^d$ to its projection onto sparse supports. For such subsets, the speed of propagation of information between the distant clusters makes the projections onto them essentially independent. Using some additional arguments, from here we can translate the above to a product chain on smaller tori $(\mathbb{Z}/r\mathbb{Z})^d$. Finally, the use of log-Sobolev inequalities provides a tight control over the L^1-mixing in terms of the aforementioned L^2-quantity, as stated in Theorem 3.1.

2. Preliminaries

Consider the Ising model on set of spins V, let μ be its Gibbs distribution as given in (1.1) and $\sigma \in \Omega = \{\pm 1\}^V$ be a configuration. For $\Lambda \subset V$, we will let $\sigma(\Lambda)$ and μ_Λ denote the projections of σ and μ onto $\{\pm 1\}^\Lambda$ respectively. We let μ_Λ^τ denote the measure on Λ given the boundary condition τ, that is, the conditional measure $\mu_\Lambda(\cdot \mid \sigma_{\partial\Lambda} = \tau)$.

2.1. **Glauber dynamics for the Ising model.** The Glauber dynamic for the Ising model on the lattice $V = \mathbb{Z}^d$, whose generator is given in (1.2), accepts any choice of transition rates $c(x,\sigma)$ which satisfy the following:

(1) *Finite range interactions*: For some fixed $R > 0$ and any $x \in V$, if $\sigma, \sigma' \in \Omega$ agree on the ball of diameter R about x then $c(x,\sigma) = c(x,\sigma')$.
(2) *Detailed balance*: For all $\sigma \in \Omega$ and $x \in V$,

$$\frac{c(x,\sigma)}{c(x,\sigma^x)} = \exp\Big(2h\sigma(x) + 2\beta\sigma(x)\sum_{y\sim x}\sigma(y)\Big)\,.$$

(3) *Positivity and boundedness*: The rates $c(x,\sigma)$ are uniformly bounded from below and above by some fixed $C_1, C_2 > 0$.
(4) *Translation invariance*: If $\sigma \equiv \sigma'(\cdot + \ell)$, where $\ell \in V$ and addition is according to the lattice metric, then $c(x,\sigma) = c(x+\ell, \sigma')$ for all $x \in V$.

The Glauber dyanmics generator with such rates defines a unique Markov process, reversible with respect to the Gibbs measure μ_V^τ. For simplicity of the exposition, the proof is given for the cases of heat-bath Glauber dynamics or Metropolis Glauber dynamics. The results for other choices of transition rates follow with minor adjustments to the arguments.

We will let $\mathbb{P}_{\sigma_0}(\cdot)$ and $\mathbb{E}_{\sigma_0}[\cdot]$ denote the probability and expectation conditioned on the Glauber dynamics having initial configuration σ_0.

2.2. **Mixing, spectral gap and the logarithmic-Sobolev constant.** The L^1 (total-variation) distance is perhaps the most fundamental notion of convergence in the theory of Markov chains. For two probability measures ν_1, ν_2 on a finite space Ω the total-variation distance is defined as

$$\|\nu_1 - \nu_2\|_{\mathrm{TV}} = \max_{A \subset \Omega} |\nu_1(A) - \nu_2(A)| = \frac{1}{2} \sum_{x \in \Omega} |\nu_1(x) - \nu_2(x)| \ ,$$

i.e. half the L^1-distance between the two measures. For an ergodic Markov chain (Y_t) with stationary distribution ν, the mixing-time notion t_{MIX} is defined with respect to $\|\mathbb{P}(Y_t \in \cdot) - \mu\|_{\mathrm{TV}}$.

The spectral gap and log-Sobolev constant of the continuous-time Glauber dynamics are given by the following Dirichlet form (see, e.g., [28, 33]):

$$\lambda = \inf_f \frac{\mathcal{E}(f)}{\mathrm{Var}(f)} \ , \qquad \alpha_{\mathsf{s}} = \inf_f \frac{\mathcal{E}(f)}{\mathrm{Ent}(f)} \ , \tag{2.1}$$

where the infimum is over all nonconstant $f \in L^2(\mu)$ and

$$\mathcal{E}(f) = \langle \mathcal{L} f, f \rangle_{L^2(\mu)} = \frac{1}{2} \sum_{\sigma, x} \mu(\sigma) c(x,\sigma) \left[f(\sigma^x) - f(\sigma) \right]^2 \ ,$$
$$\mathrm{Ent}(f) = \mathbb{E}\left[f^2(\sigma) \log\left(f^2(\sigma)/\mathbb{E} f^2(\sigma) \right) \right] \ .$$

It is well known (see e.g. [3, 10]) that for any finite ergodic reversible Markov chain $0 < 2\alpha_{\mathsf{s}} < \lambda$ and $\lambda^{-1} \le t_{\mathrm{MIX}}(1/\mathrm{e})$. In our case, since the sites are updated via rate-one independent Poisson clocks, we also have $\lambda \le 1$.

By bounding the log-Sobolev constant one may obtain remarkably sharp upper bounds not only for the total-variation mixing-time but also for the L^2-mixing (cf., e.g., [8–11, 33]). The following theorem of Diaconis and Saloff-Coste [10, Theorem 3.7] (see also [33],[3, Chapter 8]) demonstrates this powerful method.

Theorem 2.1. *Let (Y_t) be a finite reversible continuous-time Markov chain with stationary distribution ν. For any x with $\nu(x) \leq \mathrm{e}^{-1}$ and any $s > 0$,*

$$\|\mathbb{P}_x(X_s \in \cdot) - \pi\|_{L^2(\nu)} \leq \exp\left(1 - \lambda\left(s - \frac{1}{4\alpha_{\mathtt{s}}}\log\log\frac{1}{\nu(x)}\right)\right).$$

2.3. **Strong spatial mixing and logarithmic-Sobolev inequalities.** As noted in the introduction, bounds on the log-Sobolev constant of the Glauber dynamics for the Ising model were proved under a variety of increasingly general spatial mixing conditions. We will work under the assumption of strong spatial mixing (or regular complete analyticity) introduced by a Martinelli and Oliveri [30] as it holds for the largest known range of β.

Definition 2.2. For a set $\Lambda \subset \mathbb{Z}^d$ we say that $\mathtt{SM}(\Lambda, c_1, c_2)$ holds if there exist constants $c_1, c_2 > 0$ such that for any $\Delta \subset \Lambda$,

$$\sup_{\tau,y} \left\|\left(\mu_\Lambda^\tau\right)_\Delta - \left(\mu_\Lambda^{\tau^y}\right)_\Delta\right\|_{\mathrm{TV}} \leq c_1 \mathrm{e}^{-c_2 \operatorname{dist}(y,\Delta)},$$

where the supremum is over all $y \in \partial\Lambda$ and $\tau \in \{\pm 1\}^{\partial\Lambda}$ and where $(\mu_\Lambda^\tau)_\Delta$ is the projection of the measure μ_Λ^τ onto Δ. We say that strong spatial mixing holds for the Ising model with inverse temperature β and external field h on $\mathbb{Z}^d$ if there exist $L, c_1, c_2 > 0$ such that $\mathtt{SM}(Q, c_1, c_2)$ holds for all cubes Q of side-length at least L.

The above definition implies uniqueness of the Gibbs measure on the infinite lattice. Moreover, strong spatial mixing holds for all temperatures when $d = 1$ and for $d = 2$ it holds whenever $h \neq 0$ or $\beta < \beta_c$. As discussed in the introduction, this condition further implies a uniform lower bound on the log-Sobolev constant of the Glauber dynamics on cubes under any boundary condition τ (see [28, 30, 31]). We will make use of the next generalization of this result to periodic boundary conditions, i.e. the dynamics on the torus, obtained by following the original arguments as given in [28] with minor alterations (see also [5, 17, 25, 29]).

Theorem 2.3. *Suppose that the inverse-temperature β and external field h are such that the Ising model on $\mathbb{Z}^d$ has strong spatial mixing. Then there exists a constant $\alpha_{\mathtt{s}}^\star = \alpha_{\mathtt{s}}^\star(\beta, h) > 0$ such that the Glauber dynamics for the Ising model on $(\mathbb{Z}/n\mathbb{Z})^d$ with periodic boundary conditions has a log-Sobolev constant at least $\alpha_{\mathtt{s}}^\star$ independent of n.*

3. Reducing L^1 mixing to L^2 local mixing

In this section, we show that the L^1 (total-variation) distance of the dynamics on the lattice from stationarity is essentially determined by the L^2 distance from stationarity of a projection of this chain onto smaller boxes.

More precisely, consider the continuous-time Glauber dynamics (X_t) for the Ising model on $\mathbb{Z}_n^d$, the d-dimensional lattice with side-length n and periodic boundary conditions, and let μ denote its Gibbs distribution. Further, consider such a chain on a smaller lattice, namely (X_t^*) on $\mathbb{Z}_r^d$ for $r = 3\log^3 n$, and let μ^* denote its stationary distribution. (We actually have a lot of freedom in the choice of r, e.g. any poly-logarithmic value which is at least $\log^{2+\delta} n$ for some fixed $\delta > 0$ would do; this will prove useful later on for relating the cutoff location to the spectral gap of the dynamics on the infinite-volume lattice.) Within this smaller lattice we consider a d-dimensional box B with side-length $2\log^3 n$ (the location of the box B within $\mathbb{Z}_r^d$ does not play a role as the boundary is periodic). Define

$$\mathfrak{m}_t = \max_{x_0} \left\| \mathbb{P}_{x_0}\big(X_t^*(B) \in \cdot\big) - \mu_B^* \right\|_{L^2(\mu_B^*)} , \tag{3.1}$$

where $X_t^*(B)$ and μ_B^* are the projections of X_t^* and μ^* resp. onto the box B. The following theorem demonstrates how the L^2 mixing measured by the quantity $n^d \mathfrak{m}^2$ governs the L^1 mixing of (X_t).

Theorem 3.1. *Let (X_t) be the continuous-time Glauber dynamics for the Ising model on $\mathbb{Z}_n^d$, and define $\mathfrak{m}_t$ as in (3.1). The following then holds:*

1. *Let $s = s(n)$ and $t = t(n)$ satisfy $(10d/\alpha_{\mathsf{s}}^\star)\log\log n \le s < \log^{4/3} n$ and $0 < t < \log^{4/3} n$. For any sufficiently large n,*

$$\max_{x_0} \|\mathbb{P}_{x_0}(X_{t+s} \in \cdot) - \mu\|_{\mathrm{TV}} \le \frac{1}{2}\Big(\exp\big((n/\log^5 n)^d\, \mathfrak{m}_t^2\big) - 1\Big)^{1/2} + 3n^{-9d}\,.$$

In particular, if $(n/\log^5 n)^d\, \mathfrak{m}_t^2 \to 0$ as $n \to \infty$ for the above s, t then

$$\limsup_{n\to\infty} \max_{x_0} \|\mathbb{P}_{x_0}(X_{t+s} \in \cdot) - \mu\|_{\mathrm{TV}} = 0\,.$$

2. *If $(n/\log^3 n)^d \mathfrak{m}_t^2 \to \infty$ for some $t \ge (20d/\alpha_{\mathsf{s}}^\star)\log\log n$, then*

$$\liminf_{n\to\infty} \max_{x_0} \|\mathbb{P}_{x_0}(X_t \in \cdot) - \mu\|_{\mathrm{TV}} = 1\,.$$

Remark 3.2. It will be useful to apply Part 1 of the above theorem to lattices of varying sizes. Indeed, we will show that if (X_t) is the continuous-time Glauber dynamics for the Ising model on $\mathbb{Z}_m^d$ with

$$\log^3 n \le m \le n\,,$$

and with $s, t, \mathfrak{m}_t$ as in Theorem 3.1 (e.g., $s, t < \log^{4/3} n$, the box B in the definition of $\mathfrak{m}_t$ given in (3.1) has side length $2\log^3 n$, etc.), then

$$\max_{x_0} \|\mathbb{P}_{x_0}(X_{t+s} \in \cdot) - \mu\|_{\mathrm{TV}} \le \frac{1}{2}\Big(\exp\big(m^d\, \mathfrak{m}_t^2\big) - 1\Big)^{1/2} + 3n^{-9d}\,. \tag{3.2}$$

The upper and lower bounds stated in the above theorem appear in Subsections 3.2 and 3.3 respectively. We begin by describing two key ingredients in the proof of the L^1-L^2 reduction that enable us to eliminate long-range dependencies between spins and break down the lattice into smaller independent blocks. First we analyze which spins effectively influence the final configuration at some designated target time and characterize the geometric structure of components comprising such spins. Second, we introduce the *barrier-dynamics*, a variant of the Glauber dynamics that separates the lattice into weakly-dependent blocks by surrounding each one with a periodic-boundary barrier.

3.1. **Eliminating long-range dependencies.** Consider some time point $t = t(n) \asymp \log n$ just prior to mixing, and let $s = s(n) \asymp \log\log n$ be a short time-frame. Our goal in this section is to bounds the L^1-distance of the Glauber dynamics from equilibrium at time $t+s$ in terms of the L^1-distance projected onto sparse subsets of the spins.

Definition 3.3 (*Sparse set*)**.** Let $\log^3 n \leq m \leq n$. We say that the set $\Delta \subset \mathbb{Z}_m^d$ is *sparse* if for some $L \leq (n/\log^5 n)^d$ it can be partitioned into components $A_1, \ldots, A_L$ such that

1. Every A_i has diameter at most $\log^3 n$ in $\mathbb{Z}_m^d$.
2. The distance in $\mathbb{Z}_m^d$ between any distinct A_i, A_j is at least $2d\log^2 n$.

Let $\mathcal{S} = \mathcal{S}(m) = \{\Delta \subset \mathbb{Z}_m^d : \Delta \text{ is sparse}\}$.

Theorem 3.4. *For $\log^3 n \leq m \leq n$ let (X_t) be the Glauber dynamics on $\mathbb{Z}_m^d$ and μ its stationary measure. Let $(10d/\alpha_{\mathsf{s}}^\star)\log\log n \leq s \leq \log^{4/3} n$ and $t > 0$. Then there exists some distribution ν on $\mathcal{S}(m)$ such that*

$$\|\mathbb{P}_{x_0}(X_{t+s} \in \cdot) - \mu\|_{\mathrm{TV}} \leq \int_{\mathcal{S}} \|\mathbb{P}_{x_0}(X_t(\Delta) \in \cdot) - \mu_\Delta\|_{\mathrm{TV}} \, d\nu(\Delta) + 3n^{-10d}\,.$$

Proof. To prove the above theorem, we introduce the following variant of the Glauber dynamics which breaks down the dynamics into smaller blocks over which we have better control.

Definition 3.5 (*Barrier-dynamics*)**.** Let (X_t) be the Glauber dynamics for the Ising model on $\mathbb{Z}_m^d$. Define the corresponding *barrier-dynamics* as the following coupled Markov chain:

(1) Partition the lattice into disjoint d-dimensional boxes (or blocks), where each side-length is either $\log^2 n$ or $\log^2 n - 1$.
(2) For each such box B, let B^+ be the d-dimensional box centered at B with side-lengths $\log^2 n + 2\log^{3/2} n$, e.g., if B has side-length $\log^2 n$

$$B^+ = \bigcup_{v \in B} \{u : \|u - v\|_\infty \leq \log^{3/2} n\}\,.$$

Let ψ_B be a graph isomorphism mapping B^+ onto some block C^+ (and B onto $C \subset C^+$), where the C^+ blocks are pairwise disjoint.

(3) Impose a periodic boundary condition on each C^+ (to be thought of as a barrier surrounding it), and run the following dynamics: As usual, each site u in $\mathbb{Z}_m^d$ receives updates according to a unit rate Poisson clock. Updating u at time t via a variable $I \sim U[0,1]$ in the standard dynamics implies updating every $v = \psi_B(u)$ (for some B with $u \in B^+$) via the same update variable I.

The above Markov chain gives rise to the following randomized operator $\mathcal{G}_s$ (for $s > 0$) on $\{\pm 1\}^{\mathbb{Z}_m^d}$. Given an initial configuration x_0 for $\mathbb{Z}_m^d$, we translate it to a configuration for the C^+ blocks in the obvious manner, and then run the barrier-dynamics for time s. The output of the operator $\mathcal{G}_s$ is obtained by assigning each $u \in \mathbb{Z}_m^d$ the value of $\psi_B(u)$, where B is the (unique) block that contains u. In other words, we pull-back the configuration from the C's onto $\mathbb{Z}_m^d$ (while discarding the spins of the overlaps).

To simplify the notations, we identify the blocks C, C^+ with B, B^+ whenever there is no danger of confusion.

Note that for the mixing-time analysis, we are only interested in the behavior of the dynamics up to time $O(\log n)$, and the above parameters were chosen accordingly. Indeed, the next lemma shows that the barrier-dynamics can be coupled to the original one up to time $(\log n)^{4/3}$ except with a negligible error-probability.

Lemma 3.6. *Let $t_0 = (\log n)^{4/3}$. The barrier-dynamics and the original Glauber dynamics are coupled up to time t_0 except with probability n^{-10d}. That is, except with probability n^{-10d}, for any X_0 we have $X_s = \mathcal{G}_s(X_0)$ simultaneously for all $s \le t_0$.*

Proof. Let (X_t) denote the Glauber dynamics on $\mathbb{Z}_m^d$ and let $(\tilde{X}_t)$ denote the barrier-dynamics with corresponding blocks B_i and B_i^+. Apply the aforementioned coupling between the two processes, where the original Glauber dynamics runs as usual, and the barrier-dynamics uses the same updates for each of its sites. That is, if site u is updated in the original dynamics via a uniform real $I \sim U[0,1]$, we update it in every B^+ that contains it using the same I at the same time.

Consider some box B and its block B^+. Clearly, the barrier-dynamics on B^+ is identical to the original dynamics until it needs to update ∂B^+, the boundary of B^+ (in which case $(\tilde{X}_t)$ has periodic conditions whereas (X_t) uses external sites in the lattice).

Therefore, a necessary condition to have $X_t(v) \neq \tilde{X}_t(v)$ for some $v \in B$ is the existence of a path of adjacent sites $u_1, \ldots, u_\ell$ connecting v to ∂B^+, and a sequence of times $t_1 < \ldots < t_\ell \le t$ such that site u_i was updated at time t_i (note that $\ell \geq \log^{3/2} n$ by definition). Summing over all $(2d)^\ell$ possible paths originating from v, and accounting for the probability that the ℓ corresponding rate 1 Poisson clocks fire sequentially before time $t \leq t_0$ (while recalling that $t_0 = o(\ell)$) it then follows that

$$\sum_{t \leq t_0} \mathbb{P}\left(\cup_i \left\{X_t(B_i) \neq \tilde{X}_t(B_i)\right\}\right) \leq t_0 n^d \sum_{\ell \geq \log^{3/2} n} (2d)^\ell \, \mathbb{P}(\mathrm{Po}(t_0) \geq \ell)$$

$$\leq 2t_0 n^d \mathrm{e}^{-t_0} \sum_{\ell \geq \log^{3/2} n} \frac{(2dt_0)^\ell}{\ell!} \leq n^{d-\sqrt{\log n}} < n^{-10d}\,,$$

where the inequalities hold for any large n (with room to spare). ∎

In light of the above lemma, we can focus on the barrier-dynamics for the sake of proving Theorem 3.4. Crucially, suitably distant sites evolve independently in this new setting.

The random operator $\mathcal{G}_s$ is determined by the random update sequence W_s (each update is a tuple (u_j, t_j, I_j), where u_j is the site that was updated, t_j is the time of update and I_j is the unit variable determining the update result). In other words, for any such sequence W_s there exists some deterministic function $g_{W_s} : \{\pm 1\}^{\mathbb{Z}_m^d} \to \{\pm 1\}^{\mathbb{Z}_m^d}$ so that $\mathcal{G}_s(x) = g_{W_s}(x)$ for all x. Further note that g_{W_s} is monotone, by the monotonicity of the Ising model. We use the abbreviated form $\mathbb{P}(W_s)$ for the probability of encountering the specific update sequence W_s between times $(0, s)$.

Definition 3.7 (*Update support*)**.** Let W_s be an update sequence for the barrier-dynamics between times $(0, s)$. The *support* of W_s is the minimum subset $\Delta_{W_s} \subset \mathbb{Z}_m^d$ such that $\mathcal{G}_s(x)$ is a function of $x(\Delta_{W_s})$ for any x, i.e.,

$$g_{W_s}(x) = f_{W_s}(x(\Delta_{W_s})) \text{ for some } f_{W_s} : \{\pm 1\}^{\Delta_{W_s}} \to \{\pm 1\}^{\mathbb{Z}_m^d} \text{ and all } x.$$

In other words, $v \notin \Delta_{W_s}$ if and only if for every initial configuration x, modifying the spin at v does not affect the configuration $g_{W_s}(x)$. This definition uniquely defines the support of W_s.

Using this notion of the support of updates in the barrier-dynamics, we can now infer the following upper bound on the L^1-distance of the original dynamics to stationarity.

Lemma 3.8. *Let (X_t) be the Glauber dynamics on $\mathbb{Z}_m^d$, and W_s be the random update sequence for the barrier-dynamics along an interval $(0, s)$ for some $s \leq \log^{4/3} n$. For any x_0 and $t > 0$,*

$$\|\mathbb{P}_{x_0}(X_{t+s} \in \cdot) - \mu\|_{\mathrm{TV}} \leq \int \left\|\mathbb{P}_{x_0}(X_t(\Delta_{W_s}) \in \cdot) - \mu_{\Delta_{W_s}}\right\|_{\mathrm{TV}} d\mathbb{P}(W_s) + 2n^{-10d}.$$

Proof. Let X_t be the Glauber dynamics at time t started from $X_0 = x_0$, as usual let $\Omega_m = \{\pm 1\}^{\mathbb{Z}_m^d}$ and let $Y \in \Omega_m$ be distributed according to μ. Recalling that g_{W_s} denotes the deterministic function associated with an update sequence W_s for the barrier-dynamics in the interval $(0, s)$, for any random configuration $X \in \Omega_m$ we have

$$\begin{aligned}\|\mathbb{P}(\mathcal{G}_s(X) \in \cdot) - \mathbb{P}(\mathcal{G}_s(Y) \in \cdot)\|_{\mathrm{TV}} &= \max_{\Lambda \subset \Omega_m} [\mathbb{P}(\mathcal{G}_s(X) \in \Lambda) - \mathbb{P}(\mathcal{G}_s(Y) \in \Lambda)] \\ &= \max_{\Lambda \subset \Omega_m} \int [\mathbb{P}(g_{W_s}(X) \in \Lambda) - \mathbb{P}(g_{W_s}(Y) \in \Lambda)] \, d\,\mathbb{P}(W_s) \,.\end{aligned}$$

Since $g_{W_s}(X) = f_{W_s}(X(\Delta_{W_s}))$ by definition of Δ_{W_s}, the above is at most

$$\begin{aligned}&\int \max_{\Lambda \subset \Omega_m} \left[\mathbb{P}\big(f_{W_s}(X(\Delta_{W_s})) \in \Lambda\big) - \mathbb{P}\big(f_{W_s}(Y(\Delta_{W_s})) \in \Lambda\big)\right] \, d\,\mathbb{P}(W_s) \\ &\leq \int \big\|\mathbb{P}\big(X(\Delta_{W_s}) \in \cdot\big) - \mathbb{P}\big(Y(\Delta_{W_s}) \in \cdot\big)\big\|_{\mathrm{TV}} \, d\,\mathbb{P}(W_s) \\ &= \int \big\|\mathbb{P}\big(X(\Delta_{W_s}) \in \cdot\big) - \mu_{\Delta_{W_s}}\big\|_{\mathrm{TV}} \, d\,\mathbb{P}(W_s) \,,\end{aligned}$$

where the inequality in the second line used the fact that when taking a projection of two measures their total-variation can only decrease.

Since $s < \log^{4/3} n$, by Lemma 3.6 we can couple X_{t+s} and $\mathcal{G}_s(X_t)$ together except with probability n^{-10d} and hence

$$\|\mathbb{P}(X_{t+s} \in \cdot) - \mathbb{P}(\mathcal{G}_s(X_t) \in \cdot)\|_{\mathrm{TV}} \leq n^{-10d} \,.$$

Similarly, by Lemma 3.6 we can couple $\mathcal{G}_s(Y)$ with the Glauber dynamics run from Y for time s (having the stationary distribution μ) and hence

$$\|\mathbb{P}(\mathcal{G}_s(Y) \in \cdot) - \mu\|_{\mathrm{TV}} \leq n^{-10d} \,.$$

Combining these estimates, it follows that

$$\begin{aligned}\|\mathbb{P}(X_{t+s} \in \cdot) - \mu\|_{\mathrm{TV}} &\leq \|\mathbb{P}(\mathcal{G}_s(X_t) \in \cdot) - \mathbb{P}(\mathcal{G}_s(Y) \in \cdot)\|_{\mathrm{TV}} + 2n^{-10d} \\ &\leq \int \big\|\mathbb{P}\big(X_t(\Delta_{W_s}) \in \cdot\big) - \mu_{\Delta_{W_s}}\big\|_{\mathrm{TV}} \, d\,\mathbb{P}(W_s) + 2n^{-10d} \,,\end{aligned}$$

as required. ∎

Thus far, we have established an upper bound on the L^1-distance between X_{t+s} and μ in terms of $\Delta_{W_s} \subset \mathbb{Z}_m^d$, the support of the update sequence in the barrier-dynamics operator $\mathcal{G}_s$ along the interval $(0, s)$. We now wish to investigate the geometry of the set of sites comprising Δ_{W_s} for a typical update sequence W_s. The following lemma estimates the probability that Δ_{W_s} is sparse, as characterized in Definition 3.3. Figure 3 shows a realization of the update support becoming sparser with time.

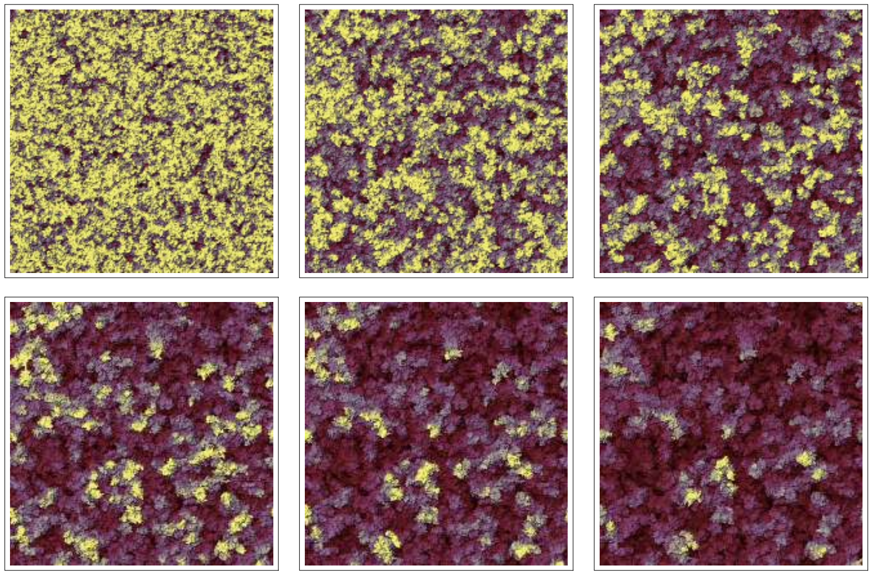

FIGURE 3. Evolution of the update support over increasing time intervals, simulated on a 500×500 square lattice at inverse-temperature $\beta = 0.4$. Highlighted regions correspond to the components comprising the sparse support.

Lemma 3.9. *Let $\mathcal{G}_s$ be the barrier-dynamics operator on $\mathbb{Z}_m^d$, let W_s be the update sequence up to time s for some $s \geq (10d/\alpha_{\mathsf{s}}^\star) \log\log n$, and $\mathcal{S}$ be the collection of sparse sets of $\mathbb{Z}_m^d$. Then $\mathbb{P}(\Delta_{W_s} \in \mathcal{S}) \geq 1 - n^{-10d}$ for any sufficiently large n.*

Proof. First, consider a single block B in the barrier-dynamics on $\mathbb{Z}_m^d$ and for simplicity let π stand for μ_{B^+}, the stationary distribution on B^+. Let E_B denote the event that $\Delta_{W_s} \cap B \neq \emptyset$ for a random update sequence W_s.

Observe that, by definition, the following holds for all s: If $\bar{B} \neq B$ are two distinct blocks, then modifying the value of a spin $u \in B$ in the initial configuration X_0 can only affect $\tilde{X}_s(\bar{B}^+)$ provided that $B, \bar{B}$ are adjacent (otherwise the barrier around $\bar{B}$ prevents the effect of this change). Hence, when assessing whether $u \in \Delta_{W_s}$ it suffices to consider the projection of $\tilde{X}_s$ onto the block B and all of its neighboring blocks; let $N(B)$ denote this set of 3^d blocks.

Let $(\tilde{X}_t^+)$ and $(\tilde{X}_t^-)$ be two instances of the barrier-dynamics restricted to $\bar{B}^+$ starting from the all-plus and all-minus states respectively, coupled via the monotone coupling (note that the restriction to one block turns these into the standard Glauber dynamics). Clearly, if these two chains coalesce under the update sequence W_s at some point $0 < t < s$ then either one of them predicts $\tilde{X}_s(\bar{B}^+)$ regardless of the value of $\tilde{X}_0(\bar{B}^+)$, and so

$$\mathbb{P}(E_B) \le \mathbb{P}\bigg(\bigcup_{\bar{B}\in N(B)} X_s^+(\bar{B}^+) \neq X_s^-(\bar{B}^+)\bigg) \le 3^d\, \mathbb{P}(\tilde{X}_s^+(B^+) \neq \tilde{X}_s^-(B^+))\,,$$

where the last inequality is by symmetry. As the system is monotone,

$$\mathbb{P}(E_B) \le 3^d\, \mathbb{P}\left(\tilde{X}_s^+(B^+) \neq \tilde{X}_s^-(B^+)\right) \le 3^d \sum_{u\in B^+} \mathbb{P}\left(\tilde{X}_s^+(u) \neq \tilde{X}_s^-(u)\right)\,. \tag{3.3}$$

Moreover, for any $u \in B^+$

$$\begin{aligned}\mathbb{P}\left(\tilde{X}_s^+(u) \neq \tilde{X}_s^-(u)\right) &\le \|\mathbb{P}(\tilde{X}_s^+ \in \cdot) - \pi\|_{\mathrm{TV}} + \|\mathbb{P}(\tilde{X}_s^- \in \cdot) - \pi\|_{\mathrm{TV}} \\ &\le \tfrac{1}{2}\|\mathbb{P}(\tilde{X}_s^+ \in \cdot) - \pi\|_{L^2(\pi)} + \tfrac{1}{2}\|\mathbb{P}(\tilde{X}_s^- \in \cdot) - \pi\|_{L^2(\pi)}\,,\end{aligned}$$

and by Theorem 2.1, if the all-plus state $\underline{1}$ has stationary measure at most e^{-1} (clearly the case for large n) then for any $s > 0$

$$\Big\|\mathbb{P}(\tilde{X}_s^+ \in \cdot) - \pi\Big\|_{L^2(\pi)} \le \exp\left(1 - \lambda\left(s - \frac{1}{4\alpha_{\mathsf{s}}}\log\log\frac{1}{\pi(\underline{1})}\right)\right),$$

where λ and α_{s} are the spectral gap and log-Sobolev constant resp. of the Glauber dynamics on B^+. Recalling that $\pi(\underline{1}) \ge 1/|\Omega_{B^+}| = 2^{-(1+o(1))\log^{2d} n}$ and that $\lambda \ge \alpha_{\mathsf{s}} \ge \alpha_{\mathsf{s}}^\star$, the assumption on s gives that

$$s \ge \frac{1}{4\alpha_{\mathsf{s}}^\star}\log\log(1/\pi(\underline{1})) + \frac{8d}{\lambda}\log\log n$$

for any sufficiently large n, and in this case

$$\|\mathbb{P}(\tilde{X}_s^+ \in \cdot) - \pi\|_{L^2(\pi)} \le 3(\log n)^{-8d}\,.$$

By the exact same argument we have

$$\|\mathbb{P}(\tilde{X}_s^- \in \cdot) - \pi\|_{L^2(\pi)} \le 3(\log n)^{-8d}\,,$$

and it now follows that

$$\sum_{u\in B^+} \mathbb{P}\left(\tilde{X}_s^+(u) \neq \tilde{X}_s^-(u)\right)\Big) \le (3+o(1))(\log n)^{-6d} < 4(\log n)^{-6d}\,,$$

where the last inequality holds for large n. Combining this with (3.3) yields that for any large n, the following holds with room to spare:

$$\mathbb{P}(E_B) \le (\tfrac{1}{2}\log n)^{-6d}\,.$$

This estimate will now readily imply a bound on the number of components in the support of W_s. Let $E^\sharp$ denote the following event: There exists a collection $\mathcal{B}$ of $L \ge (n/\log^7 n)^d$ blocks, such that E_B holds (that is, $\Delta_{W_s} \cap B \neq \emptyset$) for all $B \in \mathcal{B}$, and the pairwise distances in blocks between the blocks in $\mathcal{B}$ are all at least 4. We claim that $\mathbb{P}(E^\sharp) \le n^{-20d}$ for large n.

To see this, first notice that if a block B has a distance of at least 4 blocks from a set of blocks $\mathcal{B}$ (i.e., for any $B' \in \mathcal{B}'$, no two blocks in $N(B)$ and $N(B')$ are adjacent), then the variable $\mathbb{1}_{E_B}$ is independent of $\{\mathbb{1}_{E_{B'}} : B \in \mathcal{B}\}$. Indeed, the initial configuration on B can only affect the outcome of $\tilde{X}_s(N(B))$, and by definition this outcome is derived from the initial configuration via the updates of the sites $U_B = \cup_{\bar{B} \in N(B)} \bar{B}^+$. Our assumption on the distance between B and $\mathcal{B}$ precisely implies that the above U_B is disjoint to any $U_{B'}$ for $B' \neq B$ in $\mathcal{B}$, and the statement now follows from the independence of updates to distinct sites. Hence, as the total number of blocks is $(1+o(1))(n/\log^2 n)^d$, for large n we have

$$\mathbb{P}(E^\sharp) \leq \binom{(2n/\log^2 n)^d}{(n/\log^7 n)^d} \left(\tfrac{1}{2}\log n\right)^{-6d(n/\log^7 n)^d} < n^{-(n/\log^8 n)^d} .$$

Now suppose that there is a sequence of blocks, $(B_{i_0}, B_{i_1}, B_{i_2}, \ldots, B_{i_\ell})$ for some $\ell \geq \ell_0 = \frac{1}{3d}\log n$, such that for all k,

(1) The distance in blocks between $B_{i_{k-1}}, B_{i_k}$ is at most $3d$ (i.e., they are the endpoints of a path of at most $3d+1$ adjacent blocks).
(2) We have $\Delta_{W_s} \cap B_{i_k} \neq \emptyset$.

Clearly, if the distance in blocks between some B, B' is at least $3d+1$, then since every block has side-length at least $\log^2 n - 1$, the distance between any two sites $u \in B$ and $v \in B'$ is at least $2d\log^2 n$ for sufficiently large n.

Observe that if $\Delta_{W_s} \notin \mathcal{S}$, then either the event $E^\sharp$ holds or a sequence of blocks as described above must exist. Indeed, consider some Δ_{W_s} that is not sparse, and partition it into components as follows:

$$\begin{aligned} &u \in B \cap \Delta_{W_s} \text{ and } u' \in B' \cap \Delta_{W_s} \text{ belong to the same component} \\ \iff \quad &\text{the distance in blocks between } B, B' \text{ is at most } 3d. \end{aligned}$$

The number of components is clearly at most $(n/\log^7 n)^d$, otherwise the event $E^\sharp$ holds. Furthermore, by definition (as argued above), the distance between any two distinct components is at least $2d\log^2 n$. Hence, the assumption that $\Delta_{W_s} \notin \mathcal{S}$ implies that some component A_i must have a diameter larger than $\log^3 n$. In particular, there are two sites $u, u' \in A_i$ belonging to B, B' respectively, such that B, B' have distance of at least $\log n$ blocks between them. Moreover, by the way we defined the component A_i there are blocks $B = B_{j_0}, \ldots, B_{j_\ell} = B'$ such that $B_{j_i} \cap \Delta_{W_s} \neq \emptyset$ and the distance in blocks between $B_{j_i}, B_{j_{i+1}}$ is at most $3d$. As the assumption on u, u' implies that $\ell \geq \frac{1}{3d}\log n$, this sequence satisfies the required properties.

It therefore remains to show that, except with probability n^{-20d}, there does not exist a sequence of blocks satisfying the above properties (1),(2).

The above described sequence in particular contains a set $\mathcal{B}$ of at least $\ell/7^d$ blocks, whose pairwise distances are all at least 4, and $\Delta_{W_s} \cap B \neq \emptyset$

for all $B \in \mathcal{B}$ (for instance, take $\mathcal{B} = \{B_{i_k}\}$, process its blocks sequentially, and for each B delete from $\mathcal{B}$ any $B' \neq B$ whose distance from B is less than 4, a total of at most $7^d - 1$ blocks). By the above discussion on the independence of the events E_B for such blocks, as well as our estimate on $\mathbb{P}(E_B)$, we deduce that for any large n

$$\mathbb{P}(\Delta_{W_s} \cap B \neq \emptyset \text{ for all } B \in \mathcal{B}) \leq \left(\tfrac{1}{2}\log n\right)^{-6d\ell/7^d} .$$

Clearly, there are at most $n^d(6d+1)^{d\ell}$ sequences of blocks $\{B_{i_0}, B_{i_1}, \ldots, B_{i_\ell}\}$, where any two consecutive blocks are of distance (in blocks) at most $3d$. Hence, the probability that there exists a sequence with the aforementioned properties (1),(2) is at most

$$n^d \sum_{\ell \geq \ell_0} \left((6d+1)^{7d}(\tfrac{1}{2}\log n)^{-6}\right)^{\ell d/7^d} < n^d(\log n)^{-(5d/7^d)\ell_0} < n^{d-7^{-d}\log\log n}$$

for any large n, as required. ∎

Theorem 3.4 is now obtained as an immediate corollary of Lemma 3.8 and Lemma 3.9, with ν given by $\nu(\Delta) = \mathbb{P}(\{W_s : \Delta_{W_s} = \Delta\})$ for $\Delta \in \mathcal{S}$. ∎

3.2. **Proof of Theorem 3.1, Part 1: upper bound on the L^1 distance.** Let $\Delta \subset \mathbb{Z}_m^d$ be a sparse set and $\cup_{i=1}^L A_i$ be its partition to components as per Definition 3.3, where $L \leq m^d \wedge (n/\log^5 n)^d$. Letting $\mathrm{dist}(\cdot,\cdot)$ denote the distance according to the lattice metric of $\mathbb{Z}_m^d$, put

$$A_i^+ = \{v : \mathrm{dist}(A_i, v) < \log^{3/2} n\}\,,$$

and (recalling that the diameter of A_i is at most $\log^3 n$) let ψ_i be an isometry of A_i^+ into a box B_i of side-length $2\log^3 n$ inside $\mathbb{Z}_r^d$, the torus of side-length $r = 3\log^3 n$. A crucial point to notice is that the sets $\{A_i^+\}$ are pairwise disjoint, since the distance between distinct A_i, A_j is at least $2d\log^2 n$.

Let (X_t^*) denote the product chain of the Glauber dynamics run on each of these L copies of $\mathbb{Z}_r^d$ independently, and let μ^* denote its stationary measure. Given some initial configuration x_0 on $\mathbb{Z}_m^d$, we construct x_0^* for (X_t^*) in the obvious way: For each $v \in A_i^+$, assign $x_0^*(\psi_i(v)) = x_0(v)$, while the configuration of each $\mathbb{Z}_r^d \setminus \psi_i(A_i^+)$ can be chosen arbitrarily (e.g., all-plus).

Let $t \leq \log^{4/3} n$. We claim that we can now couple (X_s) with (X_s^*) such that $X_s(\Delta) = X_s^*(\cup\psi_i(A_i))$ for all $0 \leq s \leq t$ except with probability n^{-10d}.

Indeed, by repeating the argument of Lemma 3.6, if we let (X_s^*) use the same updates of (X_s) on $\cup A_i^+$ and run independent updates elsewhere, then each A_i can be coupled to $\psi_i(A_i)$ on $[0,t]$ except with probability n^{-11d} (here too the distance between A_i and ∂A_i^+ is at least $\log^{3/2} n$). Since the subsets $\{A_i^+\}$ are disjoint, distinct $\psi_i(A_i)$ indeed obtain independent updates in this manner. Summing the above error probability over the L copies we deduce that the coupling exists except with probability $Ln^{-11d} \leq n^{-10d}$.

We claim that by the strong spatial mixing property, for any large n

$$\|\mu_\Delta - \mu^*_{(\cup\psi_i(A_i))}\|_{\mathrm{TV}} < n^{-9d} \tag{3.4}$$

after one translates the sites in the obvious manner (that is, when comparing the two measures we apply the isometries ψ_i on the inputs).

It is easy to infer the above statement under the assumption that there is strong spatial mixing $\mathtt{SM}(\Lambda, c_1, c_2)$ for any subset Λ (so-called complete analyticity, as opposed to our assumption that strong spatial mixing $\mathtt{SM}(Q, c_1, c_2)$ holds for sufficiently large cubes). Indeed, recall that for some fixed $c_1, c_2 > 0$ we have $\|\mu^\tau_\Delta - \mu^{\tau^y}_\Delta\|_{\mathrm{TV}} \leq c_1 \exp(-c_2 \operatorname{dist}(\Delta, y))$ for all $\tau \in \{\pm 1\}^{\partial\Lambda}$ and $y \in \partial\Lambda$. By taking the boundary conditions τ_1 and τ_2 to be distributed according to μ and μ^* about $\cup\partial A_i^+$ and $\cup\partial\psi_i(A_i^+)$ respectively, we can transform τ_1 into τ_2 via a series of at most m^d spin flips. Each flip amounts to an error of at most $c_1 \exp(-c_2 \log^{3/2} n)$, leading to (3.4).

Since we only have regular complete analyticity at hand, we will obtain (3.4) via monotonicity and log-Sobolev inequalities. Take $t = (\log n)^{4/3}$ and apply Theorems 2.1 and 2.3 to get

$$\begin{aligned}
\|\mathbb{P}_{x_0}(X_t \in \cdot) - \mu\|_{\mathrm{TV}} &\leq \|\mathbb{P}_{x_0}(X_t \in \cdot) - \mu\|_{L^2(\mu)} \\
&\leq \exp\left(1 - \lambda\left(t - \frac{1}{4\alpha_{\mathtt{s}}} \log\log \mu^*(x_0)\right)\right),
\end{aligned}$$

where $\lambda, \alpha_{\mathtt{s}}$ are the spectral gap and log-Sobolev constant of (X_t) respectively. Since $\log\log(1/\mu^*(\sigma)) \geq (d + o(1)) \log n$ for all σ and since it holds that $\lambda \geq \alpha_{\mathtt{s}} \geq \alpha^\star_{\mathtt{s}} > 0$, for a suitably large n we have

$$\max_{x_0} \|\mathbb{P}_{x_0}(X_t \in \cdot) - \mu\|_{\mathrm{TV}} \leq n^{-10d}. \tag{3.5}$$

It is well known (see, e.g., [33, Lemma 2.2.11] and also [3, Chapter 8]) that the spectral gap (respectively log-Sobolev constant) of a product chain is equal to the minimum of the spectral gaps (log-Sobolev constants). In particular, the log-Sobolev constant of (X^*_t) is at least $\alpha^\star_{\mathtt{s}}$, hence similarly

$$\max_{x^*_0} \left\|\mathbb{P}_{x^*_0}(X^*_t \in \cdot) - \mu^*\right\|_{\mathrm{TV}} \leq n^{-10d}. \tag{3.6}$$

Therefore, if Y is a configuration on Δ distributed according to μ_Δ then

$$\begin{aligned}
\left\|\mu_\Delta - \mu^*_{(\cup\psi_i(A_i))}\right\|_{\mathrm{TV}} &\leq \|\mathbb{P}(\psi(Y) \in \cdot) - \mathbb{P}_{x_0}(\psi(X_t(\Delta)) \in \cdot)\|_{\mathrm{TV}} \\
&\quad + \left\|\mathbb{P}_{x_0}(\psi(X_t(\Delta)) \in \cdot) - \mathbb{P}_{x^*_0}(X^*_t(\cup\psi_i(A_i)) \in \cdot)\right\|_{\mathrm{TV}} \\
&\quad + \left\|\mathbb{P}_{x^*_0}(X^*_t(\cup\psi_i(A_i)) \in \cdot) - \mu^*_{(\cup\psi_i(A_i))}\right\|_{\mathrm{TV}} < n^{-9d},
\end{aligned}$$

where the last inequality holds for large n, combining (3.10), (3.5) and (3.6) with the fact that projections can only reduce the total variation distance.

Altogether, letting $\Gamma = \cup B_i$ and abbreviating $\mathbb{P}_{x_0^*}(X_t^* \in \cdot)$ by π^*, we have

$$\begin{aligned}\|\mathbb{P}_{x_0}\left(X_t(\Delta) \in \cdot\right) - \mu_\Delta\|_{\mathrm{TV}} &\leq 2n^{-9d} + \left\|\mathbb{P}_{x_0^*}\left(X_t^*(\cup\psi_i(A_i)) \in \cdot\right) - \mu^*_{(\cup\psi_i(A_i))}\right\|_{\mathrm{TV}} \\ &\leq 2n^{-9d} + \|\pi^*_\Gamma - \mu^*_\Gamma\|_{\mathrm{TV}}\,. \end{aligned} \tag{3.7}$$

We next seek an upper bound for the last expression, uniformly over the initial configuration x_0^*.

$$\begin{aligned}\|\pi^*_\Gamma - \mu^*_\Gamma\|_{\mathrm{TV}} &= \frac{1}{2}\sum_{b_1,\ldots,b_L} |\pi^*_\Gamma(b_1,\ldots,b_L) - \mu^*_\Gamma(b_1,\ldots,b_L)| = \frac{1}{2}\mathbb{E}_{\mu^*_\Gamma}\left|\frac{\pi^*_\Gamma}{\mu^*_\Gamma} - 1\right| \\ &\leq \frac{1}{2}\left(\mathbb{E}_{\mu^*_\Gamma}\left|\frac{\pi^*_\Gamma}{\mu^*_\Gamma} - 1\right|^2\right)^{1/2} = \frac{1}{2}\left(\mathbb{E}_{\mu^*_\Gamma}\left|\frac{\pi^*_\Gamma}{\mu^*_\Gamma}\right|^2 - 1\right)^{1/2}, \end{aligned} \tag{3.8}$$

where the inequality was by Cauchy-Schwartz, and the last equality is due to the fact that $\mathbb{E}_{\mu^*_\Gamma}[\pi^*_\Gamma/\mu^*_\Gamma] = 1$. Now, since (X_t^*) is a product of L independent instances of Glauber dynamics on $\mathbb{Z}_r^d$, we infer that

$$\mathbb{E}_{\mu^*_\Gamma}\left|\frac{\pi^*_\Gamma}{\mu^*_\Gamma}\right|^2 = \mathbb{E}_{\mu^*_\Gamma}\prod_{i=1}^L \left|\frac{\pi^*_{B_i}}{\mu^*_{B_i}}\right|^2 = \prod_{i=1}^L \left(\|\pi^*_{B_i} - \mu^*_{B_i}\|^2_{L^2(\mu^*_{B_i})} + 1\right),$$

and recalling the definition of $\mathfrak{m}_t$ in (3.1), it follows that $\mathbb{E}_{\mu^*_\Gamma}|\pi^*_\Gamma/\mu^*_\Gamma|^2$ is at most $(\mathfrak{m}_t^2 + 1)^L$. Plugging this in (3.8),

$$\begin{aligned}\max_{x_0^*}\|\mathbb{P}_{x_0}\left(X_t^*(\Gamma) \in \cdot\right) - \mu^*_\Gamma\|_{\mathrm{TV}} &\leq \frac{1}{2}\left(\left(\mathfrak{m}_t^2 + 1\right)^L - 1\right)^{1/2} \\ &\leq \frac{1}{2}\left(\exp\left(L\,\mathfrak{m}_t^2\right) - 1\right)^{1/2}.\end{aligned}$$

Altogether, using (3.7), we conclude that for any $\Delta \in \mathcal{S}$ and $t < \log^{4/3} n$,

$$\max_{x_0}\|\mathbb{P}_{x_0}\left(X_t(\Delta) \in \cdot\right) - \mu_\Delta\|_{\mathrm{TV}} \leq \frac{1}{2}\left(\exp\left(L\,\mathfrak{m}_t^2\right) - 1\right)^{1/2} + 2n^{-9d}.$$

At this point, Theorem 3.4 implies that for any s,t with $0 < t < \log^{4/3} n$ and $(10d/\alpha^\star_{\mathsf{s}})\log\log n \leq s \leq \log^{4/3} n$ and for every large n,

$$\|\mathbb{P}_{x_0}(X_{t+s} \in \cdot) - \mu\|_{\mathrm{TV}} \leq \frac{1}{2}\left(\exp\left(L\,\mathfrak{m}_t^2\right) - 1\right)^{1/2} + 3n^{-9d}. \tag{3.9}$$

Recalling that $L \leq m^d \wedge (n/\log^5 n)^d$ concludes Part 1 of Theorem 3.1 and also establishes the inequality in Remark 3.2. ■

3.3. **Proof of Theorem 3.1, Part 2: lower bound on the L^1 distance.** Let

$$r = 3\log^3 n\,, \qquad L = \lfloor n/r \rfloor^d\,,$$

and let $A_1, \ldots, A_L \subset \mathbb{Z}_n^d$ be a collection of d-dimensional boxes of side-length $\frac{2}{3}r$ satisfying $\|u - v\|_\infty > r/3$ for any $u \in A_i$ and $v \in A_j$ with $i \neq j$. Similar to the notation of the previous subsection, for each $i \in \{1, \ldots, L\}$ we define

$$A_i^+ = \{v : \mathrm{dist}(A_i, v) \leq r/6\}\,.$$

Denote the unions of these boxes by $\Delta = \cup_{i=1}^L A_i$ and $\Delta^+ = \cup_{i=1}^L A_i^+$.

Let $B_1^+, \ldots, B_L^+$ be a sequence of disconnected d-dimensional boxes of side-length r and let Γ^+ denote the graph of their union. Let ψ_i be an isometry mapping A_i^+ to B_i^+ and let ψ be the isometry that maps Δ^+ to Γ^+ such that its restriction to any individual A_i^+ is ψ_i. We define $B_i = \psi(A_i)$ and $\Gamma = \cup_{i=1}^L B_i$. For a configuration X on some $\Delta' \subseteq \Delta^+$ we let will denote $\psi(X)$ as the corresponding configuration on $\psi(\Delta')$.

We couple the Glauber dynamics on $\mathbb{Z}_n^d$ and Γ^+ as follows: Whenever a site $u \in \Delta^+$ receives an update via some unit variable I we also update the site $\psi(u)$ using the same I and a periodic boundary condition on its corresponding box B_i^+. Denote the dynamics induced on Γ^+ as X_t^*, and let μ^* be its stationary distribution. The above defined coupling satisfies the following claim.

Claim 3.10. *Let (X_t) and (X_t^*) be the above coupled Glauber dynamics on $\mathbb{Z}_n^d$ and Γ^+ respectively. Suppose X_0, X_0^* satisfy $\psi(X_0(\Delta^+)) = X_0^*(\Gamma^+)$. Then with probability at least $1 - n^{-10d}$, for all $0 \leq s \leq (\log n)^{4/3}$ we have*

$$\psi(X_s(\Delta)) = X_s^*(\Gamma)\,. \tag{3.10}$$

Furthermore,

$$\max_{x_0} \|\mathbb{P}_{x_0}(X_s \in \cdot) - \mu\|_{\mathrm{TV}} \geq \max_{x_0^*} \left\|\mathbb{P}_{x_0^*}(X_s^*(\Gamma) \in \cdot) - \mu_\Gamma^*\right\|_{\mathrm{TV}} - 4n^{-10d}\,.$$

Proof. Equation (3.10) holds by a simple adaption of the proof of Lemma 3.6 as the initial conditions and updates agree on A_i^+ and B_i^+ for all i, and since each B_i is distance $r/6 = \frac{1}{2}\log^3 n$ from the boundary of B_i^+.

For the second statement of the claim, we repeat the argument that yielded inequality (3.4) in the previous subsection. Take $t = (\log n)^{4/3}$ and apply Theorems 2.1 and 2.3 to get

$$\|\mathbb{P}_{x_0}(X_t \in \cdot) - \mu\|_{\mathrm{TV}} \leq \exp\left(1 - \lambda\left(t - \frac{1}{4\alpha_{\mathsf{s}}}\log\log\mu^*(x_0)\right)\right)\,,$$

where λ and $\alpha_{\mathtt{s}}$ are the spectral gap and log-Sobolev constant of (X_t) respectively. This yields the following for any sufficiently large n:

$$\max_{x_0} \|\mathbb{P}_{x_0}(X_t \in \cdot) - \mu\|_{\mathrm{TV}} \leq n^{-10d} . \tag{3.11}$$

Since the log-Sobolev constant of the dynamics on Γ^+ is at least $\alpha_{\mathtt{s}}^\star$,

$$\max_{x_0^*} \left\|\mathbb{P}_{x_0^*}(X_t^* \in \cdot) - \mu^*\right\|_{\mathrm{TV}} \leq n^{-10d} . \tag{3.12}$$

Now, if $Y \in \{\pm 1\}^\Delta$ is distributed according to μ_Δ then

$$\begin{aligned}\|\mathbb{P}(\psi(Y) \in \cdot) - \mu^*_\Gamma\|_{\mathrm{TV}} \leq{}& \|\mathbb{P}(\psi(Y) \in \cdot) - \mathbb{P}_{x_0}(\psi(X_t(\Delta)) \in \cdot)\|_{\mathrm{TV}} \\ &+ \left\|\mathbb{P}_{x_0}(\psi(X_t(\Delta)) \in \cdot) - \mathbb{P}_{x_0^*}(X_t^*(\Gamma) \in \cdot)\right\|_{\mathrm{TV}} \\ &+ \left\|\mathbb{P}_{x_0^*}(X_t^*(\Gamma) \in \cdot) - \mu^*_\Gamma\right\|_{\mathrm{TV}} \leq 3n^{-10d} ,\end{aligned}$$

where the last inequality added (3.10), (3.11) and (3.12) to the fact that projections can only reduce the total variation distance. It follows that

$$\begin{aligned}\max_{x_0} \|\mathbb{P}_{x_0}(X_s \in \cdot) - \mu\|_{\mathrm{TV}} &\geq \max_{x_0} \|\mathbb{P}_{x_0}(\psi(X_s(\Delta)) \in \cdot) - \mathbb{P}(\psi(Y) \in \cdot)\|_{\mathrm{TV}} \\ &\geq \max_{x_0^*} \left\|\mathbb{P}_{x_0^*}(X_s^*(\Gamma) \in \cdot) - \mu^*_\Gamma\right\|_{\mathrm{TV}} - 4n^{-10d} ,\end{aligned}$$

where we used the facts that $\|\mathbb{P}(\psi(Y) \in \cdot) - \mu^*_\Gamma\|_{\mathrm{TV}} \leq 3n^{-10d}$ and that $\|\mathbb{P}(\psi(X_s(\Delta)) \in \cdot) - \mathbb{P}_{x_0^*}(X_s^*(\Gamma) \in \cdot)\|_{\mathrm{TV}} \leq n^{-10d}$. ∎

Recall that the B_i boxes have side-length $\frac{2}{3}r = 2\log^3 n$, matching the boxes B in the definition (3.1) of $\mathfrak{m}_t$. Let $x_0^* = x_0^*(t)$ be a configuration on a box B of side-length $3\log^3 n$ which achieves $\mathfrak{m}_t$, i.e.

$$\mathfrak{m}_t = \left\|\mathbb{P}_{x_0^*}\big(X_t^*(B) \in \cdot\big) - \mu^*_B\right\|_{L^2(\mu^*_B)} .$$

We define i.i.d. random variables

$$Y_i = \frac{\mathbb{P}\left(X_t^*(B_i) = U_i \mid X_0^*(B_i^+) = x_0^*\right)}{\mu^*_{B_i}(U_i)} , \tag{3.13}$$

where the U_i are i.i.d. configurations on B_i distributed according to $\mu^*_{B_i}$. As the dynamics on different tori are independent it follows that the Y_i are independent. As we will soon show, these random variables provide crucial insight into the mixing of (X_t) in the L^1-distance. First we need to obtain some estimates on their moments. Clearly,

$$\mathbb{E}Y_i = \sum_{b_i} \frac{\mathbb{P}\left(X_t^*(B_i) = b_i \mid X_0^*(B_i^+) = x_0^*\right)}{\mu^*_{B_i}(b_i)} \mu^*_{B_i}(b_i) = 1 ,$$

and

$$\operatorname{Var} Y_i = \sum_{b_i} \left| \frac{\mathbb{P}\left(X_t^*(B_i) = b_i \mid X_0^*(B_i^+) = x_0^*\right)}{\mu^*_{B_i}(b_i)} - 1 \right|^2 \mu^*_{B_i}(b_i) = \mathfrak{m}_t^2 .$$

Moreover, by a standard L^∞ to L^2 reduction (cf., e.g., [33]),

$$\begin{aligned}\|Y_i - 1\|_\infty &= \|\mathbb{P}\left(X_t^*(B_i) \in \cdot \mid X_0^*(B_i^+) = x_0^*\right) - \mu^*_{B_i}\|_{L^\infty(\mu^*_{B_i})} \\ &\leq \|\mathbb{P}\left(X_{t/2}^*(B_i) \in \cdot \mid X_0^*(B_i^+) = x_0^*\right) - \mu^*_{B_i}\|^2_{L^2(\mu^*_{B_i})}\,,\end{aligned}$$

and hence by Theorems 2.1 and 2.3 this is at most $\mathrm{e}^{-c\log\log n}$ for some absolute constant $c > 0$, yielding

$$\mathbb{E}|Y_i - 1|^3 \leq \|Y_i - 1\|_\infty \operatorname{Var} Y_i \leq \mathrm{e}^{-c\log\log n}\mathfrak{m}_t^2 = o(\mathfrak{m}_t^2)\,.$$

Define $Z_i = \log Y_i$. Taking Taylor series expansions gives

$$\mathbb{E}Z_i = \mathbb{E}(Y_i - 1) - \tfrac{1}{2}\mathbb{E}(Y_i - 1)^2 + O\left(\mathbb{E}|Y_i - 1|^3\right) = -\tfrac{1-o(1)}{2}\mathfrak{m}_t^2\,,$$

and similarly,

$$\mathbb{E}Z_i^2 = \mathbb{E}(Y_i - 1)^2 + O\left(\mathbb{E}|Y_i - 1|^3\right) = (1 + o(1))\mathfrak{m}_t^2\,.$$

We now derive the required lower bound on the L^1-distance of (X_t) from stationarity at time $t \geq (20d/\alpha^\star_{\mathfrak{s}})\log\log n$, provided that $\lim_{n\to\infty} L\mathfrak{m}_t^2 = \infty$. Consider a starting configuration X_0 such that $\psi(X_0(A_i^+)) = x_0^*$ for all i, and similarly choose X_0^* so that $X_0^*(B_i^+) = x_0^*$ for all i. By Claim 3.10 it is sufficient to show that under our hypothesis

$$\|\mathbb{P}_{x_0}(X_t^*(\Gamma) \in \cdot) - \mu^*_\Gamma\|_{\mathrm{TV}} \to 1\,.$$

By the definition Y_i's,

$$\begin{aligned}&\|\mathbb{P}(X_t^*(\Gamma) \in \cdot) - \mu^*_\Gamma\|_{\mathrm{TV}} \\ &\quad= \frac{1}{2}\sum_{b_1,\ldots,b_L} \left|\mathbb{P}_{x_0}\left(X_t^*(B_1) = b_1, \ldots, X_t^*(B_L) = b_L\right) - \mu^*_\Gamma\left(b_1, \ldots, b_L\right)\right| \\ &\quad= \frac{1}{2}\sum_{b_1,\ldots,b_L} \left|\prod_{i=1}^L \frac{\mathbb{P}_{x_0}(X_t^*(B_i) = b_i)}{\mu^*_{B_i}(b_i)} - 1\right| \prod_{i=1}^L \mu^*_{B_i}(b_i) \\ &\quad= \frac{1}{2}\mathbb{E}\left|\prod_{i=1}^L Y_i - 1\right| = \mathbb{E}\left|\exp\left(\sum_{i=1}^L Z_i\right) - 1\right|^-\,,\end{aligned}$$

where $|a|^-$ denotes $\max\{-a, 0\}$. The random variables Z_i are independent, $\|Z_i\|_\infty = o(1)$ and $\sum_{i=1}^L \operatorname{Var} Z_i \to \infty$, hence the Central Limit Theorem implies that

$$\lim_{n\to\infty} \frac{\sum_{i=1}^L \left(Z_i - \mathbb{E}Z_i\right)}{\sqrt{L\mathfrak{m}_t^2}} = \mathcal{N}(0,1)\,.$$

Since $L\mathbb{E}Z_1/\sqrt{L\mathfrak{m}_t^2} = -\frac{1-o(1)}{2}\sqrt{L\mathfrak{m}_t^2} \to -\infty$, it then follows that $\sum_{i=1}^L Z_i$ converges in probability to $-\infty$ and therefore

$$\mathbb{E}\left|\exp\left(\sum_{i=1}^{L} Z_i\right) - 1\right|^{-} \to 1\,,$$

as required. ∎

4. Cutoff for the Ising model

In this section we prove the main results, Theorems 1 and 2. We first describe how the L^1-L^2 reduction from the previous section (Theorem 3.1) establishes the existence of cutoff. A refined analysis of the L^2 distance to stationarity (captured by the quantity $\mathfrak{m}_t$ in the aforementioned theorem) via log-Sobolev inequalities then allows us to pinpoint the precise location of the cutoff in terms of spectral gaps of the dynamics on tori of prescribed sizes. Finally, by applying this tool on tori of varying sizes we obtain as a biproduct that, as the tori side-length tends to infinity, these spectral gaps tend to λ_∞, the spectral gap of the dynamics on the infinite-volume lattice. In turn, this yields the asymptotics of the mixing time in terms of λ_∞ and establishes Theorem 4.

4.1. **Existence of cutoff.** Theorem 3.1 already establishes cutoff for the Ising model on $(\mathbb{Z}/n\mathbb{Z})^d$, although not its precise location. To see this, recall the definition of $\mathfrak{m}_t$ given in (3.1) and choose t^* as follows

$$t^* = \inf\left\{t : \mathfrak{m}_t^2 \leq \frac{\log^{3d+1}}{n^d}\right\}.$$

As before, the log-Sobolev inequalities of Theorems 2.1 and 2.3 imply that $t^* = O(\log n)$. Let $s = (10d/\alpha_{\mathtt{s}}^\star)\log\log n$. Since $\mathfrak{m}_t$ is a continuous function, we have $(n/\log^5 n)^d \mathfrak{m}_{t^*}^2 = \log^{1-2d} n = o(1)$, and so by Part 1 of Theorem 3.1

$$\max_{x_0} \|\mathbb{P}_{x_0}(X_{t^*+s} \in \cdot) - \mu\|_{\mathrm{TV}} \leq \frac{1}{2}\Big(\exp\big(\log^{1-2d} n\big) - 1\Big)^{1/2} + 6n^{-9d} = o(1)\,.$$

Next, the results of [18] imply that the L^1 mixing time of the Glauber dynamics for the Ising model on $(\mathbb{Z}/n\mathbb{Z})^d$ has order at least $\log n$, hence (by the above inequality) $t^*(n)$ is also of order at least $\log n$. In particular, $t^* \geq (20d/\alpha_{\mathtt{s}})\log\log n$ for any sufficiently large n, and since by definition $(n/\log^3 n)^d \mathfrak{m}_{t^*}^2 = \log n$, it follows from Part 2 of Theorem 3.1 that

$$\max_{x_0} \|\mathbb{P}_{x_0}(X_{t^*} \in \cdot) - \mu\|_{\mathrm{TV}} = 1 - o(1)\,.$$

This establishes cutoff at t^* with a window of $O(\log\log n)$. ∎

4.2. **Cutoff location (asymptotics of the mixing time).** To obtain the asymptotics of the mixing time, it remains to estimate the parameter t^* introduced above, that is, to understand the threshold $t(n)$ for $n^d \mathfrak{m}_t^2$ to tend to infinity faster than some poly-logarithmic function of n.

In what follows, let $\lambda(r), \alpha_{\mathtt{s}}(r)$ be the spectral gap and log-Sobolev constant of the Glauber dynamics on a d-dimensional torus of side-length r.

Lemma 4.1. *Set* $c_0 = \frac{12d}{\alpha_{\mathtt{s}}^\star}$, *let* $\frac{20d}{\alpha_{\mathtt{s}}^\star \lambda^\star} \log\log n \leq t \leq \log^{4/3} n$ *and* $r = 3\log^3 n$. *For n sufficiently large,*

$$\mathrm{e}^{-\lambda(r)t - c_0 \log\log n} - n^{-9d} \leq \mathfrak{m}_t \leq \mathrm{e}^{-\lambda(r)t + c_0 \log\log n} . \tag{4.1}$$

Proof. Let X_t^* denote the Glauber dynamics on $\mathbb{Z}_r^d$ with periodic boundary conditions, let μ_r^* be its stationary distribution and let Ω_r^* denote its state space. Since $\log\log(1/\mu_r^*(\sigma)) \geq (3d + o(1)) \log\log n$ for all $\sigma \in \Omega_r^*$ and since $\lambda(r) \leq 1$ (vertices are updated at rate 1), another application of Theorem 2.1 implies that for large n

$$\begin{aligned}\mathfrak{m}_t &\leq \max_{\sigma \in \Omega_r^*} \exp\left(1 - \lambda(r)\left(t - \frac{1}{4\alpha_{\mathtt{s}}(r)} \log\log\left(1/\mu_r^*(\sigma)\right)\right)\right) \\ &\leq \mathrm{e}^{-\lambda(r)t + \frac{3d+o(1)}{4\alpha_{\mathtt{s}}^\star} \log\log n} \leq \mathrm{e}^{-\lambda(r)t + c_0 \log\log n} . \end{aligned} \tag{4.2}$$

This establishes the upper bound on $\mathfrak{m}_t$. Further, as $t \geq (20d/\alpha_{\mathtt{s}}^\star \lambda^\star) \log\log n$ it follows that $r^{d/2} \mathfrak{m}_t \leq \log^{-6d} n = o(1)$.

A standard lower bound on the total variation distance in terms of the spectral gap (cf. its discrete-time analogue [23, equation (12.13)]) gives that

$$\mathrm{e}^{-\lambda(r)t} \leq 2 \left\| \mathbb{P}(X_t^* \in \cdot) - \mu^* \right\|_{\mathrm{TV}} \text{ for all } t > 0 .$$

Set $s = (10d/\alpha_{\mathtt{s}}^\star) \log\log n$. Applying Part 1 of Theorem 3.1 to $\mathbb{Z}_r^d$ with these s, t (recalling Remark 3.2 and plugging in $m = r$ in (3.2)) gives

$$\begin{aligned}\mathrm{e}^{-\lambda(r)(t+s)} &\leq 2 \left\| \mathbb{P}(X_{t+s}^* \in \cdot) - \mu^* \right\|_{\mathrm{TV}} \\ &\leq \left(\exp\left(r^d \mathfrak{m}_t^2\right) - 1\right)^{1/2} + 6n^{-9d} \leq 2r^{d/2} \mathfrak{m}_t + 6n^{-9d} , \end{aligned} \tag{4.3}$$

where the last inequality used the fact that for $x < 1$ we have $\mathrm{e}^x - 1 \leq 2x$ and $r^d \mathfrak{m}_t^2 = o(1)$. Rearranging equation (4.3) we have that

$$\mathfrak{m}_t \geq \mathrm{e}^{-\lambda(r)(t+s) - \log(2r^{d/2})} - \frac{3n^{-9d}}{r^{d/2}} \geq \mathrm{e}^{-\lambda(r)t - c_0 \log\log n} - n^{-9d} \tag{4.4}$$

Combining equations (4.2) and (4.3) completes the proof. ∎

The following theorem now establishes the position of the mixing time in terms of $\lambda(r)$ with a window of $O(\log\log n)$.

Theorem 4.2. *Let* (X_t) *the Glauber dynamics on the* $(\mathbb{Z}/n\mathbb{Z})^d$*, and set*

$$t^* = t^*(n) = \frac{d}{2\lambda(r)} \log n\,,$$
$$t_n^- = t^* - \frac{15d}{\alpha_{\mathfrak{s}}^\star \lambda^\star} \log\log n\,, \qquad t_n^+ = t^* + \frac{25d}{\alpha_{\mathfrak{s}}^\star \lambda^\star} \log\log n\,.$$

The following then holds:

$$\max_{x_0} \left\| \mathbb{P}_{x_0}(X_{t_n^-} \in \cdot) - \mu \right\|_{\mathrm{TV}} \to 1\,,$$
$$\max_{x_0} \left\| \mathbb{P}_{x_0}(X_{t_n^+} \in \cdot) - \mu \right\|_{\mathrm{TV}} \to 0\,.$$

Proof. We begin by applying Lemma 4.1 to t_n^-. The left-hand-side of (4.1) establishes that

$$(n/\log^3 n)^d \mathfrak{m}_{t_n^-}^2 \geq (n/\log^3 n)^d \left(n^{-d} (\log n)^{\frac{6d}{\alpha_{\mathfrak{s}}^\star}} - n^{-9d} \right) \geq (1+o(1)) \log^{3d} n\,,$$

and in particular $(n/\log^3 n)^d \mathfrak{m}_{t_n^-}^2 \to \infty$, hence by Part 2 of Theorem 3.1

$$\max_{x_0} \left\| \mathbb{P}_{x_0}(X_{t_n^-} \in \cdot) - \mu \right\|_{\mathrm{TV}} \to 1\,.$$

Using the right-hand-side of (4.1) for $t = t_n^+ - (10d/\alpha_{\mathfrak{s}}^\star) \log\log n$ gives

$$(n/\log^5 n)^d \mathfrak{m}_t^2 \leq (n/\log^5 n)^d \cdot n^{-d} (\log n)^{-\frac{6d}{\alpha_{\mathfrak{s}}^\star}} \leq \log^{-11d} n = o(1)\,.$$

Therefore, by Part 1 of Theorem 3.1, at time $t_n^+ = t + s$ we have

$$\max_{x_0} \left\| \mathbb{P}_{x_0}(X_{t_n^+} \in \cdot) - \mu \right\|_{\mathrm{TV}} \to 0\,,$$

as required. ∎

4.3. **Limit of spectral gaps.** Theorem 4.2 established the location of the mixing time in terms of $\lambda(r)$, the spectral gap of the Glauber dynamics on the d-dimensional torus of side-length r. As commented at the beginning of Section 3, we are allowed some latitude in our choice of r. This provides a way of relating the spectral gaps for different values of r, ultimately proving that they converge to λ_∞, the spectral gap on the infinite-volume lattice.

Lemma 4.3. *Let* $\beta \geq 0$*, and let* $\lambda(r)$ *be the spectral gap of continuous-time Glauber dynamics for the Ising model on* $\mathbb{Z}_r^d$ *at inverse-temperature* β*. If there is strong spatial mixing at inverse temperature* β *then there exists some* $\hat\lambda > 0$ *such that*

$$\left| \lambda(r) - \hat\lambda \right| \leq r^{-1/2+o(1)}\,. \tag{4.5}$$

Proof. As noted in the beginning of Section 3, there is a lot of freedom in the choice of $r = 3\log^3 n$ for the definition of $\mathfrak{m}_t$ in (3.1), and the proofs hold as is (while resulting in slightly different absolute constants, e.g., c_0 in Theorem 4.2 etc.) for $r = \log^{2+\delta} n$ with an arbitrary fixed $\delta > 0$. With this in mind, fix some small $\delta > 0$ and take $r_1 = \log^{2+\delta} n$ and $r_1 \leq r_2 \leq r_1^2$.

In Theorem 4.2 we provided upper and lower bounds on the quantity $\max_{x_0} \|\mathbb{P}_{x_0}(X_t \in \cdot) - \mu\|_{\mathrm{TV}}$ in terms of $\lambda(r), \alpha_{\mathsf{s}}^\star, \lambda^\star$ and some absolute constants. As argued above, there exists some $C = C(\delta) > 0$ such that for every $r_1 \leq r \leq r_1^2$ the statement of the theorem holds with parameters

$$t^* = t^*(n) = \frac{d}{2\lambda(r)} \log n\,,$$
$$t_n^- = t^* - \frac{Cd}{\alpha_{\mathsf{s}}^\star \lambda^\star} \log\log n\,, \qquad t_n^+ = t^* + \frac{Cd}{\alpha_{\mathsf{s}}^\star \lambda^\star} \log\log n\,.$$

Applying this theorem both on r_1 and r_2 we must have $t_n^+(r_1) \geq t_n^-(r_2)$ for sufficiently large n, and so

$$\frac{d}{2\lambda(r_1)} \log n + \frac{Cd}{\alpha_{\mathsf{s}}^\star \lambda^\star} \log\log n \geq \frac{d}{2\lambda(r_2)} \log n - \frac{Cd}{\alpha_{\mathsf{s}}^\star \lambda^\star} \log\log n\,.$$

Rearranging we obtain that

$$\lambda(r_1) - \lambda(r_2) \leq 4C \frac{\lambda(r_1)\lambda(r_2)}{\alpha_{\mathsf{s}}^\star \lambda^\star} \frac{\log\log n}{\log n} \leq r_1^{-1/2+\delta}\,,$$

where the last inequality holds for any sufficiently large n. As we can clearly reverse the role of r_1 and r_2 it follows that for any large n,

$$\max_{r_1 < r_2 \leq r_1^2} |\lambda(r_1) - \lambda(r_2)| \leq r_1^{-1/2+\delta}\,. \tag{4.6}$$

By the above inequality, if r is a sufficiently large integer then

$$\sum_{i=0}^{\infty} \left|\lambda\big(r^{2^i}\big) - \lambda\big(r^{2^{i+1}}\big)\right| \leq \sum_{i=0}^{\infty} r^{-2^{i-1} + 2^i \delta} \leq 2r^{-1/2+\delta} < \infty\,. \tag{4.7}$$

Combining equations (4.6) and (4.7) establishes that $\{\lambda(r)\}_{r=0}^\infty$ converges to some limit $\hat\lambda$ and that for large r,

$$\left|\lambda(r) - \hat\lambda\right| \leq r^{-1/2+\delta}\,.$$

Letting $\delta \to 0$ completes the proof. ∎

It remains to show that the above $\hat\lambda$ is equal to λ_∞, the spectral gap of the dynamics on the infinite-volume lattice. Let (σ_t) denote the Glauber dynamics on the infinite volume lattice, let (σ_t^+) be the dynamics starting from the all-plus configuration and define

$$\xi_t = \mathbb{P}\left(\sigma_t^+(o) = 1\right) - \mathbb{P}\left(\sigma_t^-(o) = 1\right)\,,$$

which in the special case of no external field is simply $2\left(\mathbb{P}\left(\sigma_t^+(o)=1\right)-\frac{1}{2}\right)$.

Claim 4.4. *The above defined* ξ_t *and* $\hat\lambda$ *satisfy* $\left|\xi_t^{1/t}-\exp(-\hat\lambda)\right| = O\left(\frac{\log t}{t}\right)$, *and in particular* $\lim_{t\to\infty}\xi_t^{1/t}=\exp(-\hat\lambda)$.

Proof. Put $n=\exp(t)$ and let (X_t^+) and (X_t^-) denote the Glauber dynamics starting from the all-plus and all-minus configurations respectively on the torus with side-length $r=3\log^3 n$. Consider the monotone coupling of (X_t^+) and (X_t^-). By the symmetry of the torus, the expected number of disagreements at time t is given by $r^d\left(\mathbb{P}(X_t^+(o)=1)-\mathbb{P}(X_t^-(o)=1)\right)$ and hence

$$\begin{aligned}\mathbb{P}(X_t^+(o)=1)-\mathbb{P}(X_t^-(o)=1) &\le 2\max_{x_0}\|\mathbb{P}_{x_0}(X_t\in\cdot)-\mu\|_{\mathrm{TV}}\\ &\le 2r^d\left(\mathbb{P}(X_t^+(o)=1)-\mathbb{P}(X_t^-(o)=1)\right)\,,\end{aligned}\tag{4.8}$$

where the first inequality is by definition of the total variation distance.

Next, identify the vertices of $\mathbb{Z}_r^d$ with those in $\{x\in\mathbb{Z}^d:\|x\|_\infty\le r/2\}$ and couple (X_t^+) and (σ_t) via identical updates to identified vertices. By another simple application of the disagreement percolation argument (as in the proof of Lemma 3.6),

$$\mathbb{P}\left(\sigma_t^+(o)\neq X_t^+(o)\right)\le n^{-10d}\,.$$

Combining this with (4.8),

$$r^{-d}\max_{x_0}\|\mathbb{P}_{x_0}(X_t\in\cdot)-\mu\|_{\mathrm{TV}}-n^{-10d}\le\xi_t\le 2\max_{x_0}\|\mathbb{P}_{x_0}(X_t\in\cdot)-\mu\|_{\mathrm{TV}}+n^{-10d}.$$

Applying (4.1) and (4.3) for a choice of $s=(10d/\alpha_{\mathsf{s}}^\star)\log\log n$ now gives

$$\xi_t\le O(n^{-9d})+2r^{d/2}\mathrm{e}^{-\lambda(r)(t-s)+c_0\log\log n}\le\mathrm{e}^{-\lambda(r)t+c_1\log\log n}$$

and

$$\xi_t\ge\frac{1}{2r^d}\mathrm{e}^{-\lambda(r)t}-n^{-10d}\ge\mathrm{e}^{-\lambda(r)t-c_1\log\log n}$$

for some constant $c_1>0$. Altogether, as $r=3\log^3 n$ and $t=\log n$,

$$\left|\log\frac{\xi_t^{1/t}}{\exp(-\lambda(3t^3))}\right|\le c_1\frac{\log\log n}{t}=c_1\frac{\log t}{t}\,,$$

which combined with Lemma 4.3 completes the proof. ∎

A result of Holley [19] shows that $\lim_{t\to\infty}\xi_t^{1/t}=\exp(-\lambda_\infty)$. We will show how this result is quickly recovered from our proof. Plugging the test-function $f(\sigma)=\mathbb{1}_{\{\sigma(o)=1\}}-\mathbb{E}\mathbb{1}_{\{\sigma(o)=1\}}$ into the characterization of the spectral gap as the slowest rate of exponential decay of the semigroup gives

$$\exp(-\lambda_\infty)\ge\lim_{t\to\infty}\xi_t^{1/t}=\exp(-\hat\lambda)\,.$$

Now fix $\varepsilon > 0$ and recall the Dirichlet form (2.1), according to which

$$\lambda_\infty = \inf_{f \in L^2(\{\pm 1\}^{\mathbb{Z}^d}, \mu_\infty)} \frac{\mathcal{E}_{\mu_\infty}(f,f)}{\mathrm{Var}_{\mu_\infty}(f)}\,,$$

where μ_∞ is the stationary measure of the infinite-volume Ising model. For any $f \in L^2(\{\pm 1\}^{\mathbb{Z}^d}, \mu_\infty)$ with $\mathcal{E}_{\mu_\infty}(f,f)$ we can find a sequence of functions $f_n \in L^2(\{\pm 1\}^{\mathbb{Z}^d}, \mu_\infty)$ each of which depends only on a finite number of spins such that $f_n \to f$ in $L^2(\{\pm 1\}^{\mathbb{Z}^d}, \mu_\infty)$ and $\mathcal{E}_{\mu_\infty}(f_n, f_n) \to \mathcal{E}_{\mu_\infty}(f,f)$ (see e.g. the proof of [24, Lemma 4.3]). So take $g \in L^2(\{\pm 1\}^{\mathbb{Z}^d}, \mu_\infty)$ depending only on a finite number of spins such that

$$\frac{\mathcal{E}_{\mu_\infty}(g,g)}{\mathrm{Var}_{\mu_\infty}(g)} \leq \frac{\mathcal{E}_{\mu_\infty}(f,f)}{\mathrm{Var}_{\mu_\infty}(f)} + \varepsilon.$$

For some large enough M we have that g is a function of the spins in the box $B_M = \{x \in \mathbb{Z}^d : \|x\|_\infty \leq M\}$. We compare the Glauber dynamics on $\mathbb{Z}^d$ to that on $\mathbb{Z}_r^d$ for some large r, identifying the vertices of the latter with those in $\{x \in \mathbb{Z}^d : \|x\|_\infty \leq r/2\}$ and denoting its stationary distribution by μ_r. By the strong spatial mixing property, the projection of μ_r on B_M converges to the projection of μ_∞ on B_M as $r \to \infty$ (in fact, an assumption weaker than strong spatial mixing would already infer this, e.g. uniqueness), hence

$$\frac{\mathcal{E}_{\mu_\infty}(g,g)}{\mathrm{Var}_{\mu_\infty}(g)} = \lim_{r\to\infty} \frac{\mathcal{E}_{\mu_r}(g,g)}{\mathrm{Var}_{\mu_r}(g)} \geq \lim_{r\to\infty} \lambda(r) = \hat{\lambda}\,,$$

where the inequality follows from the characterization of the spectral gap by the Dirichlet form. This implies that

$$\lambda_\infty = \inf_f \frac{\mathcal{E}_{\mu_\infty}(f,f)}{\mathrm{Var}_{\mu_\infty}(f)} \geq \hat{\lambda} - \varepsilon\,,$$

and letting $\varepsilon \to 0$ gives that $\lambda_\infty = \hat{\lambda}$, as required. This completes the proof of Theorems 1 and 2. ∎

5. Cutoff for other spin-systems

While the proof above was given for the ferromagnetic Ising model, its arguments naturally extend to other well-studied spin-system models. In this section we outline the minor modifications one needs to make in order to obtain cutoff for general monotone and anti-monotone systems, thereby proving Theorems 5 and 6.

A key prerequisite for our proofs is a uniformly bounded log-Sobolev constant for the dynamics. This was proved in great generality in [28] for Glauber dynamics on spin-system models on the lattice where strong spatial mixing holds, and that result carries to periodic boundary with minor adjustments.

Otherwise, the only specific property of the Ising model used in our arguments is the monotonicity of the model. The Glauber dynamics (1.2) for a spin-system is *monotone* (or *attractive*) if there is a partial ordering of the state space $\preceq$ according to which the transition rates satisfy

$$\sigma(x)c(x,\sigma) \leq \eta(x)c(x,\eta) \text{ for all } \eta \preceq \sigma \text{ with } \eta(x) = \sigma(x)\,. \tag{5.1}$$

Under this condition, the system allows a monotone coupling of the Glauber dynamics. We use the monotonicity of the model in precisely two locations.

(1) In Lemma 3.9 we prove that the update support is sparse w.h.p. by showing that in most of the tori that make up the barrier dynamics the all-plus and all-minus configurations couple (hence the final configuration is independent of the projection of the starting configuration onto those tori). This is a crucial application of the system's monotonicity.
(2) In equation (4.8) of Claim 4.4 we relate the expected number of disagreements in the monotone coupling of the dynamics to the total-variation distance of the dynamics from stationarity.

The arguments in both of these steps hold essentially unchanged for any monotone dynamics.

An anti-monotone system is one where the reverse inequality in (5.1) always holds. When the underlying geometry is a bipartite graph, there is a standard transformation of an anti-monotone system into a monotone one. Let $V = V_e \cup V_o$ be a partition of the sites such that there are no edges within V_e or V_o. We define a new partial ordering $\preceq^*$ on $\{\pm 1\}^V$ as follows: For two configurations σ, η we have $\sigma \preceq^* \eta$ if $\sigma(v) \leq \eta(v)$ for all $v \in V_e$ and $\sigma(v) \geq \eta(v)$ for all $v \in V_o$. It is easy to verify that an anti-monotone system under the standard partial ordering $\preceq$ is a monotone system under $\preceq^*$.

To derive Theorems 5 and 6, note that the d-dimensional lattice with periodic boundary conditions $(\mathbb{Z}/r\mathbb{Z})^d$ is bipartite if and only if r is even. Crucially, our proof for cutoff on $(\mathbb{Z}/n\mathbb{Z})^d$ only required the dynamics to be monotone on the smaller tori $(\mathbb{Z}/r\mathbb{Z})^d$ for $r = r(n)$ as defined above (3.1). As we noted there, we have freedom for our choice of r (basically any choice between $\log^{2+\varepsilon} n$ and $\log^{O(1)} n$ would do) and in particular we can let r be even. Applying the above transformation therefore establishes cutoff for the anti-ferromagnetic Ising model and the gas hard-core model.

As a side note, in our companion paper [26] we establish that any model with soft interactions (e.g. the Potts model) has cutoff at high enough temperatures. One of the key challenges there is to show that the update support is typically sparse. We further address non-periodic boundary conditions, which break down the symmetric structure of the torus, thereby letting the smaller tori (which we analyze in our L^1-L^2 reduction) play different roles in the dynamics according to their vicinity to the boundary.

Acknowledgments

We are grateful to Yuval Peres for inspiring us to pursue this research. We thank him and Fabio Martinelli for useful discussions.

This work was initiated while the second author was an intern at the Theory Group of Microsoft Research as a doctoral student at UC Berkeley.

References

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

[2] D. Aldous, *Random walks on finite groups and rapidly mixing Markov chains*, Seminar on probability, XVII, 1983, pp. 243–297.

[3] 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`.

[4] D. Aldous and P. Diaconis, *Shuffling cards and stopping times*, Amer. Math. Monthly **93** (1986), 333–348.

[5] F. Cesi, *Quasi-factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields*, Probab. Theory Related Fields **120** (2001), no. 4, 569–584.

[6] G.-Y. Chen and L. Saloff-Coste, *The cutoff phenomenon for ergodic Markov processes*, Electronic Journal of Probability **13** (2008), 26–78.

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

[8] P. Diaconis and L. Saloff-Coste, *Comparison techniques for random walk on finite groups*, Ann. Probab. **21** (1993), no. 4, 2131–2156.

[9] P. Diaconis and L. Saloff-Coste, *Comparison theorems for reversible Markov chains*, Ann. Appl. Probab. **3** (1993), no. 3, 696–730.

[10] P. Diaconis and L. Saloff-Coste, *Logarithmic Sobolev inequalities for finite Markov chains*, Ann. Appl. Probab. **6** (1996), no. 3, 695–750.

[11] P. Diaconis and L. Saloff-Coste, *Nash inequalities for finite Markov chains*, J. Theoret. Probab. **9** (1996), no. 2, 459–510.

[12] P. Diaconis and M. Shahshahani, *Generating a random permutation with random transpositions*, Z. Wahrsch. Verw. Gebiete **57** (1981), no. 2, 159–179.

[13] J. Ding, E. Lubetzky, and Y. Peres, *The mixing time evolution of Glauber dynamics for the mean-field Ising model*, Comm. Math. Phys. **289** (2009), no. 2, 725–764.

[14] J. Ding, E. Lubetzky, and Y. Peres, *Total-variation cutoff in birth-and-death chains*, Probab. Theory Related Fields, to appear.

[15] R. L. Dobrushin and S. B. Shlosman, *Completely analytical interactions: constructive description*, J. Statist. Phys. **46** (1987), no. 5-6, 983–1014.

[16] R. J. Glauber, *Time-dependent statistics of the Ising model*, J. Mathematical Phys. **4** (1963), 294–307.

[17] A. Guionnet and B. Zegarlinski, *Decay to equilibrium in random spin systems on a lattice*, Comm. Math. Phys. **181** (1996), no. 3, 703–732.

[18] T. P. Hayes and A. Sinclair, *A general lower bound for mixing of single-site dynamics on graphs*, Foundations of Computer Science, 2005. FOCS 2005. 46th Annual IEEE Symposium on, 2005, pp. 511–520.

[19] R. A. Holley, *On the asymptotics of the spin-spin autocorrelation function in stochastic Ising models near the critical temperature*, Spatial stochastic processes, Progr. Probab., vol. 19, Birkhäuser Boston, Boston, MA, 1991, pp. 89–104.

[20] R. A. Holley and D. W. Stroock, *Logarithmic Sobolev inequalities and stochastic Ising models*, J. Statist. Phys. **46** (1987), no. 5-6, 1159–1194.

[21] R. A. Holley and D. W. Stroock, *Uniform and L^2 convergence in one-dimensional stochastic Ising models*, Comm. Math. Phys. **123** (1989), no. 1, 85–93.

[22] 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.

[23] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov chains and mixing times*, 2008.

[24] T. M. Liggett, *Interacting particle systems*, Classics in Mathematics, Springer-Verlag, Berlin, 2005. Reprint of the 1985 original.

[25] S. L. Lu and H.-T. Yau, *Spectral gap and logarithmic Sobolev inequality for Kawasaki and Glauber dynamics*, Comm. Math. Phys. **156** (1993), no. 2, 399–433.

[26] E. Lubetzky and A. Sly, *Cutoff for general spin systems with arbitrary boundary conditions*. preprint.

[27] E. Lubetzky and A. Sly, *Cutoff phenomena for random walks on random regular graphs*, Duke Math. J., to appear.

[28] F. Martinelli, *Lectures on Glauber dynamics for discrete spin models*, Lectures on probability theory and statistics (Saint-Flour, 1997), Lecture Notes in Math., vol. 1717, Springer, Berlin, 1999, pp. 93–191.

[29] F. Martinelli, *Relaxation times of Markov chains in statistical mechanics and combinatorial structures*, Probability on discrete structures, Encyclopaedia Math. Sci., vol. 110, Springer, Berlin, 2004, pp. 175–262.

[30] F. Martinelli and E. Olivieri, *Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case*, Comm. Math. Phys. **161** (1994), no. 3, 447–486.

[31] F. Martinelli and E. Olivieri, *Approach to equilibrium of Glauber dynamics in the one phase region. II. The general case*, Comm. Math. Phys. **161** (1994), no. 3, 487–514.

[32] F. Martinelli, E. Olivieri, and R. H. Schonmann, *For 2-D lattice spin systems weak mixing implies strong mixing*, Comm. Math. Phys. **165** (1994), no. 1, 33–47.

[33] L. Saloff-Coste, *Lectures on finite Markov chains*, Lectures on probability theory and statistics (Saint-Flour, 1996), Lecture Notes in Math., vol. 1665, Springer, Berlin, 1997, pp. 301–413.

[34] L. Saloff-Coste, *Random walks on finite groups*, Probability on discrete structures, 2004, pp. 263–346.

[35] D. W. Stroock and B. Zegarliński, *The equivalence of the logarithmic Sobolev inequality and the Dobrushin-Shlosman mixing condition*, Comm. Math. Phys. **144** (1992), no. 2, 303–323.

[36] D. W. Stroock and B. Zegarliński, *The logarithmic Sobolev inequality for continuous spin systems on a lattice*, J. Funct. Anal. **104** (1992), no. 2, 299–326.

[37] D. W. Stroock and B. Zegarliński, *The logarithmic Sobolev inequality for discrete spin systems on a lattice*, Comm. Math. Phys. **149** (1992), no. 1, 175–193.

[38] B. Zegarliński, *Dobrushin uniqueness theorem and logarithmic Sobolev inequalities*, J. Funct. Anal. **105** (1992), no. 1, 77–111.

[39] B. Zegarliński, *On log-Sobolev inequalities for infinite lattice systems*, Lett. Math. Phys. **20** (1990), no. 3, 173–182.

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

Allan Sly
Microsoft Research, One Microsoft Way, Redmond, WA 98052-6399, USA.
E-mail address: allansly@microsoft.com