On sampling diluted Spin-Glasses using Glauber Dynamics

Charilaos Efthymiou CHARILAOS.EFTHYMIOU@WARWICK.AC.UK
Kostas Zampetakis KOSTAS.ZAMPETAKIS@WARWICK.AC.UK
University of Warwick

Editors: Shipra Agrawal and Aaron Roth

Abstract

Spin-glasses are natural Gibbs distributions that have been studied in theoretical computer science for many decades. Recently, they have been gaining renewed attention from the community as they emerge naturally in *neural computation* and *learning*, *network inference*, *optimisation* and many other areas.

Here we consider the *2-spin model* at inverse temperature β when the underlying graph is an instance of $\boldsymbol{G}(n, d/n)$, i.e., the random graph on n vertices such that each edge appears independently with probability d/n, where the expected degree $d = \Theta(1)$. We study the problem of efficiently sampling from the aforementioned distribution using the well-known Markov chain called *Glauber dynamics*.

For a certain range of β, that depends only on the expected degree d of the graph, and for typical instances of the 2-spin model on $\boldsymbol{G}(n, d/n)$, we show that the corresponding (single-site) Glauber dynamics exhibits mixing time $O\left(n^{2+\frac{3}{\log^2 d}}\right)$. The range of β for which we obtain our rapid mixing result corresponds to the expected influence being smaller than $1/d$.

We establish our results by utilising the well-known *path-coupling* technique. In the standard setting of Glauber dynamics on $\boldsymbol{G}(n, d/n)$ one has to deal with the so-called effect of high degree vertices. Here, with the spin-glasses, rather than considering vertex-degrees, it is more natural to use a different measure on the vertices of the graph, that we call *aggregate influence*.

We build on the block-construction approach proposed by [Dyer, Flaxman, Frieze and Vigoda: 2006] to circumvent the problem with the high degrees in the path-coupling analysis. Specifically, to obtain our results, we first establish rapid mixing for an appropriately defined block-dynamics. We design this dynamics such that vertices of large aggregate influence are placed deep inside their blocks. Then, we obtain rapid mixing for the (single-site) Glauber dynamics by utilising a comparison argument.

Keywords: Spin glasses, Phase transitions, Gibbs Sampler, Random graphs

1. Introduction

Spin-glasses are natural, high dimensional Gibbs distributions that have been studied in theoretical computer science for many decades. Recently, they have been gaining renewed attention from the community as they emerge naturally in *neural computation* and *learning*, e.g., in the Hopfield model, as models of *network inference*, e.g., in the stochastic block model, in *optimisation*, *counting-sampling* and many other areas, e.g., see Stein and Newman (2013); Gamarnik et al.

(2020); Coja-Oghlan et al. (2018); Alaoui et al. (2021); Koehler et al. (2022); Eldan et al. (2022); Efthymiou (2022).

Furthermore, spin-glasses are widely considered to be canonical models of *extremely* disordered systems Mézard et al. (1987); Stein and Newman (2013) and, as such, they have beed studied extensively in mathematics & mathematical physics, e.g., see Talagrand (2006); Panchenko (2013); Franz et al. (2001); Guerra and Toninelli (2004), and also in statistical physics Mézard et al. (1987); Stein and Newman (2013); Parisi (1979). In particular, as far as physics is concerned, spin-glasses have been studied intensively since the early '80s, while the seminal, groundbreaking work of Giorgio Parisi on spin-glasses got him the Nobel Prize in Physics in 2021.

In this work, we consider the natural problem of sampling from these distributions. To this end, our endeavour is to employ the powerful Markov Chain Monte Carlo (MCMC) method. Typically, the MCMC sampling algorithms are very simple to describe and implement in practice, however, analysing their performance can be extremely challenging.

We focus on a very natural and well-known case of spin glasses: the *2-spin model*. Given a fixed graph $G = (V, E)$, and the vector $\boldsymbol{\mathcal{J}} = \{\boldsymbol{J}_e : e \in E\}$ of independent, identically distributed (i.i.d.) *standard Gaussians*, the 2-spin model with inverse temperature $\beta > 0$, corresponds to the Gibbs distribution $\mu = \mu_{G,\boldsymbol{\mathcal{J}},\beta}$ on the configuration space $\{\pm 1\}^V$ such that

$$\mu(\sigma) \propto \exp\left(\beta \sum\nolimits_{\{w,u\}\in E} \mathbb{1}\{\sigma(u) = \sigma(w)\} \cdot \boldsymbol{J}_{\{u,w\}}\right) \quad , \tag{1}$$

where $\propto$ stands for "proportional to". We usually refer to vector $\boldsymbol{\mathcal{J}}$ as the *edge couplings*.

The above distribution is encountered in the literature with different names, i.e., depending on the underlying graph. For example, the 2-spin model on the complete graph corresponds to the well-known *Sherrington-Kirkpatrick* model (SK model) Sherrington and Kirkpatrick (1975). For the d-dimensional square lattice $\mathbb{Z}^d$, we use the name *Edwards-Anderson model*. The choice of terminology, here, is standard and follows, e.g., Panchenko and Talagrand (2004).

On a first account, the 2-spin model may look innocent, i.e., it looks similar to the standard Ising model with just the addition of the Gaussian couplings. It turns out, though, that it is a *fascinating* distribution with a lot of intricacies, while the configuration space has extremely rich structure, e.g. in various settings, it is conjectured to exhibit the "*infinite* Replica Symmetry Breaking" Mézard et al. (1987).

Interestingly, with the Gaussian couplings, we expect to have edges e in the graph such that $\boldsymbol{J}_e$ is *unbounded*. There is a related, more tame spin-glass distribution where the couplings are independent, random $\{\pm 1\}$, rather than Gaussian. In this case, the disorder comes from interactions choosing between being ferromagnetic, or antiferromagnetic, while their intensity always remains *bounded*. Here, unless otherwise is specified, we *always* assume that the couplings are standard Gaussians.

We use the *Glauber dynamics* to sample from the 2-spin model. We assume that the underlying geometry is an instance of the *sparse* random graph $\boldsymbol{G}(n, d/n)$. This is a random graph on n vertices, such that each edge appears, independently, with probability $p = d/n$, for some constant

$d > 0$. Note that we obtain an instance of the problem by first drawing the underlying graph form the distribution $\boldsymbol{G}(n, d/n)$ and then, given the graph, we generate the random Gibbs distribution.

Sampling from Gibbs distributions induced by instances of $\boldsymbol{G}(n, d/n)$, or, more generally, instances of so-called random constraint satisfaction problems, is at the heart of recent endeavors to investigate connections between *phase transitions* and the efficiency of algorithms, e.g. Achlioptas and Coja-Oghlan (2008); Alaoui and Montanari (2020); Coja-Oghlan and Efthymiou (2011); Galanis et al. (2015); Gamarnik and Sudan (2017); Sly and Sun (2012); Mossel and Sly (2013). The MCMC sampling problem on random graphs has garnered a lot of attention, e.g., see Dyer et al. (2006); Bezáková et al. (2022); Efthymiou and Feng (2023); Mossel and Sly (2010, 2013); Efthymiou et al. (2018a); Dyer and Frieze (2010); Chen et al. (2023), as it is considered to be an intriguing case to study. So far, the focus has been on sampling standard Gibbs distributions, which, already, is a very challenging problem. The study of spin-glasses takes us a step further. Working with the 2-spin model, we introduce an *extra level of disorder* which is due to the Gaussian couplings at the edges of the graph. Hence, in our analysis, we need to deal with the disorder of both $\boldsymbol{G}(n, d/n)$ and the random couplings.

For typical instances of *both* the random graph $\boldsymbol{G}(n, d/n)$ and the 2-spin model, we show that Glauber dynamics exhibits mixing time $O(n^{2+\frac{3}{\log^2 d}})$ for any inverse temperature $\beta > 0$ such that

$$\mathbb{E}\left[\tanh\left(\tfrac{\beta}{2}\,|\boldsymbol{J}|\right)\right] < 1/d \ , \tag{2}$$

where the expectation is with respect to the standard Gaussian random variable $\boldsymbol{J}$. For large d, the above condition corresponds to (roughly) having $0 < \beta < \frac{\sqrt{2\pi}}{d}$. In turn, this implies that the *expected influence* is smaller than $1/d$, a quite natural requirement to have.

Until recently, the focus of the study for MCMC sampling has been on the *mean-field* SK model, with the most recent works being Alaoui et al. (2022); Bauerschmidt and Bodineau (2019); Koehler et al. (2022); Eldan et al. (2022). For the cases we consider here, these results imply a much weaker bound on β compared to what we get with our approach, i.e., they require β to be a vanishing function of n.[1] This comes without surprise as the SK model lives on the complete graph, hence, there is *no geometry* in the problem. On the other hand, the 2-spin model on $\boldsymbol{G}(n, d/n)$ has a very rich, and as it turns out, intricate geometry.

Perhaps, more related to our paper is the recent work in Anari et al. (2024) which studies the distribution with random $\{\pm 1\}$ couplings, on the random d-regular graph. Interestingly, they obtain a bound $\beta = O(1/\sqrt{d})$ for rapid mixing of Glauber dynamics. As opposed to the distribution we consider here, where both vertex degrees and couplings take unbounded values, in the distribution considered in Anari et al. (2024), both these quantities remain *bounded*. Subsequently to our work here, Liu et al. (2024) obtains a $\beta = O(1/\sqrt{d})$ rapid mixing bound of Glauber dynamics for the distribution with random $\{\pm 1\}$ couplings on $\boldsymbol{G}(n, d/n)$. The approach in Liu et al. (2024) is different than ours and uses *stochastic localization* ideas from Eldan et al. (2022). It seems, though,

1. More concretely, it is elementary to verify that they require $\beta = O\left(\sqrt{\frac{\log\log n}{\log n}}\right)$.

that the block-construction analysis, there, cannot produce similar bounds, or improvements, for the case of Gaussian couplings.

In light of Anari et al. (2024); Liu et al. (2024), it seems natural to have a bound $\beta = O(1/\sqrt{d})$ for rapid mixing of Glauber dynamics on the 2-spin model on $G(n, d/n)$, i.e., with Gaussian couplings. Our conjecture is that the mixing time in the aforementioned range of β is $n \exp(\Theta(\sqrt{\log n}))$. This is because we typically have isolated stars with rather large couplings at their edges, i.e., we expect to see couplings of magnitude as large as $\Theta(\sqrt{\log n})$. Let us remark that the conjectured bound for the mixing time does not apply to the $\{\pm 1\}$-case, as the couplings there are bounded.

Our analysis does not rely on the newly introduced methods such as the *Spectral Independence* (SI) Anari et al. (2020); Chen et al. (2021, 2022), even though there are rapid mixing results for Gibbs distributions on $\boldsymbol{G}(n, d/n)$ that utilise SI, e.g., see Bezáková et al. (2022); Efthymiou and Feng (2023). When it comes to spin-glasses, the natural quantities that arise with SI turn out to be too complicated to work with. Our approach exploits the classic *path-coupling* technique Bubley and Dyer (1997). Specifically, we build on the machinery developed in the sequence of results in Dyer et al. (2006); Efthymiou (2014); Efthymiou et al. (2018b) that establish fast mixing of the Glauber dynamics for (standard) Gibbs distributions on $\boldsymbol{G}(n, d/n)$. Basically, our approach builds on these ideas so that we can accommodate the extra disorder that the problem exhibits. Note that these previous works are about distributions such as the colouring model, the hard-core model, etc., and not for spin-glasses.

For typical instances of $\boldsymbol{G}(n, d/n)$ all but an $\exp(-\Theta(d))$ fraction of the vertices are of degree close to d, while we expect to have vertices with degrees as huge as $\Theta\left(\frac{\log n}{\log\log n}\right)$. In that respect, it is natural to have the parameters of the problem expressed in terms of the expected degree d, rather than, e.g., the maximum degree. Hence, a major challenge in the analysis is how to circumvent the so-called "effect of high degrees".

Roughly speaking, the approach underlying Efthymiou et al. (2018b) is as follows: rather than considering (single site) Glauber dynamics, we consider *block dynamics*. That is, there is an appropriately constructed block-partition of the set of vertices such that, at each step, we update the configuration of a randomly chosen block. As it was already observed in Dyer et al. (2006), the typical instances of $\boldsymbol{G}(n, d/n)$ admit a block-partition, such that the high-degree vertices are hidden deep inside the blocks, in a way that makes their effect vanish. This allows one to circumvent the problem that the high-degree vertices pose in the path-coupling analysis and show fast mixing for the block dynamics. Subsequently, one obtains the bounds on the mixing time for the single-site dynamics using a *comparison* argument.

A primary challenge of the above approach is to construct the desirable block-partition. The contribution coming from the aforementioned works amounts to introducing a *weighting-scheme* (a set of potentials) for the paths in the graph, leveraged for the block construction. Typically, this approach entails heavy probabilistic analysis.

It turns out that the use of potentials for the paths is quite natural in our setting, too. One of our main contributions here is to introduce *new* weights (new potentials) for the paths of the graph,

which accommodate the richer structure of the problem. Unlike the previous works that focus only on vertex-degrees, we utilise concepts from Spectral Independence and introduce a new measure for the vertices of the graph, which we call *aggregate influence*.

Overall, getting a handle on the behaviour of the weights of the paths is one of the most technically demanding parts of this work. It is worth mentioning that the set of blocks we obtain here is quite different from the one that appears in Efthymiou et al. (2018b). Note that in this work, the block partition is such that vertices with large degree are hidden deep inside the blocks. Here, it is typical to have single-vertex blocks consisting of a high-degree vertex, or having multi-vertex blocks whose vertices are all low-degree ones. This comes without surprise, as the notions of the aggregate influence and the degree of a vertex are different from each other.

1.1. Results

We let $\boldsymbol{G} = \boldsymbol{G}(n, d/n)$ be the Erdős–Rényi graph on a set V_n of n vertices, with edge probability d/n, where $d > 0$ is a fixed number. The 2-spin model on $\boldsymbol{G}$ at *inverse temperature* $\beta > 0$, is defined as follows: for $\boldsymbol{\mathcal{J}} = \{\boldsymbol{J}_e : e \in E(\boldsymbol{G})\}$ a family of independent, *standard Gaussians*, and for $\sigma \in \{\pm 1\}^{V_n}$, we let

$$\mu_{\boldsymbol{G},\boldsymbol{\mathcal{J}},\beta}(\sigma) = \frac{1}{Z_\beta(\boldsymbol{G},\boldsymbol{\mathcal{J}})} \cdot \exp\left(\beta \sum_{x\sim y} \mathbb{1}\{\sigma(y) = \sigma(x)\} \cdot \boldsymbol{J}_{\{x,y\}}\right) \quad , \tag{3}$$

where

$$Z_\beta(\boldsymbol{G},\boldsymbol{\mathcal{J}}) = \sum_{\tau\in\{\pm 1\}^{V_n}} \exp\left(\beta \sum_{x\sim y} \mathbb{1}\{\tau(y) = \tau(x)\} \cdot \boldsymbol{J}_{\{x,y\}}\right) \quad .$$

Typically, we study this distribution as $n \to \infty$.

We use the discrete time, (single site) *Glauber dynamics* $(X_t)_{t\geq 0}$ to approximately sample from the aforementioned distribution $\mu = \mu_{\boldsymbol{G},\boldsymbol{\mathcal{J}},\beta}$. Glauber dynamics is a Markov chain with state space the support of the distribution μ. We assume that the chain starts from an arbitrary configuration $X_0 \in \{\pm 1\}^{V_n}$. For $t \geq 0$, the transition from the state X_t to X_{t+1} is according to the following steps:

1. choose uniformly at random a vertex v,
2. for every vertex w different than v, set $X_{t+1}(w) = X_t(w)$,
3. set $X_{t+1}(v)$ according to the marginal of μ at v, conditional on the neighbours of v having the configuration specified by X_{t+1}.

It is standard to show that when a Markov chain satisfies a set of technical conditions, called *ergodicity*, then it converges to a unique stationary distribution. For the cases we consider here, the Glauber dynamics is trivially ergodic.

Let P be the transition matrix of an ergodic Markov chain (X_t) with a finite state space Ω, and equilibrium distribution μ. For $t \geq 0$, and $\sigma \in \Omega$, let $P^t(\sigma, \cdot)$ denote the distribution of X_t, when the initial state of the chain is σ, i.e., $X_0 = \sigma$. The *mixing time* of the chain $(X_t)_{t\geq 0}$ is defined by

$$\mathrm{T}_{\mathrm{Mix}} = \max_{\sigma\in\Omega} \min\left\{t > 0 : \|P^t(\sigma, \cdot) - \mu\|_{\mathrm{TV}} \leq 1/\mathrm{e}\right\} \ .$$

Our focus is on the mixing time of the aforementioned Markov chain.

Finally, for $k > 1$ we let

$$\beta_c(k) = \frac{\sqrt{2\pi}}{k} \ . \tag{4}$$

Theorem 1 *For any $\varepsilon \in (0, 1)$, there exists $d_0 = d_0(\varepsilon) \geq 1$, such that for $d \geq d_0$, for $\beta \leq (1 - \varepsilon)\beta_c(d)$, there is a constant $C > 0$ such that the following is true:*

Let $\boldsymbol{G} = \boldsymbol{G}(n, d/n)$, while let μ be the 2-spin model on $\boldsymbol{G}$, at inverse temperature β. Then, with probability $1 - n^{-1/4}$ over the instances of $\boldsymbol{G}$ and $\boldsymbol{\mathcal{J}}$, the Glauber dynamics on μ exhibits mixing time

$$\mathrm{T}_{\mathrm{Mix}} \leq C \cdot n^{2+\frac{3}{\log^2 d}} \ .$$

A couple of remarks are in order. First, since we take d to be sufficiently large, rather than using the condition in (2), we equivalently write $\beta < \beta_c(d)$ in Theorem 1. This leads to cleaner derivations.

Second, we note that Theorem 1 continues to hold if we replace the Gaussian couplings with couplings drawn from an arbitrary *sub-Gaussian* distribution. To see this, notice that if J' is sub-Gaussian, and J is a zero-mean Gaussian, there exist a constant $c \geq 0$, such that for all $t > 0$ we have that $\Pr[|J'| \geq t] \leq c \cdot \Pr[|J| \geq t]$. Therefore, via a standard coupling, we can reduce any sub-Gaussian model to the 2-spin model we consider here.

2. Approach

In this section, we give a high-level overview of the construction we use to prove Theorem 1. Note that the construction relies on a good number of potentials.

2.1. Aggregate Influence

Consider the graph $\boldsymbol{G} = \boldsymbol{G}(n, d/n)$, and the vector of real numbers $\boldsymbol{\mathcal{J}} = \{\boldsymbol{J}_e : e \in E(\boldsymbol{G})\}$, where $\boldsymbol{J}_e$'s are i.i.d. standard Gaussians, i.e., $\mathcal{N}(0, 1)$. For $\beta > 0$, consider also the 2-spin model $\mu = \mu_{\boldsymbol{G},\boldsymbol{\mathcal{J}},\beta}$ on $\{\pm 1\}^{V_n}$, that is,

$$\mu(\sigma) \propto \exp\left(\beta \sum\nolimits_{x\sim y} \mathbb{1}\{\sigma(y) = \sigma(x)\} \cdot \boldsymbol{J}_{\{x,y\}}\right), \qquad \forall \sigma \in \{\pm 1\}^{V_n} \ . \tag{5}$$

We usually refer to each $\boldsymbol{J}_e$ as the *coupling* at the edge e in the graph.

For each edge $e \in E(\boldsymbol{G})$ the *influence* Γ_e is defined as

$$\Gamma_e = \frac{|1 - \exp(\beta \boldsymbol{J}_e)|}{1 + \exp(\beta \boldsymbol{J}_e)} \ . \tag{6}$$

A natural way of viewing the influence Γ_e is as a measure of correlation decay over the edge e. Note that the influence is a very natural quantity to consider and emerges in many different contexts such as Chen et al. (2021); Efthymiou (2022); Efthymiou and Zampetakis (2023), just to mention a few.

To get some intuition, note that $\Gamma_e \in [0, 1]$ is an increasing function of $|J_e|$. For a "heavy edge", i.e., an edge e that $|J_e|$ is "big", the corresponding influence is close to 1. On the other hand, for a "light edge", i.e., when $|J_e|$ is "small", the corresponding influence is close to 0.

Given the set of influences $\{\Gamma_e\}_{e \in E(\boldsymbol{G})}$, for every vertex $w \in V(\boldsymbol{G})$, we define its *aggregate influence*, $\mathsf{A}(w)$, with

$$\mathsf{A}(w) = \sum\nolimits_{z \sim w} \Gamma_{\{w,z\}} \ . \tag{7}$$

The quantities $\mathsf{A}(w)$'s play a key role in the analysis, as we use them to construct the blocks.

The choice of parameters in Theorem 1, implies that for each vertex w, we have that $\mathbb{E}[\mathsf{A}(w)] = (1 - \varepsilon)$, where the expectation is with respect to both the degree of w, and the couplings at its incident edges.

In the full version of this work, Efthymiou and Zampetakis (2024), we show that $\mathsf{A}(w)$ is well-concentrated, i.e., for any fixed $\delta > 0$, we have that

$$\Pr[\mathsf{A}(w) > \mathbb{E}[\mathsf{A}(w)] + \delta] \leq \exp(-\Omega(d)) \ .$$

In our analysis, we would have liked each vertex in $\boldsymbol{G}$ to have aggregate influence < 1. From the above tail bound, we expect to have many, but relatively rare, heavy vertices in $\boldsymbol{G}$, i.e., vertices with aggregate influence > 1. Note that, we typically have $\mathrm{poly}(n)$ many vertices each of aggregate influence as huge as $\Theta(\log^{4/3} n)$.

2.2. Path Weights

We introduce a weighting-scheme for the vertices of the graph $\boldsymbol{G}$ that uses $\mathsf{A}(w)$'s. There are two parameters $d, \delta > 0$ in the scheme, where $d > 0$ is a large number, while $\delta \in (0, 1)$. Each vertex w is assigned weight $\mathsf{M}(w)$ defined by

$$\mathsf{M}(w) = \begin{cases} 1 - \delta/2 & \text{if } \mathsf{A}(w) \leq 1 - \delta, \\ d \cdot \mathsf{A}(w) & \text{otherwise.} \end{cases} \tag{8}$$

Given the above weights for the vertices, we define a weight for each path in $\boldsymbol{G}$. Specifically, the path $P = (v_0, \ldots, v_\ell)$ is assigned weight $\mathsf{M}(P)$ defined by

$$\mathsf{M}(P) = \prod_{i=0}^{\ell} \mathsf{M}(v_i) \ .$$

Subsequently, we introduce the notion of *block-vertices* in $\boldsymbol{G}$. A vertex w is called block-vertex if every path P of length at most $\log n$ that emanates from w is "light", i.e., it has weight $\mathsf{M}(P) < 1$. Intuitively, w being a block-vertex implies that $\mathsf{A}(w) < 1$, while every heavy vertex v, i.e., having $\mathsf{A}(v) > 1$, needs to be far from w.

For the range of the parameters we consider here, it turns out that there is a plethora of block-vertices in $\boldsymbol{G}$. Specifically, we show the following result (see the full version Efthymiou and Zampetakis (2024)):

Let $\mathcal{P}$ be the set of paths P in $\boldsymbol{G}$ of length $|P| = \frac{\log n}{\sqrt{d}}$, such that there is no block-vertex in P. Then, we have that

$$\Pr[\mathcal{P} \text{ is empty}] = 1 - o(1) \ . \tag{9}$$

Establishing the above, is a major technical challenge in this wprk, as the quantities we consider are inherently quite involved.

Below, we show how we use (9) for the block construction.

2.3. Block construction

The aim is to obtain a block partition $\mathcal{B} = \{B_1, B_2, \ldots, B_N\}$ such that each $B \in \mathcal{B}$ is small, simply structured, while $\partial_{\text{out}} B$, the outer boundary of B, consists exclusively of block vertices. Note that $\partial_{\text{out}} B$ is the set of vertices outside B which have a neighbour inside the block.

Let us give a high-level description of how $\mathcal{B}$ looks like.

Recall that, typically, $\boldsymbol{G}$ is locally tree-like, however, there are some relatively rare short cycles, i.e., cycles of length less than $4\frac{\log n}{\log^4 d}$, which are far apart from each other.

Each block $B \in \mathcal{B}$ can be one of the following

1. single vertex,
2. a tree,
3. a unicyclic graph.

If B consists of a single vertex, then this vertex must be a block-vertex. If B is multi-vertex, then $\partial_{\text{out}} B$ consists of block-vertices. This, somehow, guarantees that the heavy vertices are hidden deep inside the blocks. The unicyclic blocks contain only short cycles.

Intuitively, (9) guarantees that the blocks are not extensive structures. Note that, for a heavy vertex w, we can reach the boundary of its block, by following any path of length as small as $\frac{\log n}{\sqrt{d}}$ that emanates from w.

Compared to the blocks we have in Efthymiou (2014); Efthymiou et al. (2018b) the ones we get here are quite different. Note that the actual structure of the blocks depends on both graph $\boldsymbol{G}$ and the 2-spin model $\mu_{\boldsymbol{G},\mathcal{J},\beta}$ on this graph. Furthermore, here, it is typical to see single vertex blocks consisting of a high-degree vertex, i.e., degree $\gg d$, or having multi-vertex blocks whose vertices are all low-degree ones, i.e., degree $\approx d$. Of course, this has to do with the fact that the aggregate influence is a different quantity than the degree of a vertex.

2.4. Rapid Mixing of Block Dynamics

Suppose that we have a block partition $\mathcal{B} = \{B_1, \dots, B_N\}$ as the one we describe in Section 2.3.

We consider the *block dynamics* $(X_t)_{t\geq 0}$ with respect to the set of blocks $\mathcal{B}$. The transition from X_t to X_{t+1} is according to the following steps:

1. Choose uniformly at random a block $B \in \mathcal{B}$.
2. For every vertex w outside B, set $X_{t+1}(w) = X_t(w)$.
3. Draw $X_{t+1}(B)$, the configuration at B, according to the marginal of μ at B, conditional on the vertices in $\partial_{\text{out}} B$ having the configuration specified by X_{t+1}.

At this stage, the goal is to show that $(X_t)_{t\geq 0}$ exhibits mixing time T_{Mix} such that

$$\mathrm{T}_{\text{Mix}} = O(N \log N) \ , \tag{10}$$

where N is the number of blocks in $\mathcal{B}$.

We use path coupling Bubley and Dyer (1997) to show (10). That is, we consider $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$, two copies of the block dynamics. Assume that at some time $t \geq 0$, the configurations X_t and Y_t differ at a single vertex, u^*. It suffices to show that we have one-step contraction i.e., we can couple the two copies of the block dynamics such that the expected distance between X_{t+1} and Y_{t+1} is smaller than that of X_t and Y_t.

Typically, we establish contraction with respect to the *Hamming metric* between two configurations. It turns out that, for block dynamics, this metric is suboptimal. In contrast to the single-site dynamics, when we update a block B that is adjacent to the disagreeing vertex u^*, the number of disagreements grows by the size of B. For this reason, we follow an analysis for path coupling which adapts to the setting of block dynamics.

For each vertex z, we write B_z for the block that z belongs to. Furthermore, let

$$\mathsf{A}_{\text{out}}(z) = \sum_{\substack{z\sim w \\ w\notin B_z}} \Gamma_{\{z,w\}} \ . \tag{11}$$

That is, $\mathsf{A}_{\text{out}}(z)$ is the sum of influences over the edges that connect z with its neighbours outside the block B_z. Typically, if vertex z has at least one neighrbour outside B_z, then $\mathsf{A}_{\text{out}}(z) > n^{-7/3}$.

We let $\mathrm{W} \subseteq V$ be the set that of vertices z such that $\mathsf{A}_{\text{out}}(z) > 0$. We call W the set of *external* vertices. On the other hand, we call *internal* all the vertices in $V \setminus \mathrm{W}$. Note that the internal vertices have no neighbours outside their block and hence $\mathsf{A}_{\text{out}}(z) = 0$.

For the path coupling, we introduce the following distance metric for any two $\sigma, \tau \in \{\pm 1\}^V$

$$\texttt{dist}(\sigma,\tau) = \sum_{z\in V\setminus \mathrm{W}} \mathbb{1}\{z \in \sigma \oplus \tau\} + n^4 \cdot \sum_{z\in \mathrm{W}} \mathsf{A}_{\text{out}}(z) \cdot \mathbb{1}\{z \in \sigma \oplus \tau\} \ ,$$

where $\sigma \oplus \tau$ is the set of vertices $w \in V$ that the two configurations disagree, i.e., it consists of the vertices w such that $\sigma(w) \neq \tau(w)$.

The above metric assigns completely different weights to the internal and external vertices, respectively. If some vertex z is internal, then its disagreement gets (tiny) weight 1. On the other hand, if z is external, its disagreement gets weight which is equal to $n^4 \times \mathsf{A}_{\text{out}}(z) \gg 1$. Particularly, we have that $n^4 \times \mathsf{A}_{\text{out}}(z) = \Omega(n^{4/3})$, for all external vertices z.

The above metric essentially captures that the disagreements that do matter in the path coupling analysis, are those which involve vertices at the boundary of blocks, i.e., external vertices. In particular, the "potential" for an external vertex to spread disagreements to adjacent blocks increases with $\mathsf{A}_{\text{out}}(z)$. Let us remark that this observation was first introduced and exploited in Efthymiou et al. (2018b) in an analogous setting.

In the path coupling analysis, we also exploit properties of the block partition $\mathcal{B}$. Particularly, we use that the block vertices are far from the heavy ones, i.e., the vertices that have aggregate influence larger than 1. A heavy vertex in block B, once it becomes disagreeing, tends to create higher than typical number of new disagreements. This is highly undesirable. Having a large distance between the heavy vertices inside B and the boundary $\partial_{\text{out}} B$ implies that the probability of a heavy vertex becoming disagreeing is very small. As a consequence, the overall expected contribution of the heavy vertices becomes negligible.

In light of all the above, we conclude that there is a constant $C > 0$ such that for any vertex $u^* \in V$, for any pair of configurations X_t, Y_t that differ on u^* there is a coupling such that

$$\mathbb{E}\left[\texttt{dist}(X_{t+1}, Y_{t+1}) \mid X_t, Y_t\right] \leq (1 - C/N) \cdot \texttt{dist}(X_t, Y_t) \ . \tag{12}$$

Using path coupling, and arguing that $N = \Omega(\sqrt{n})$, it is standard to obtain (10) from the above inequality. For further details, see the full vestion, Efthymiou and Zampetakis (2024).

2.5. Rapid Mixing for Glauber Dynamics - Comparison

This part is a bit technical. Recall that our aim is to obtain a bound on the mixing time for the single-site Glauber dynamics, whereas so far we only have a bound for the mixing time of the block-dynamics. To this end, we utilise a well-known comparison argument form Martinelli (1999), which relates the relaxation times of the Glauber dynamics and the block dynamics. Recall that the relaxation time of a Markov chain with transition matrix P is equal to $\frac{1}{1-\lambda^*}$, where λ^* is the second largest eigenvalue in magnitude of P.

Letting τ_{rel} and τ_{block} be the relaxation times of the Glauber dynamics and the block dynamics, respectively, from Martinelli (1999), we obtain that

$$\tau_{\text{rel}} \leq \tau_{\text{block}} \cdot \left(\max_{B \in \mathcal{B}} \{\tau_B\}\right) \ , \tag{13}$$

where τ_B is the relaxation time of the (single site) Glauber dynamics on each block $B \in \mathcal{B}$, under worst-case condition σ at the boundary $\partial_{\text{out}} B$.

From the rapid mixing result of block dynamics it is standard to obtain a bound on τ_{block}, hence, it remains to get a bound on τ_B, for all $B \in \mathcal{B}$. Our endeavour to bound τ_B gives rise to

a new weight over the paths in $\boldsymbol{G} = \boldsymbol{G}(n, d/n)$. Specifically, for a path P in $\boldsymbol{G}$, we define the weight

$$\Upsilon(P) = \beta \sum_e |J_e| + \sum_v \log \deg(v) \ . \tag{14}$$

In the first sum, which involves the couplings, the variable e varies over all edges having at least one endpoint in P. The second sum varies over all the vertices in P.

Building on a recursive argument from Efthymiou et al. (2018b); Mossel and Sly (2010), for every tree-like block B rooted at vertex v, we show

$$\tau_B \leq \exp\left(\max_P\{\Upsilon(P)\}\right) \ , \tag{15}$$

where the maximum is over all the paths P in B from the root v to $\partial_{\text{out}}B$. Note that such a path P is at most $\frac{\log n}{\sqrt{d}}$ long. We also get a similar bound for the relaxation time when B is unicyclic.

We show that with probability $1 - o(1)$ over the instances of $\boldsymbol{G}$ and μ, every path P in $\boldsymbol{G}$ of length at most $\frac{\log n}{\log^4 d}$, satisfies $\Upsilon(P) \leq \frac{\log n}{\log^2 d}$. This implies that for every block B we have that

$$\tau_B \leq n^{\frac{3}{\log^2 d}} \ . \tag{16}$$

As mentioned earlier, it is standard to obtain an estimate for τ_{block} from our rapid mixing results for the block-dynamics. Hence, plugging τ_B's and τ_{block} into (13), gives the desired bound on τ_{rel}, the relaxation time for Glauber dynamics. From this point on, it is standard to get Theorem 1. For further details, see the full version Efthymiou and Zampetakis (2024).

Acknowledgments

We are grateful to the anonymous reviewers for their thorough review of our submission, and for their insightful comments and suggestions.

Both authors are supported by the EPSRC New Investigator Award (grant no. EP/V050842/1), and the Center of Discrete Mathematics and Applications (DIMAP), The University of Warwick.

References

Dimitris Achlioptas and Amin Coja-Oghlan. Algorithmic barriers from phase transitions. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008*, pages 793–802. IEEE Computer Society, 2008. URL https://doi.org/10.1109/FOCS.2008.11.

Ahmed El Alaoui and Andrea Montanari. Algorithmic thresholds in mean field spin glasses. *arXiv preprint arXiv:2009.11481*, 2020.

Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Optimization of mean-field spin glasses. *The Annals of Probability*, 49(6):2922 – 2960, 2021. URL https://doi.org/10.1214/21-AOP1519.

Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Sampling from the sherrington-kirkpatrick gibbs measure via algorithmic stochastic localization. In *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022*, pages 323–334. IEEE, 2022.

Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. In *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020*, pages 1319–1330. IEEE, 2020.

Nima Anari, Frederic Koehler, and Thuy-Duong Vuong. Trickle-down in localization schemes and applications. *STOC 2024 - To appear*, 2024.

Roland Bauerschmidt and Thierry Bodineau. A very simple proof of the lsi for high temperature spin systems. *Journal of Functional Analysis*, 276(8):2582–2588, 2019.

Ivona Bezáková, Andreas Galanis, Leslie Ann Goldberg, and Daniel Stefankovic. Fast sampling via spectral independence beyond bounded-degree graphs. In *49th International Colloquium on Automata, Languages, and Programming, ICALP 2022*, volume 229, pages 21:1–21:16, 2022.

Russ Bubley and Martin Dyer. Path coupling: A technique for proving rapid mixing in markov chains. In *Proceedings 38th Annual Symposium on Foundations of Computer Science*, pages 223–231. IEEE, 1997.

Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Optimal mixing for two-state anti-ferromagnetic spin systems. In *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022*, pages 588–599. IEEE, 2022.

Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of glauber dynamics: entropy factorization via high-dimensional expansion. In *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 1537–1550. ACM, 2021.

Zongchen Chen, Nitya Mani, and Ankur Moitra. From algorithms to connectivity and back: Finding a giant component in random k-sat. In *Proceedings of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023*, pages 3437–3470. SIAM, 2023. URL https://doi.org/10.1137/1.9781611977554.ch132.

Amin Coja-Oghlan and Charilaos Efthymiou. On independent sets in random graphs. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011*, pages 136–144. SIAM, 2011. URL https://doi.org/10.1137/1.9781611973082.12.

Amin Coja-Oghlan, Charilaos Efthymiou, Nor Jaafari, Mihyun Kang, and Tobias Kapetanopoulos. Charting the replica symmetric phase. *Communications in Mathematical Physics*, 359:603–698, 2018.

Martin Dyer, Abraham D Flaxman, Alan M Frieze, and Eric Vigoda. Randomly coloring sparse random graphs with fewer colors than the maximum degree. *Random Structures & Algorithms*, 29(4):450–465, 2006.

Martin E. Dyer and Alan M. Frieze. Randomly coloring random graphs. *Random Struct. Algorithms*, 36(3):251–272, 2010. URL https://doi.org/10.1002/rsa.20286.

Charilaos Efthymiou. MCMC sampling colourings and independent sets of $G(n, d/n)$ near uniqueness threshold. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014*, pages 305–316. SIAM, 2014.

Charilaos Efthymiou. On Sampling Symmetric Gibbs Distributions on Sparse Random Graphs and Hypergraphs. In *49th International Colloquium on Automata, Languages, and Programming, ICALP 2022, July 4-8, 2022*, volume 229, pages 57:1–57:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022.

Charilaos Efthymiou and Weiming Feng. On the mixing time of glauber dynamics for the hard-core and related models on g(n, d/n). *CoRR*, abs/2302.06172, 2023. URL https://doi.org/10.48550/arXiv.2302.06172.

Charilaos Efthymiou and Kostas Zampetakis. Broadcasting with random matrices. *CoRR*, abs/2302.11657, 2023. doi: 10.48550/arXiv.2302.11657. URL https://doi.org/10.48550/arXiv.2302.11657.

Charilaos Efthymiou and Kostas Zampetakis. On sampling diluted spin glasses using glauber dynamics. *CoRR*, abs/2403.08921, 2024. URL https://arxiv.org/abs/2403.08921.

Charilaos Efthymiou, Thomas P. Hayes, Daniel Stefankovic, and Eric Vigoda. Sampling random colorings of sparse random graphs. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018*, pages 1759–1771. SIAM, 2018a.

Charilaos Efthymiou, Thomas P Hayes, Daniel Štefankovič, and Eric Vigoda. Sampling random colorings of sparse random graphs. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1759–1771. SIAM, 2018b.

Ronen Eldan, Frederic Koehler, and Ofer Zeitouni. A spectral condition for spectral gap: fast mixing in high-temperature ising models. *Probability theory and related fields*, 182(3-4):1035–1051, 2022.

Silvio Franz, Michele Leone, Federico Ricci-Tersenghi, and Riccardo Zecchina. Exact solutions for diluted spin glasses and optimization problems. *Physical review letters*, 87(12):127209, 2001.

Andreas Galanis, Daniel Stefankovic, and Eric Vigoda. Inapproximability for antiferromagnetic spin systems in the tree nonuniqueness region. *J. ACM*, 62(6):50:1–50:60, 2015.

David Gamarnik and Madhu Sudan. Performance of sequential local algorithms for the random NAE-K-SAT problem. *SIAM J. Comput.*, 46(2):590–619, 2017.

David Gamarnik, Aukosh Jagannath, and Alexander S. Wein. Low-degree hardness of random optimization problems. In *61st IEEE Annual Symposium on Foundations of Computer Science, FOCS 2020*, pages 131–140. IEEE, 2020. URL https://doi.org/10.1109/FOCS46700.2020.00021.

Francesco Guerra and Fabio Lucio Toninelli. The high temperature region of the Viana-Bray diluted spin glass model. *Journal of statistical physics*, 115:531–555, 2004.

Frederic Koehler, Holden Lee, and Andrej Risteski. Sampling approximately low-rank ising models: MCMC meets variational methods. In *Conference on Learning Theory, 2022*, volume 178 of *Proceedings of Machine Learning Research*, pages 4945–4988. PMLR, 2022.

Kuikui Liu, Sidhanth Mohanty, Amit Rajaraman, and David X Wu. Fast mixing in sparse random ising models. *arXiv preprint arXiv:2405.06616*, 2024.

Fabio Martinelli. Lectures on glauber dynamics for discrete spin models. *Lectures on probability theory and statistics (Saint-Flour, 1997)*, 1717:93–191, 1999.

M. Mézard, G. Parisi, and M. Virasoro. *Spin glass theory and beyond*. World Scientific, 1987.

Elchanan Mossel and Allan Sly. Gibbs rapidly samples colorings of g (n, d/n). *Probability theory and related fields*, 148(1-2):37–69, 2010.

Elchanan Mossel and Allan Sly. Exact thresholds for ising–gibbs samplers on general graphs. *The Annals of Probability*, 41(1):294–328, 2013.

Dmitry Panchenko. The parisi ultrametricity conjecture. *Annals of Mathematics*, pages 383–393, 2013.

Dmitry Panchenko and Michel Talagrand. Bounds for diluted mean-fields spin glass models. *Probability Theory and Related Fields*, 130:319–336, 2004.

Giorgio Parisi. Infinite number of order parameters for spin-glasses. *Physical Review Letters*, 43 (23):1754, 1979.

David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Physical review letters*, 35(26):1792, 1975.

Allan Sly and Nike Sun. The computational hardness of counting in two-spin models on d-regular graphs. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012*, pages 361–369. IEEE Computer Society, 2012. URL https://doi.org/10.1109/FOCS.2012.56.

Daniel L Stein and Charles M Newman. *Spin glasses and complexity*, volume 4. Princeton University Press, 2013.

Michel Talagrand. The Parisi formula. *Ann. Math. (2)*, 163(1):221–263, 2006. ISSN 0003-486X. doi: 10.4007/annals.2006.163.221.