Glauber dynamics for Ising models on random regular graphs: cut-off and metastability

Van Hao Can, Remco van der Hofstad and Takashi Kumagai

Research Institute for Mathematical Sciences, Kyoto University
Kyoto 606-8502, Japan
& Institute of Mathematics, Vietnam Academy of Science and Technology
18 Hoang Quoc Viet, 10307 Hanoi, Vietnam
E-mail address: cvhao89@gmail.com
URL: https://sites.google.com/site/vanhaocan/home

Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
E-mail address: r.w.v.d.hofstad@tue.nl
URL: https://www.win.tue.nl/~rhofstad/

Research Institute for Mathematical Sciences, Kyoto University
Kyoto 606-8502, Japan
E-mail address: kumagai@kurims.kyoto-u.ac.jp
URL: https://www.kurims.kyoto-u.ac.jp/~kumagai/

Abstract. Consider random d-regular graphs, i.e., random graphs such that there are exactly d edges from each vertex for some $d \geq 3$. We study both the configuration model version of this graph, which has occasional multi-edges and self-loops, as well as the simple version of it, which is a d-regular graph chosen uniformly at random from the collection of all d-regular graphs.

In this paper, we discuss mixing times of Glauber dynamics for the Ising model with an external magnetic field on a random d-regular graph, both in the quenched as well as the annealed settings. Let β be the inverse temperature, β_c be the critical temperature and B be the external magnetic field. Concerning the annealed measure, we show that for $\beta > \beta_c$ there exists $\hat{B}_c(\beta) \in (0, \infty)$ such that the model is *metastable* (i.e., the mixing time is exponential in the graph size n) when $\beta > \beta_c$ and $0 \leq B < \hat{B}_c(\beta)$, whereas it exhibits the cut-off phenomenon at $c_\star n \log n$ with a window of order n when $\beta < \beta_c$ or $\beta > \beta_c$ and $B > \hat{B}_c(\beta)$. Interestingly, $\hat{B}_c(\beta)$ coincides with the critical external field of the Ising model on the d-ary tree (namely, above which the model has a unique Gibbs measure). Concerning the quenched measure, we show that there exists $B_c(\beta)$ with $B_c(\beta) \leq \hat{B}_c(\beta)$ such that for $\beta > \beta_c$, the mixing time is at least exponential along some subsequence $(n_k)_{k \geq 1}$ when $0 \leq B < B_c(\beta)$, whereas it is less than or equal to $Cn \log n$ when $B > \hat{B}_c(\beta)$. The quenched results also hold for the model conditioned on simplicity, for the annealed results this is unclear.

1. Introduction

The Ising model is a paradigmatic model in statistical mechanics. It was invented by Ising and his PhD-supervisor Lenz to model magnetism, for which it was considered on regular lattices (see Niss, 2005, 2009 for the interesting history of the Ising model, as well as the standard books Bovier, 2006; Ellis, 2006 and the references therein). With the view that the Ising model can also model cooperative behavior, its relevance in the area of network science has increased, and the literature on the Ising model on random graphs, invented to model complex networks, has exploded. See Dorogovtsev et al. (2008) for the physics perspective on critical phenomena on random graphs. Initially, the focus was on establishing the thermodynamic limit of the quenched Ising model on random graphs (Dembo and Montanari, 2010b,a; Dommers et al., 2010) as well as on its critical behavior (Dommers et al., 2014; Giardinà et al., 2015). In the past years, also the annealed Ising model has attracted considerable attention Can (2019, 2017); Can et al. (2021+); Dommers et al. (2016); Giardinà et al. (2016). As we explain in more detail below, the quenched and annealed settings for the Ising model describe different physical realities in the dynamics of the underlying graph and the Ising model on it. The local weak limit of the Ising model on locally tree-like random graphs was studied by Basak and Dembo (2017); Montanari et al. (2012).

Recently, the *dynamical* properties of the Ising model have attracted attention, focusing on its metastable behavior Bovier et al. (2019); Dommers (2017); Dommers et al. (2017); den Hollander and Jovanovski (2021) in two frameworks. In the first one (see Dommers, 2017; Dommers et al., 2017), large random graphs were considered, on which an Ising model lives with a slightly positive field at very low temperature. In such settings, the all-plus state is the ground state, and thus the most likely state for the system to be in. However, due to the strong coupling, the all-minus state is a metastable state, and it takes the system a very long time to leave this state when started from it. For large network size, the main results of Dommers (2017); Dommers et al. (2017) give detailed estimates for the transition time to move from the metastable state to the stable state, as the temperature tends to zero, for graphs of fixed large size. In the second framework, the temperature of the system is fixed and we are interested in what happens as the system size tends to infinity The concrete analyses of metastable hitting time for the Glauber dynamics on dense Erdős-Rényi random graphs have been given by Bovier et al. (2019); den Hollander and Jovanovski (2021), by using a potential theoretic approach as discussed in detail by Bovier and den Hollander (2015).

In this paper, we follow the second framework. However, rather than treating it as a *metastable* system, we approach it as a Markov chain mixing-time problem. Our paper provides the first results dealing with *sparse random graphs* instead of dense ones. In general, it is believed that for supercritical temperature, mixing is fast (mixing time of order $n \log n$ where n is the size of the graph), see Mossel and Sly (2013), while for subcritical temperatures and small external fields, the mixing time is exponentially large in the graph size. We focus on both the quenched as well as the annealed Ising model on random regular graphs, where our results are the strongest in the annealed setting. In particular, our main results and innovations are as follows:

(a) For subcritical temperatures, where the Ising model at zero external field has two Gibbs measures, we identify the critical value for the field in the annealed setting. More precisely, for large field, mixing is rapid (mixing time of order $n \log n$), while for small field, mixing is slow (mixing time exponential in the system size). The latter corresponds to the metastable setting. The proofs rely on a close relation between the annealed Ising model and birth-death chains, for which such results have been established by Barrera et al. (2009); Chen and Saloff-Coste (2013); Ding et al. (2009a,b); Levin et al. (2010).

(b) For the quenched Ising model, we prove similar properties, however, we are not able to identify the *exact* critical value of the external magnetic field, but instead resort to bounds on it. The results rely on proofs of mixing times for Glauber dynamics on general graphs

such as proved in the literature, and most effort is in proving that the necessary conditions hold for degree-regular configuration models.

(c) While the mixing time is at least $e^{\lambda n}$ in both models for some appropriate $\lambda > 0$ in the slow-mixing regime, in the annealed setting, the constant λ is *independent* of β, while in the quenched setting, the constant λ is *linear* in β for large β.

The remainder of this section is organized as follows. We start in Section 1.1 by defining the random regular graphs that we will be working on. In Section 1.2, we define the Ising model on random graphs in its quenched and annealed settings. In Section 1.3, we recall some previous results on the Ising model proved by the first author Can (2019) in the annealed setting and by Dembo and Montanari (2010b) in the quenched setting. In Section 1.4, we state our main results. We close in Section 1.5 with discussion and open problems.

1.1. *Random regular graphs.*

Let us start by defining the configuration model introduced by Bollobás (1980) in the degree-regular case, and by Molloy and Reed (1995) in the general degree case. We consider a sequence $(G_n)_{n \geq 1}$ of such graphs. To define it, start for each n with the vertex set $[n] = \{1, \ldots, n\}$. Construct the edge set as follows. Consider a sequence of degrees $(d_i)_{i \in [n]}$ and assume that $\ell_n = \sum_{i \in [n]} d_i$ is even. For each vertex $i \in [n]$, start with d_i half-edges incident to i. Denote the set of all the half-edges by $\mathcal{H}$. Select $h_1 \in \mathcal{H}$ arbitrarily, and then choose a half-edge h_2 uniformly from $\mathcal{H} \setminus \{h_1\}$, and pair h_1 and h_2 to form an edge. Next, select an arbitrarily half-edge $h_3 \in \mathcal{H} \setminus \{h_1, h_2\}$, and pair it to h_4 uniformly chosen from $\mathcal{H} \setminus \{h_1, h_2, h_3\}$. Continue this procedure until there are no more half-edges. The resulting graph is called the *configuration model*, see van der Hofstad (2017, Chapter 7) for an extensive introduction. In particular, it is known that the configuration model conditioned on simplicity is a *uniform* random graph with the prescribed degrees $(d_i)_{i \in [n]}$ - see van der Hofstad (2017, Proposition 7.7).

In this paper, we consider the *random d-regular graph*, that is $d_i = d$ for all $i \in [n]$, with $d \geq 3$ and nd assumed to be even. We let $\mathbb{P}$ and $\mathbb{E}$ denote the probability measure and expectation with respect to the random regular graph. We say that a sequence of events $(A_n)_{n \geq 1}$ occurs with high probability (which we abbreviate as whp) if $\mathbb{P}(A_n) = 1 - o(1)$ as $n \to \infty$.

In the degree-regular setting, it is also known that the probability that the configuration model is simple converges to a positive value van der Hofstad (2017, Theorem 7.12) (see also Angel et al., 2019; Bollobás, 1980; Janson, 2009, 2014), which implies that any result that holds whp for the configuration model, also holds whp for the random regular graph. This implies that all the results in the *quenched* setting also hold for the random regular graph. In the annealed setting, this is less obvious, as we are taking expectations with respect to exponential functionals in the Ising model, as we now explain in more detail.

1.2. *Ising model and Glauber dynamics.*

In this section, we define the quenched and annealed Ising models, as well as Glauber dynamics for it.

1.2.1. *Ising model.*

Let $\Omega_n = \{-1, 1\}^n$ be the space of spin configurations. For any spin vector $\sigma = (\sigma_1, \ldots, \sigma_n) \in \Omega_n$, the Hamiltonian is given by

$$H_n(\sigma) = H_n^{\beta,B}(\sigma) = -\beta \sum_{i,j \in [n]:\, i \leq j} k_{i,j} \sigma_i \sigma_j - B \sum_{i \in [n]} \sigma_i,$$

where $k_{i,j}$ is the number of edges between vertices i and j, $\beta \geq 0$ is the inverse temperature and $B \in \mathbb{R}$ is the uniform external magnetic field.

The *quenched* measure is defined by Boltzmann law, defined, for any $\sigma \in \Omega_n$, as

$$\mu_n^{\beta,B}(\sigma) = \frac{\exp(-H_n(\sigma))}{Z_n^{\beta,B}},$$

where $Z_n^{\beta,B}$ is the partition function given by

$$Z_n^{\beta,B} = \sum_{\sigma \in \Omega_n} \exp(-H_n^{\beta,B}(\sigma)).$$

Similarly, the *annealed* measure is defined, for any $\sigma \in \Omega_n$, as

$$\hat{\mu}_n^{\beta,B}(\sigma) = \frac{\mathbb{E}[\exp(-H_n(\sigma))]}{\mathbb{E}[Z_n^{\beta,B}]},$$

where $\mathbb{E}$ is the expectation with respect to the random graph under consideration. These two measures concern different physical realities (see Can et al., 2021+). While the random graph in the quenched measure is fixed, or can be thought of as varying very slowly compared to the Ising Glauber dynamics, in the annealed law, the Glauber dynamics only observes an *average* graph instant, which, by the ergodic theorem, can be thought of as an expectation with respect to the graph randomness. As discussed before, we will work on the d-regular configuration model, but in our discussion we will also discuss more general degree settings. Throughout the paper, we consider the case $B \geq 0$; the other case can be treated identically by symmetry.

1.2.2. *Glauber dynamics.* Let us first recall the Glauber dynamics for a given reversible measure ν_n on Ω_n. Let $(\xi_t)_{t \geq 0}$ be a discrete Markov chain on Ω_n with transitions as follows. Assume that $\xi_t = \sigma$, let I be a random index in $[n]$ chosen uniformly at random. Then,

$$\xi_{t+1} = \begin{cases} \sigma^{+I} & \text{with probability } \frac{\nu_n(\sigma^{+I})}{\nu_n(\sigma^{+I})+\nu_n(\sigma^{-I})}, \\ \sigma^{-I} & \text{with probability } \frac{\nu_n(\sigma^{-I})}{\nu_n(\sigma^{+I})+\nu_n(\sigma^{-I})}, \end{cases} \tag{1.1}$$

where, for all $\sigma \in \Omega_n$ and $i \in [n]$, we define σ^{+i} and σ^{-i} to be the i-spin-flipped version of σ, i.e.,

$$\sigma_j^{+i} = \sigma_j^{-i} = \sigma_j \text{ for } j \neq i, \quad \text{and } \sigma_i^{+i} = -\sigma_i^{-i} = 1.$$

We denote by $(\xi_t^\sigma)_{t \geq 0}$ the Glauber dynamics starting from the configuration σ. It is well known that $(\xi_t)_{t \geq 0}$ is a reversible Markov chain with stationary measure ν_n. Hence, as $t \to \infty$, the distribution of $(\xi_t)_{t \geq 0}$ converges to ν_n. The distance to stationary of the Glauber dynamics is defined as

$$d_n(t) = \max_{\sigma \in \Omega_n} \|\mathbb{P}(\xi_t^\sigma \in \cdot) - \nu_n(\cdot)\|_{\mathrm{TV}},$$

where $\|\mu - \nu\|_{\mathrm{TV}}$ is the total variation distance between the probability measures μ and ν. Then, the *mixing time* is defined as

$$t_{\mathrm{mix}} = \min\{t \colon d_n(t) \leq \tfrac{1}{4}\}.$$

The value $\frac{1}{4}$ is arbitrary and can be replaced by any other value $\varepsilon \in (0,1)$.

In this paper, we study the mixing time with respect to the quenched and annealed measures $\mu_n^{\beta,B}$ and $\hat{\mu}_n^{\beta,B}$ defined previously. From now on, we let t_{mix} be the quenched mixing time for $\mu_n^{\beta,B}$, and $\hat{t}_{\mathrm{mix}}$ the annealed mixing time for $\hat{\mu}_n^{\beta,B}$, respectively. Notice that while t_{mix} is random as it depends on the random graph, $\hat{t}_{\mathrm{mix}}$ is non-random.

For $d \geq 3$, the Ising model on random regular graphs exhibits a phase transition at the critical value $\beta_c = \mathrm{atanh}(1/(d-1))$, see e.g. Dembo and Montanari (2010b). (Note that this model does not exhibit a phase transition for $d = 1, 2$, and that is why we consider $d \geq 3$.) The mixing time has been studied in the high-temperature regime:

Theorem 1.1 (Mixing times high-temperature Ising model, Mossel and Sly, 2013, Theorem 1). *There exists a positive constant C, such that if $\beta < \beta_c$, then whp $t_{\mathrm{mix}} \leq Cn \log n$.*

We notice that the results by Mossel and Sly (2013) hold for the general class of finite graphs with degree bounded by d for any $\beta < \mathrm{atanh}(1/(d-1))$. For us, it is crucial that $\mathrm{atanh}(1/(d-1))$ is the

critical value for the Ising model on the configuration model, so that Theorem 1.1 holds throughout the high-temperature regime.

In this paper, apart from studying the mixing times of the Ising model, we also study its cut-off behavior. This is related to the speed at which $t \mapsto d_n(t)$ decreases. Formally, with $t_{\mathrm{mix}}(\varepsilon) = \min\{t \colon d_n(t) \leq \varepsilon\}$, we say that the cut-off phenomenon occurs at a_n with a window of order $b_n \ll a_n$ when, for all $\varepsilon \in (0, \frac{1}{2})$,

$$t_{\mathrm{mix}}(\varepsilon)/a_n \to 1, \qquad \text{while} \qquad t_{\mathrm{mix}}(\varepsilon) - t_{\mathrm{mix}}(1 - \varepsilon) = \Theta(b_n). \tag{1.2}$$

1.3. *Previous results about Ising models on the configuration model.* In this section, we state some important previous results about the annealed and quenched Ising model that we shall rely upon. We first define the *fixed-spin partition function*, which will play an important role in our results. For $\sigma \in \Omega_n$, we write

$$\sigma_+ = \{i \in [n] \colon \sigma_i = 1\} \qquad \text{and} \qquad \sigma_- = \{i \in [n] \colon \sigma_i = -1\}.$$

For any $t \in (0, 1)$, we define

$$\varphi_n^{\beta, B}(t) = \frac{1}{n} \log Z_n^{\beta, B}(t),$$

where

$$Z_n^{\beta, B}(t) = \sum_{\sigma \colon |\sigma_+| = \lceil nt \rceil} \mathrm{e}^{-H_n(\sigma)},$$

and

$$\hat{\varphi}_n^{\beta, B}(t) = \frac{1}{n} \log \mathbb{E}\big[Z_n^{\beta, B}(t)\big].$$

It has been shown in the proof of Can (2019, Theorem 1.1(i)) that

$$\sup_{0 \leq k \leq n} \left| \frac{1}{n} \log \sum_{\sigma \colon |\sigma_+| = k} \mathbb{E}\left(\mathrm{e}^{-H_n(\sigma)}\right) - \hat{\varphi}_{\beta, B}(k/n) \right| = \mathcal{O}(1/n), \tag{1.3}$$

with

$$\hat{\varphi}_{\beta, B}(t) = \frac{\beta d}{2} - B + I(t) + 2Bt + d \int_0^{t \wedge 1 - t} \log f_\beta(s) ds, \tag{1.4}$$

where $I(0) = I(1) = 0$ and, for $t \in (0, 1)$

$$I(t) = (t - 1) \log(1 - t) - t \log t,$$

and

$$f_\beta(t) = \frac{\mathrm{e}^{-2\beta}(1 - 2t) + \sqrt{1 + (\mathrm{e}^{-4\beta} - 1)(1 - 2t)^2}}{2(1 - t)}. \tag{1.5}$$

Thus, for all $t \in [0, 1]$,

$$\lim_{n \to \infty} \hat{\varphi}_n^{\beta, B}(t) = \hat{\varphi}_{\beta, B}(t).$$

Denote the quenched and annealed pressures by, respectively,

$$\varphi(\beta, B) = \lim_{n \to \infty} \frac{1}{n} \log Z_n^{\beta, B}, \qquad \hat{\varphi}(\beta, B) = \lim_{n \to \infty} \frac{1}{n} \log \mathbb{E}(Z_n^{\beta, B}). \tag{1.6}$$

It is shown in Dembo et al. (2014, Theorem 1) and Can (2019, Theorem 1.1 & Proposition 3.2) that the annealed and quenched pressures are equal and have a variational expression as

$$\varphi(\beta, B) \overset{\mathrm{a.s.}}{=} \hat{\varphi}(\beta, B) = \max_{t \in [0, 1]} \hat{\varphi}_{\beta, B}(t). \tag{1.7}$$

These results will be crucial to establish the 'energy landscape' of Glauber dynamics for the Ising model.

1.4. *Main results.* In this section, we state our main results. We start with the results for the annealed Ising model, followed by our results on the quenched Ising model.

1.4.1. *Main results for the annealed Ising model.* For the annealed case with $\beta > \beta_c$, we can show (see Lemma 2.2 below) that there is a threshold $\hat{B}_c = \hat{B}_c(\beta)$, such that if $B > \hat{B}_c$, then the function $t \mapsto \hat{\varphi}_{\beta,B}(t)$ is unimodular, i.e., $t \mapsto \hat{\varphi}_{\beta,B}(t)$ has only one critical point which is the maximizer characterizing the annealed pressure in (1.7). On the other hand, if $0 \leq B < \hat{B}_c$, then $t \mapsto \hat{\varphi}_{\beta,B}(t)$ has three critical points (one local maximizer, one global maximizer and one local minimizer), and thus the graph of $t \mapsto \hat{\varphi}_{\beta,B}(t)$ has a valley. This particular observation suggests that for $\beta > \beta_c$, there is a phase transition in the mixing time of the annealed Ising model when B crosses the value $\hat{B}_c$. Indeed, we will show in Theorem 1.2 below that for $\beta > \beta_c$, the mixing time increases exponentially in the graph size when $B < \hat{B}_c$ but it is of order $n \log n$ when $\beta > \beta_c, B > \hat{B}_c$, or when $\beta \in [0, \beta_c)$.

We observe that $\hat{\varphi}_n^{\beta,B}(t) = \hat{\varphi}_n^{\beta,0}(t) + B(2t-1)$ and thus $\hat{\varphi}_{\beta,B}(t) = \hat{\varphi}_{\beta,0}(t) + B(2t-1)$. Therefore, the unimodularity of $\hat{\varphi}_{\beta,B}$ is strongly related to the reflection point of $t \mapsto \hat{\varphi}_{\beta,0}(t)$. Indeed, we will show in Lemma 2.2 below that

$$\hat{B}_c = \hat{B}_c(\beta) := -\frac{1}{2}\hat{\varphi}'_{\beta,0}(t_u) = \sup_{t \in (0,\frac{1}{2})} \left(-\frac{1}{2}\hat{\varphi}'_{\beta,0}(t) \right) = \sup_{0<s<t<\frac{1}{2}} \frac{\hat{\varphi}_{\beta,0}(s) - \hat{\varphi}_{\beta,0}(t)}{2(t-s)}, \qquad (1.8)$$

where t_u is the reflection point of $\hat{\varphi}_{\beta,0}$ determined as in (2.2). With this notation in hand, we now state our main result for the Glauber dynamics on the annealed Ising model on the d-regular configuration model:

Theorem 1.2 (Annealed Glauber dynamics). *Consider the annealed Glauber dynamics on the d-regular configuration model.*

(i) *For $\beta > \beta_c$ and $0 \leq B < \hat{B}_c$, there exist positive constants C and λ such that*

$$C^{-1} \exp(\lambda n) \leq \hat{t}_{\mathrm{mix}} \leq Cn^4 \exp(\lambda n).$$

Moreover,

$$\sup\{\lambda \colon \beta > \beta_c, 0 \leq B < \hat{B}_c\} < \infty. \qquad (1.9)$$

(ii) *For $\beta < \beta_c$ or $\beta > \beta_c$ but $B > \hat{B}_c$, there exists a positive constant $c_\star$, such that the cut-off phenomenon occurs at $c_\star n \log n$ with a window of order n.*

Independence of inverse temperature in (1.9)**.** The fact that the constant λ in the exponential growth of the mixing time in Theorem 1.2(i) is independent of β for large β in (1.9) is rather remarkable. It can be understood as follows. Think of the curve $t \mapsto \hat{\varphi}_{\beta,B}(t)$ as an energy landscape. It takes an exponential amount of time to cross any energy barrier, meaning a difference in energy, so that it takes time of order $e^{n\lambda}$ to cross an energy barrier λ. For the annealed Ising model, $\hat{\varphi}_{\beta,B}(s)$ acts as the energy of a configuration with roughly ns plus spins, where $s \in (0,1)$. Then, let us start from a configuration σ for which $|\sigma^+| \approx sn$ where $s \neq t^\star$ and $t^\star$ is such that $\hat{\varphi}(\beta, B) = \hat{\varphi}_{\beta,B}(t^\star)$ (recall (1.7)). Let $B > 0$, so that $t^\star > \frac{1}{2}$. Then, the amount of time to go from $|\sigma^+| \approx sn$ to equilibrium is close to $e^{n\lambda(s)}$, where

$$\lambda(s) = \sup_{t \in [s,t^\star]} \left(\hat{\varphi}_{\beta,B}(t^\star) - \hat{\varphi}_{\beta,B}(t) \right). \qquad (1.10)$$

It can be expected that the worst case for this is when $s = s^\star$ for some specific $s^\star < \frac{1}{2}$, which suggests that $\lambda = \lambda(s^\star)$. Then, to go between any σ with $|\sigma^+| \approx s^\star n$ and the stationary distribution (having approximately $t^\star n$ with $t^\star > \frac{1}{2}$ plus spins), the dynamics has to pass through a spin configuration σ' with roughly $n/2$ plus spins.

In many settings, for example also in the quenched setting, this will lead to an energy difference that grows linearly in n as well as in β for large β, since the number of edges between any two disjoint sets of size approximately $n/2$ will be *linear* in n for n large, so that the energy barrier will be of order $n\beta$. However, the annealed dynamics on the configuration model is special, as it allows to update the graph as well when changing the spins, which makes the energy barrier of order n with a constant that is *independent* of β. Indeed, there exists a graph configuration whose probability is exponentially small in n but with an exponential rate that is independent of β, and a partition of the vertices in two sets each of size approximately $n/2$, for which the number of edges between these two sets equals zero. This can be seen by noting that the probability of splitting the d-regular random graph into *two* disjoint d-regular graphs of about equal size is exponentially small in n, with an exponential rate that is obviously independent of β. By then taking all spins to be plus on one part, and all spins to be minus on the other part, we see that we have a roughly equal number of plusses and minuses, while at the same time having an exponentially small cost whose exponential rate is *independent of β*.

The above argument implies that $\hat{\varphi}_{\beta,B}(\frac{1}{2}) \le C$ with C independent of β, which, in particular, suggests also that $\lambda(s^\star) \le C$ where C is independent of n. This explains (1.9).

Cut-off in subcritical regimes. In the setting where $\beta < \beta_c$, or $\beta > \beta_c$ but $B > \hat{B}_c$, there is no valley in the energy landscape, meaning that the dynamics will move quickly from any spin configuration to the stationary distribution. The fact that Theorem 1.2(ii) proves that this dynamics satisfies a cut-off phenomenon is a substantial improvement from the general result in Theorem 1.1, which is a restatement of Mossel and Sly (2013, Theorem 1), however, it is restricted to the annealed setting.

1.4.2. *Main results for the quenched Ising model.* We next state our results for the quenched Glauber dynamics. We start by investigating the fixed-spin partition function:

Proposition 1.3 (Quenched fixed-spin partition function). *The following assertions hold.*
   (i) *For all β, B and $t \in (0,1)$*

$$\varphi_n^{\beta,B}(t) - \tilde{\varphi}_n^{\beta,B}(t) \xrightarrow{a.s.} 0,$$

   where $\tilde{\varphi}_n^{\beta,B}(t) = \mathbb{E}\big[\varphi_n^{\beta,B}(t)\big]$.
   (ii) *There exists a subsequence $(n_k)_{k \ge 1}$, such that for all β, B the sequence of functions $(\tilde{\varphi}_{n_k}^{\beta,B})_{k \ge 1}$ converges uniformly to a continuous (non-random) function $\varphi_{\beta,B}$ in every compact subset of $[0,1]$. Moreover,*

$$\varphi_{\beta,B}(t) = \varphi_{\beta,0}(t) + B(2t-1).$$

   Furthermore, let

$$B_c = B_c(\beta) := \sup_{0<s<t<\frac{1}{2}} \frac{\varphi_{\beta,0}(s) - \varphi_{\beta,0}(t)}{2(t-s)}. \tag{1.11}$$

   Then $B_c \in (0, \beta d)$ for all $\beta > \beta_c$. Moreover, there exists a positive constant c such that $B_c \ge c\beta$ if $\beta \ge 12$.

It follows directly from the definition of B_c that when $B < B_c$, the function $t \mapsto \varphi_{\beta,B}(t)$ is non-unimodular and it has a valley; otherwise it is unimodular. While it is not difficult to prove that the mixing time is of exponential order when $B < B_c$, it is not clear to us how to show that the mixing time is of $\exp(o(n))$ when $B > B_c$. Instead, we can show the following. Let B_c^G be the critical external field of Ising model on d-ary tree defined by

$$B_c^G = \inf\{B \ge 0 \colon \text{ Ising model on infinite } d\text{-ary tree has a unique Gibbs measure}\}. \tag{1.12}$$

Then, the mixing time is of logarithmic order for $B > B_c^G$:

Theorem 1.4 (Quenched Glauber dynamics). *Consider the quenched Glauber dynamics on the d-regular configuration model. There exist positive constants c_1, c, C where c, C may depend on β, B and c_1 is independent of β, B, such that the following statements hold.*

(ia) *If $\beta > \beta_c$ and $0 \leq B < B_c$ then*

$$\limsup_{n \to \infty} \mathbb{P}(t_{\mathrm{mix}} \geq \exp(cn)) = 1.$$

(ib) *If $\beta > \beta_c$ and $0 \leq B < c_1 \beta$ then $t_{\mathrm{mix}} \geq \exp(c_1 \beta n)$ whp.*

(ii) *If $\beta > \beta_c$ and $B > B_c^G$, then $t_{\mathrm{mix}} \leq Cn \log n$ whp.*

The same results hold for the d-regular random graph.

It follows directly from our results that $B_c \leq B_c^G$. Surprisingly, we can show that $B_c^G = \hat{B}_c$. This means that the critical magnetic field for the annealed Ising model on G_n is the same as that of the Ising model on the d-ary tree, which equals the local limit of G_n:

Proposition 1.5 (Identification of B_c^G). *For any $\beta > \beta_c$, $B_c^G(\beta) = \hat{B}_c(\beta)$.*

However, we do not know whether the quenched quantity B_c is equal to B_c^G or not. In particular, we do not know the quenched behavior of t_{mix} when $B_c \leq B \leq B_c^G$.

1.5. *Discussion.* Here we give some comments on our results, and state some open problems.

Metastability for Ising models on configuration models. There are some related works on the metastability of Glauber dynamics on random graphs at zero temperature. Dommers (2017); Dommers et al. (2017) study the Ising model on configuration models. They show that the hitting time to all-plus configuration for the dynamics starting from all minus (which we denote by T_n) grows exponentially fast at zero temperature, i.e., $\lim_{\beta \to \infty} \mu_n^{\beta, B}(T_n \geq \exp(c\beta n)) = 1$ whp for the random graph.

Cut-off for quenched setting. In Theorem 1.2(ii), we show the cut-off phenomenon for the dynamics under the annealed law in the high-temperature regime and in the low-temperature regime with high external field. We guess that the same phenomenon occurs in the quenched setting. As far as we know, the best result is due to Lubetzky and Sly (2017), who prove the cut-off phenomenon for the Ising model on graphs with bounded degrees at sufficiently high temperature (more precisely, $\beta \leq \varepsilon/\Delta$ with Δ the maximal degree and ε a universal constant).

Extension to configuration models with general degrees. It is natural to extend our results to the setting of Ising models on configuration models with general degrees. However, it is not immediate how to appropriately define the critical external fields. One may be tempted to conjecture that the definitions in (1.8) and (1.11) are still the right critical values for B. However, in the non-regular case, this is quite unclear. Indeed, to our best knowledge there is no result for the critical external field B_c^{GW} for the uniqueness of Gibbs measure on a Galton-Watson tree (the weak limit of configuration model), see also (1.12). Thus, extensions of our results to non-regular cases require considerable novel ideas.

Slow mixing times and their dependence on β for low temperatures. Fix β large. In Theorem 1.2, we show that $\hat{t}_{\mathrm{mix}} = \exp(\lambda n(1 + o(1)))$ with λ bounded by a universal constant for the whole regime of temperatures and external fields. In contrast, Theorem 1.4 says that $t_{\mathrm{mix}} \geq \exp(c_1 \beta n)$ whp when $B \leq c_1 \beta$. That means that the annealed dynamics mixes much faster than the quenched dynamics at low temperature (i.e., with large β).

Local spin and graph configurations for wrong magnetizations. Recall the discussion of (1.9) below Theorem 1.2(i). It would be of interest to investigate the local neighborhoods and Ising spin configurations around a uniform vertex when there are around sn plusses, for general s. It can be expected that the local graph configuration remains on being a d-ary regular tree. If the discussion below Theorem 1.2(i) is indeed correct, then the spin configuration either equals the plus or the minus configuration, each with a specific probability.

Organisation of the proof. We start in Section 2 by quantifying various preliminaries on the critical external field for the Ising model on d-regular configuration models, as well as general results on Markov chain mixing times. We continue in Section 3 by discussing the mixing of annealed Ising models and prove Theorem 1.2. This is achieved by comparing the annealed Ising model to a generalized Curie-Weiss model that makes the message that the annealed setting is close to mean-field precise. In Section 4, we identify the quenched fixed-spin partition function in Proposition 1.3. In Section 5, we investigate the mixing of quenched Ising models and prove Theorem 1.4. In Section 6, we prove Proposition 1.5, i.e. the identification of B_c^G. In Section 7, we investigate the annealed cut-off behavior stated formally in Proposition 3.4, and give a sketch of its proof. The full proof of Proposition 3.4 is given in the appendix of the extended version of this paper Can et al. (2019).

2. Preliminaries: Critical external fields and mixing times

In this section, we list some preliminary results that are used later. We start by investigating critical externals fields.

2.1. *Critical external fields.* The following lemma gives a quantitative definition of the critical external field B_c^G. We will omit the proof, which follows by standard calculations as in Martinelli et al. (2004, Proposition 4.5):

Lemma 2.1 (Critical external field for uniqueness of the Gibbs measure). *The critical external field B_c^G for the uniqueness of the Ising Gibbs measure on a d-ary tree satisfies*

$$B_c^G = \inf\{B \geq 0 \colon \text{solution to } \theta = -2B + (d-1)\log L_\beta(\mathrm{e}^\theta) \text{ is unique}\},$$

where

$$L_\beta(x) = \frac{\mathrm{e}^{2\beta}x + 1}{\mathrm{e}^{2\beta} + x}. \tag{2.1}$$

In the next lemma, we give a characterization of the annealed critical external field $\hat{B}_c$.

Lemma 2.2 (Critical external field for fast mixing annealed Ising). *Fix $\beta > \beta_c$, then the following assertions hold:*

(i) *For any $B \geq 0$, the function $t \mapsto \hat{\varphi}''_{\beta,B}(t) = \hat{\varphi}''_{\beta,0}(t)$ has two zeros t_u and $1 - t_u$, with*

$$t_u = \frac{1}{2}\left(1 - \sqrt{1 - \frac{4\mathrm{e}^{-4\beta}(d-1)}{(d-2)^2(1 - \mathrm{e}^{-4\beta})}}\right) \in (0, \tfrac{1}{2}). \tag{2.2}$$

(iia) *The equation (1.8) holds. Furthermore, $\hat{\varphi}'_{\beta,B}(t)$ has three solutions $1 > t_1^B > \frac{1}{2} \geq t_2^B > t_3^B > 0$ for $0 \leq B < \hat{B}_c$. In particular, t_1^B, t_2^B, t_3^B are the global maximizer, the local minimizer, and the local maximizer, respectively, of the function $t \mapsto \hat{\varphi}_{\beta,B}(t)$. Moreover, $t_2^0 = \frac{1}{2}$ for $B = 0$. Further, $t_1^0 = 1 - t_3^0$ and t_3^0 are both global maximizers of $\hat{\varphi}_{\beta,0}(t)$. In addition,*

$$\sup\{\hat{\varphi}_{\beta,B}(t_3^B) - \hat{\varphi}_{\beta,B}(t_2^B) \colon \beta > \beta_c, 0 \leq B < \hat{B}_c\} < \infty.$$

(A) $\beta < \beta_c$, or $\beta > \beta_c$, $B > \hat{B}_c$ (B) $\beta > \beta_c$, $0 < B < \hat{B}_c$
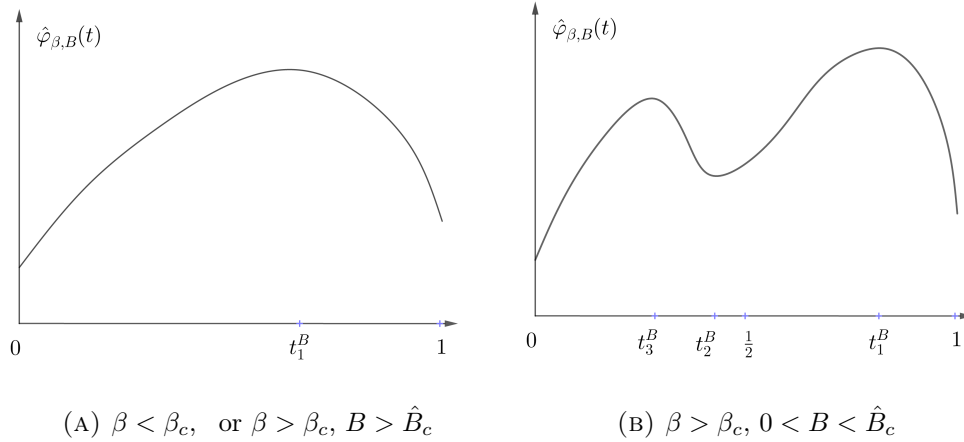
FIGURE 2.1. Graphs of $\hat{\varphi}_{\beta,B}$

(iib) *If $\beta < \beta_c$, or $\beta > \beta_c$ and $B > \hat{B}_c$, then $\hat{\varphi}'_{\beta,B}(t) = 0$ has a unique solution $t_1^B \in (\frac{1}{2}, 1)$, which is the global maximizer of $\hat{\varphi}_{\beta,B}(t)$.*

See Figure 2.1 for what $t \mapsto \hat{\varphi}_{\beta,B}$ looks like. Figure 2.1(B) gives an example of the setting in Lemma 2.2(iia), while Figure 2.1(A) gives an example of the setting in Lemma 2.2(iib).

Proof: It has been shown in the proof of Can (2019, Theorem 1.1 (ii)) that the equation $\hat{\varphi}''_{\beta,B}(t) = \hat{\varphi}''_{\beta,0}(t) = 0$ is equivalent to

$$t^2 - t + \frac{(e^{-4\beta} - 1)(d - 1)}{(d - 2)^2(1 - e^{-4\beta})} = 0.$$

The above equation has two solutions $0 < t_u < \frac{1}{2} < 1 - t_u$ as stated in (i), where t_u is given in (2.2).

We turn to proving (ii). Since t_u is the unique solution of $\hat{\varphi}''_{\beta,0}(t) = 0$ in $[0, \frac{1}{2}]$ and $\hat{\varphi}''_{\beta,0}(0^+) = -\infty$ (since $\hat{\varphi}''_{\beta,0}(t)$ equals $I''(t)$+ bounded terms), the function $-\hat{\varphi}''_{\beta,0}$ changes its sign from plus to minus at t_u and thus t_u is the maximizer of $-\hat{\varphi}'_{\beta,0}$ in $[0, \frac{1}{2}]$. Thus, the equation (1.8) holds. Since $\hat{\varphi}''_{\beta,B}(t) = 0$ has two solutions, $\hat{\varphi}'_{\beta,B}$ has at most three zeros. Moreover, $\hat{\varphi}'_{\beta,B}(0^+) = +\infty$ and $\hat{\varphi}'_{\beta,B}(1^-) = -\infty$ and t_u is the local minimizer and $1 - t_u$ is the local maximizer of $\hat{\varphi}'_{\beta,B}$. Therefore, $\hat{\varphi}'_{\beta,B}(t) = 0$ has a unique solution if $\hat{\varphi}'_{\beta,B}(t_u) > 0$. Moreover, $\hat{\varphi}'_{\beta,B}(t) = \hat{\varphi}'_{\beta,0}(t) + 2B$. Thus $\hat{\varphi}'_{\beta,B}(t_u) > 0$ is equivalent to $B > \hat{B}_c$. That gives (iib).

On the other hand, if $B < \hat{B}_c$ (or equivalently $\hat{\varphi}'_{\beta,B}(t_u) < 0$) then $\hat{\varphi}'_{\beta,B}$ has three solutions $0 < t_3^B < t_u < t_2^B < 1 - t_u < t_1^B < 1$. Note that $f_\beta(t)f_\beta(1 - t) = 1$ for all $t \in (0, 1)$. Thus for all $0 \leq t \leq 1$,

$$\hat{\varphi}'_{\beta,B}(t) = \log\left(\frac{1 - t}{t}\right) + d\log f_\beta(t) + 2B. \tag{2.3}$$

Thus $\hat{\varphi}'_{\beta,B}(\frac{1}{2}) = 2B$, since $f_\beta(\frac{1}{2}) = 1$. Therefore, if $B = 0$, then $t_2^0 = \frac{1}{2}$ and if $B > 0$ then $t_2^B < \frac{1}{2}$.

It remains to bound the gap $\hat{\varphi}_{\beta,B}(t_3^B) - \hat{\varphi}_{\beta,B}(t_2^B)$ uniformly in $\beta > \beta_c$. It has been proved by Can (2019, Lemma 3.1) that

$$\hat{\varphi}_{\beta,B}(t) = \frac{\beta d}{2} - B + I(t) + 2Bt + dt\log f_\beta(t) + J(t), \tag{2.4}$$

where $t \mapsto J(t)$ is a uniformly bounded function on $[0, \frac{1}{2}]$. Since t_3^B and t_2^B are solutions of $\hat{\varphi}'_{\beta,B}(t) = 0$, using (2.3) and (2.4), we get that

$$
\begin{aligned}
\hat{\varphi}_{\beta,B}(t_3^B) &= \frac{\beta d}{2} - B + I(t_3^B) - t_3^B \log\left(\frac{1 - t_3^B}{t_3^B}\right) + J(t_3^B) \\
&= \frac{\beta d}{2} - B - \log(1 - t_3^B) + J(t_3^B),
\end{aligned}
$$

and the same equation holds for t_2^B. Hence,

$$
|\hat{\varphi}_{\beta,B}(t_3^B) - \hat{\varphi}_{\beta,B}(t_2^B)| \leq \log 2 + 2 \max_{0 \leq t \leq \frac{1}{2}} |J(t)|.
$$

This completes the proof of (iia), and thus of the lemma. $\square$

2.2. *Bottleneck and spectral gap bounds for mixing times.* We next recall a result relating the hitting time and spectral gap of a Markov chain with the bottleneck ratio, which has been proved in the book Levin and Peres (2017, Theorem 12.3, 12.4 and 13.14):

Lemma 2.3 (Mixing time bounds). *Let P be the transition matrix of a Markov chain on a finite state space Ω with reversible measure π and let γ be the spectral gap of this chain.*

(i) *In terms of the above notation,*

$$
\left(\gamma^{-1} - 1\right) \log 2 \leq t_{\mathrm{mix}} \leq \log\left(\frac{4}{\min_{x \in \Omega} \pi(x)}\right) \gamma^{-1}.
$$

(ii) *Define the bottleneck ratio as*

$$
\Phi^\star = \min_{S:\, \pi(S) \leq \frac{1}{2}} \Phi(S), \qquad where \qquad \Phi(S) = \frac{Q(S, S^c)}{\pi(S)},
$$

and

$$
Q(x, y) = \pi(x) P(x, y), \qquad Q(K, L) = \sum_{x \in K, y \in L} Q(x, y), \qquad \pi(S) = \sum_{x \in S} \pi(x).
$$

Then

$$
(\Phi^\star)^2 \leq \gamma \leq 2\Phi^\star.
$$

As a consequence,

$$
t_{\mathrm{mix}} \geq \left(\frac{1}{2\Phi^\star} - 1\right) \log 2.
$$

3. Mixing of annealed Ising models: Proof of Theorem 1.2

In this section, we consider a generalized Curie-Weiss model whose Hamiltonian depends only on the number of positive spin. This section is organized as follows. In Section 3.1, we state the model and the main result under some smoothness conditions on the Hamiltonian in Theorem 3.1. The remainder of the section is devoted to the proof of Theorem 3.1, as well as on its application to the proof of Theorem 1.2. In Section 3.2, we prove Theorem 3.1(i), in Section 3.3, we use Theorem 3.1(i) to prove Theorem 1.2(i). A major part of this analysis consists in proving that the Hamiltonian appearing in the annealed Ising model satisfies the requested smoothness conditions. We conclude Section 3.4 with the proof of the cut-off phenomenon for generalized Curie-Weiss models in Theorems 3.1(ii) and 1.2(ii).

3.1. *Mixing of a generalized Curie-Weiss model.* We first recall the transition probabilities of the Glauber dynamics, which are given in equation (1.1). Assume that $\xi_t = \sigma$, let I be a random index in $[n]$ chosen uniformly at random. Then,

$$\xi_{t+1} = \begin{cases} \sigma^{+I} & \text{with probability } \frac{\nu_n(\sigma^{+I})}{\nu_n(\sigma^{+I}) + \nu_n(\sigma^{-I})}, \\ \sigma^{-I} & \text{with probability } \frac{\nu_n(\sigma^{-I})}{\nu_n(\sigma^{+I}) + \nu_n(\sigma^{-I})}. \end{cases}$$

Assume that we are considering Glauber dynamics on $\Omega_n = \{-1, +1\}^n$ with Hamiltonian given by

$$- H_n(\sigma) = n F_n(|\sigma_+|), \tag{3.1}$$

for some function $F_n \colon [0, n] \mapsto \mathbb{R}^+$.

For any $0 \le a < b \le 1$, let us further consider the Glauber dynamics constrained to the subspace

$$\Omega_n^{[a,b]} = \{\sigma \colon [an] \le |\sigma_+| \le [bn]\}.$$

Then, the constrained Gibbs measure is defined as

$$\mu_n^{\beta,B;[a,b]}(\sigma) = \frac{e^{n F_n(|\sigma_+|)}}{\sum_{k=[an]}^{[bn]} \binom{n}{k} e^{n F_n(k)}}, \qquad \text{for all } \sigma \in \Omega_n^{[a,b]}. \tag{3.2}$$

Clearly, $\Omega_n = \Omega_n^{[0,1]}$, and $\mu_n^{\beta,B} = \mu_n^{\beta,B;[0,1]}$.

We first formulate a smoothness condition on F_n that quantifies how close $F_n(k)$ is to $F(k/n)$ for some limiting function $F \colon [0, 1] \mapsto \mathbb{R}^+$. For this, we assume that there is a function $F \in \mathcal{C}^2$ and a constant C such that for all $[an] \le k \le [bn]$,

$$\text{(C1)} \qquad \left| F_n(k) - F(k/n) \right| \le \frac{C}{n} \quad \text{and} \quad \left| F_n(k+1) - F_n(k) - \frac{F'(k/n)}{n} \right| \le \frac{C}{n^2}.$$

Define

$$G(s) = F(s) + I(s) \quad \text{with} \quad I(s) = (s-1)\log(1-s) - s \log s. \tag{3.3}$$

Our aim is to show some sufficient conditions on G under which the Glauber dynamics on the generalized Curie-Weiss model exhibits the metastability or the cut-off phenomenon. For this, we consider the following two further conditions:

(C2) There exist $a < s_3 < s_2 < s_1 < b$ such that G is strictly increasing in the intervals (a, s_3) and (s_2, s_1) and strictly decreasing in the intervals (s_3, s_2) and (s_1, b).

(C3) The function $G'(s) = 0$ has a unique solution $s^\star \in (a, b)$, which is the maximizer of $G(s)$.

We remark that the condition (C3) implies that G is strictly increasing in $(a, s^\star)$ and strictly decreasing in $(s^\star, b)$, $G'(s^\star) = 0$ and $G''(s^\star) < 0$. Condition (C2) implies that the Glauber dynamics mixes slowly, while condition (C3) implies that the Glauber dynamics mixes quickly, as formalized in the following theorem:

Theorem 3.1 (Mixing times of generalized constrained Curie-Weiss models). *For $0 \le a < b \le 1$, consider the Glauber dynamics on the generalized constrained Curie-Weiss model defined by* (3.2).

(i) *Suppose that (C1) and (C2) hold. Then there exists a positive constant C, such that, for all n large enough,*

$$C^{-1} \exp(\lambda n) \le t_{\mathrm{mix}} \le C n^4 \exp(\lambda n),$$

with

$$\lambda = \min\{G(s_3), G(s_1)\} - G(s_2) > 0.$$

(ii) *Suppose that (C1) and (C3) hold. Then the dynamics exhibits the cut-off phenomenon at $c_\star n \log n$ with window of order n, where $c_\star = (2s^\star(1 - s^\star)|G''(s^\star)|)^{-1}$.*

Below we only present the proof for the case $a = 0$ and $b = 1$, since the proof for the general case is exactly the same.

3.2. *Slow mixing for generalized Curie-Weiss models: Proof of Theorem 3.1(i).* Let $(\xi_t)_{t\geq 0}$ be the Glauber dynamics. Define the projection chain $(X_t)_{t\geq 0}$ by

$$X_t = |\{i \colon \xi_t(i) = 1\}|.$$

Then $(X_t)_{t\geq 0}$ is a birth-death process on $\{0, \ldots, n\}$ with probability transitions given by

$$
\begin{aligned}
p_n(k) &= \mathbb{P}(X_{t+1} = k + 1 \mid X_t = k) = \frac{n-k}{n} \times \frac{\mathrm{e}^{nF_n(k+1)}}{\mathrm{e}^{nF_n(k+1)} + \mathrm{e}^{nF_n(k)}}, \\
q_n(k) &= \mathbb{P}(X_{t+1} = k - 1 \mid X_t = k) = \frac{k}{n} \times \frac{\mathrm{e}^{nF_n(k-1)}}{\mathrm{e}^{nF_n(k-1)} + \mathrm{e}^{nF_n(k)}}, \\
r_n(k) &= 1 - p_n(k) - q_n(k).
\end{aligned}
\tag{3.4}
$$

The crucial observation for the proof of Theorem 3.1(i) is that the Glauber dynamics of generalized Curie-Weiss models and their projections have the same spectral gap:

Proposition 3.2 (Spectral gap Curie-Weiss and its projection, Ding et al., 2009b, Proposition 3.9)**.** *The Glauber dynamics $(\xi_t)_{t\geq 0}$ of the generalized Curie-Weiss model and the projection chain $(X_t)_{t\geq 0}$ have the same spectral gap.*

Combining this result with Lemma 2.3(i), we can derive bounds for the mixing time of $(\xi_t)_{t\geq 0}$ from the spectral gap of $(X_t)_{t\geq 0}$. The spectral gap of birth-death processes are well understood, as shown in the following proposition:

Proposition 3.3 (Spectral gaps of birth-death chains, Chen and Saloff-Coste, 2013, Theorem 1.2)**.** *The spectral gap γ of an irreducible birth-death chain on $\{0, \ldots, n\}$ with transition probabilities $(p_n(k), q_n(k), r_n(k))$ and stationary measure ν_n satisfies*

$$\frac{1}{4\ell(i_0)} \leq \gamma \leq \frac{2}{\ell(i_0)}, \tag{3.5}$$

where i_0 is the state such that $\nu_n([0, i_0]) \geq \frac{1}{2}$ and $\nu_n([i_0, n]) \geq \frac{1}{2}$, and

$$\ell(i) = \ell_\nu(i) := \max\left(\max_{j\colon j<i} \sum_{k=j}^{i-1} \frac{\nu_n([0,j])}{\nu_n(k)p_n(k)}, \ \max_{j\colon j>i} \sum_{k=i+1}^{j} \frac{\nu_n([j,n])}{\nu_n(k)q_n(k)} \right), \tag{3.6}$$

where, for $i \leq j$,

$$\nu_n([i,j]) = \sum_{k=i}^{j} \nu_n(k).$$

Now we are ready to give the proof of Theorem 3.1(i). We investigate the birth-death chain $(X_t)_{t\geq 0}$. Under condition (C1),

$$\left| \frac{\mathrm{e}^{nF_n(k+1)}}{\mathrm{e}^{nF_n(k+1)} + \mathrm{e}^{nF_n(k)}} - \frac{\mathrm{e}^{F'(k/n)}}{\mathrm{e}^{F'(k/n)} + 1} \right| + \left| \frac{\mathrm{e}^{nF_n(k-1)}}{\mathrm{e}^{nF_n(k-1)} + \mathrm{e}^{nF_n(k)}} - \frac{1}{\mathrm{e}^{F'(k/n)} + 1} \right| \leq \frac{C}{n}, \tag{3.7}$$

for some constant $C > 0$. Hence,

$$\frac{n-k}{An} \leq p_n(k) \leq \frac{A(n-k)}{n} \quad \text{and} \quad \frac{k}{An} \leq q_n(k) \leq \frac{Ak}{n}, \tag{3.8}$$

for some universal constant $A \geq 1$. The stationary measure of $(X_t)_{t\geq 0}$ is given by

$$\nu_n(k) = \frac{\pi_n(k)}{\pi_n(0) + \cdots + \pi_n(n)}, \tag{3.9}$$

where (see e.g., Levin and Peres, 2017, Section 2.5 or Barrera et al., 2009, (2.3))

$$\pi_n(k) = \prod_{j=0}^{k-1} \frac{p_n(j)}{q_n(j+1)} = \prod_{j=0}^{k-1} \left(\frac{n-j}{j}\right) \left(\frac{e^{nF_n(j+1)}}{e^{nF_n(j)}}\right) = \binom{n}{k} e^{n[F_n(k)-F_n(0)]}. \qquad (3.10)$$

It follows from Stirling's formula that

$$\binom{n}{k} = \exp(n[I(k/n) + o(1)]).$$

Combining the last two estimates with (C1) implies that

$$\begin{aligned}
\pi_n(k) &= \exp\left(n\Big[I(k/n) + F(k/n) - F(0) + o(1)\Big]\right) \\
&= \exp\left(n\Big[G(k/n) - G(0) + o(1)\Big]\right),
\end{aligned} \qquad (3.11)$$

and thus

$$\frac{\pi_n(s)}{\pi_n(k)} = \exp\left(n\Big[G(s/n) - G(k/n) + o(1)\Big]\right). \qquad (3.12)$$

Let us define

$$\begin{aligned}
\ell_1(i,j) &= \sum_{k=j}^{i-1} \frac{\pi_n([0,j])}{\pi_n(k)p_n(k)} = \frac{1}{p_n(k)} \sum_{k=j}^{i-1} \sum_{s=0}^{j} \frac{\pi_n(s)}{\pi_n(k)} \qquad \text{for } i > j, \\
\ell_2(i,j) &= \sum_{k=i+1}^{j} \frac{\pi_n([j,n])}{\pi_n(k)q_n(k)} = \frac{1}{q_n(k)} \sum_{k=i+1}^{j} \sum_{s=j}^{n} \frac{\pi_n(s)}{\pi_n(k)} \qquad \text{for } i < j.
\end{aligned}$$

Then recall (3.6) to see that

$$\ell_\nu(i) = \max\left(\max_{j:\, j<i} \ell_1(i,j),\ \max_{j:\, j>i} \ell_2(i,j)\right). \qquad (3.13)$$

Using (3.8) and (3.12) we obtain that there exists a positive constant C, such that, for all $1 \le j < i \le n-1$,

$$C^{-1} \exp\left(nL_1(i,j)\right) \le \ell_1(i,j) \le Cn^3 \exp\left(nL_1(i,j)\right), \qquad (3.14)$$

where

$$L_1(i,j) = \max_{0 \le s < j \le k < i} \Big[G(s/n) - G(k/n)\Big],$$

and, for all $1 \le i < j \le n-1$,

$$C^{-1} \exp\left(nL_2(i,j)\right) \le \ell_2(i,j) \le Cn^3 \exp\left(nL_2(i,j)\right), \qquad (3.15)$$

where

$$L_2(i,j) = \max_{i < k \le j \le s \le n} \Big[G(s/n) - G(k/n)\Big].$$

By assumption (C2), the function G has two local maximizers s_3 and s_1. We consider the case that $G(s_3) < G(s_1)$, the other case is exactly the same.

Since $G(s_3) < G(s_1)$, s_1 is the global maximizer, and thus there exist $\varepsilon, \delta > 0$, such that $s_2 < s_1 - \delta < s_1 + \delta < 1$ and

$$\varepsilon \le G(s_1) - \max_{|x-s_1| \ge \delta} G(x), \quad \text{and} \quad G(s_1) - G(s_1 \pm \delta) \le \lambda/2, \qquad (3.16)$$

where

$$\lambda = G(s_3) - G(s_2).$$

By (3.12) and (3.16), if $|k - [ns_1]| \ge \delta n$ and n is sufficiently large, then

$$\pi_n(k) \le \pi_n([ns_1]) \exp(-(\varepsilon + o(1))n).$$

Therefore,
$$\nu_n([0, [n(s_1 - \delta)]] \cup [[n(s_1 + \delta)], n]) \le 1/4,$$
so that i_0 satisfies $[n(s_1 - \delta)] \le i_0 \le [n(s_1 + \delta)]$. If $[n(s_1 - \delta)] \le i_0 \le [ns_1]$, then, by (3.16) and the assumption (C2),
$$\max_{j < i_0} L_1(i_0, j) = G(s_3) - G(s_2) + \mathcal{O}(1/n) = \lambda + \mathcal{O}(1/n),$$
$$\max_{j > i_0} L_2(i_0, j) \le G(s_1) - G(s_1 - \delta) + \mathcal{O}(1/n) \le \lambda/2 + \mathcal{O}(1/n).$$
Similarly, if $[ns_1] < i_0 \le [n(s_1 + \delta)]$, then
$$\max_{j < i_0} L_1(i_0, j) = \max\{G(s_3) - G(s_2), G(s_1) - G(s_1 + \delta)\} + \mathcal{O}(1/n) = \lambda + \mathcal{O}(\tfrac{1}{n}),$$
while
$$\max_{j > i_0} L_2(i_0, j) = 0.$$

Combining the above estimates with (3.13), (3.14) and (3.15), we obtain
$$A^{-1} \exp(\lambda n) \le \ell_\nu(i_0) \le An^3 \exp(\lambda n),$$
for some constant A. By (3.5), the same bound (with a slightly larger A) holds for the inverse spectral gap. Hence, using Lemma 2.3, Propositions 3.2 and 3.3, and noting that $-\log(\min_{\sigma \in \Omega_n} \mu_n(\sigma)) \asymp n$, we get that the mixing time of the Glauber dynamics on the generalized Curie-Weiss model satisfies
$$C^{-1} \exp(\lambda n) \le t_{\mathrm{mix}} \le Cn^4 \exp(\lambda n),$$
for some constant C independent of n. $\qquad\square$

3.3. *Verifying conditions generalized Curie-Weiss model: Proof Theorem 1.2(i).* By Theorem 3.1(i), we only need to verify the conditions (C1) and (C2) for the annealed Ising model on random regular graphs.

It is known (see for instance Can, 2019, Lemma 2.1(i) and (3.2)) that if $|\sigma_+| = j$ then
$$\hat{\mu}_n(\sigma) = \frac{\exp(nF_n(j))}{\sum_{k=0}^{n} \binom{n}{k} \exp(nF_n(k))},$$
with
$$F_n(k) = \frac{\beta d}{2} + \frac{1}{n} \log g(dk, dn) + B\left(\frac{2k}{n} - 1\right),$$
where $g(k, m)$ satisfies that, for all $0 \le k \le \ell \le m$,
$$\left| \frac{1}{m} \log g(k, m) - \frac{1}{m} \log g(\ell, m) - \int_{k/m \wedge (1-k/m)}^{\ell/m \wedge (1-\ell/m)} \log f_\beta(s)ds \right| \le C \frac{|k - \ell|}{m^2}, \tag{3.17}$$
with C being a universal constant independent of k, ℓ, m. By (3.17),
$$\left| F_n(k) - F(k/n) \right| \le \frac{C}{n^2},$$
and
$$\left| [F_n(k+1) - F_n(k)] - \frac{F'(k/n)}{n} \right| \le \frac{C}{n^2}, \tag{3.18}$$
where
$$F(t) = \frac{\beta d}{2} + B(2t - 1) + d \int_0^{t \wedge 1 - t} \log f_\beta(s)ds, \tag{3.19}$$
which implies that (C1) holds. The function $G(t) = I(t) + F(t)$ is indeed the function $\hat{\varphi}^{\beta, B}(t)$. Hence, Lemma 2.2 implies that (C2) holds.

Finally, by Lemma 2.2(ii) and Theorem 3.1(i),

$$\sup_{\substack{\beta > \beta_c \\ 0 \le B < \hat{B}_c}} \lambda = \sup_{\substack{\beta > \beta_c \\ 0 \le B < \hat{B}_c}} G(t_3^B) - G(t_2^B) < \infty, \tag{3.20}$$

which completes the proof of Theorem 1.2(i).   □

3.4. *Cut-off generalized Curie-Weiss models: Proof Theorems 3.1(ii) and 1.2(ii).* For $\gamma > 0$, define

$$T_n^+(\gamma) = c_\star n \log n + \gamma n, \qquad T_n^-(\gamma) = c_\star n \log n - \gamma n,$$

where

$$c_\star = -[s^\star(1 - s^\star)G''(s^\star)]^{-1} > 0,$$

since $G''(s^\star) < 0$. Theorem 3.1(ii) follows from the following proposition:

Proposition 3.4 (Annealed cut-off behavior). *Suppose that Conditions (C1) and (C3) hold. Then,*

$$\lim_{\gamma \to \infty} \liminf_{n \to \infty} d_n(T_n^-(\gamma)) = 1, \qquad \lim_{\gamma \to \infty} \limsup_{n \to \infty} d_n(T_n^+(\gamma)) = 0,$$

where we recall that, for $T \ge 1$,

$$d_n(T) = \sup_{A \subset \Omega_n} \sup_{\sigma \in \Omega_n} |\mu_n(A) - \mathbb{P}_\sigma(\xi_T \in A)|.$$

For the proof, we will closely follow the strategy of proving cut-off phenomena for censored Curie-Weiss model in Ding et al. (2009a). Since the proof is long and technical, we defer it to Section 7.

Proof of Theorem 1.2(ii). Thanks to Theorem 3.1, we only need to show that Conditions (C1) and (C3) hold for the annealed Ising model. The condition (C1) is already proved in Section 3.3. Condition (C3) follows from Lemma 2.2 (iib), since $G(t) = F(t) + I(t) = \hat{\varphi}_{\beta,B}(t)$.   □

4. Quenched fixed-spin partition function: Proof of Proposition 1.3

In this section, we prove Proposition 1.3. We start in Section 4.1 by relating the fixed-spin pressures for different values of the total spin. We continue in Section 4.2 to prove concentration properties of the fixed-graph pressure, and prove Proposition 1.3(i). In Section 4.3, we show that the mean pressure converges along a subsequence and use this to prove Proposition 1.3(ii). We conclude in Section 4.4 by showing where the quenched and annealed pressures agree using large deviation ideas.

4.1. *Relating pressures with different total spins.* Let $G = (V, E)$ be a graph with degrees bounded by d and consider the Ising model on G with Hamiltonian given by

$$H_G^{\beta,0}(\sigma) = -\beta \sum_{(i,j) \in E} \sigma_i \sigma_j = -(\beta|E| - 2\beta e(\sigma_+, \sigma_-)),$$

where $e(A, B)$ is the number of edges between A and B in G. For any $0 \le k \le |V|$, define

$$Z_{G,k}^{\beta,0} = \sum_{\sigma \,:\, |\sigma_+| = k} \exp(-H_G^{\beta,0}(\sigma)).$$

We have

$$
\begin{aligned}
Z_{G,k+\ell}^{\beta,0} &= \mathrm{e}^{\beta|E|} \sum_{\substack{U \subset V \\ |U|=k+\ell}} \mathrm{e}^{-2\beta e(U,U^c)} \\
&= \mathrm{e}^{\beta|E|} \frac{1}{\binom{k+\ell}{\ell}} \sum_{\substack{A \subset V \\ |A|=k}} \sum_{\substack{B \subset V \setminus A \\ |B|=\ell}} \mathrm{e}^{-2\beta e(A \cup B,(A \cup B)^c)} \\
&= \mathrm{e}^{\beta|E|} \sum_{\substack{A \subset V \\ |A|=k}} \mathrm{e}^{-2\beta e(A,A^c)} \frac{1}{\binom{k+\ell}{\ell}} \sum_{\substack{B \subset V \setminus A \\ |B|=\ell}} \mathrm{e}^{2\beta[e(A,B)-e(B,(A \cup B)^c)]}.
\end{aligned}
$$

Using the naive bound

$$
-d\ell \le -d|B| \le e(A,B) - e(B,(A \cup B)^c) \le d|B| \le d\ell,
$$

we get

$$
\binom{|V|-k}{\ell} \exp(-2\beta d\ell) \le \sum_{\substack{B \subset V \setminus A \\ |B|=\ell}} \mathrm{e}^{2\beta[e(A,B)-e(B,(A \cup B)^c)]} \le \binom{|V|-k}{\ell} \exp(2\beta d\ell).
$$

Therefore,

$$
\mathrm{e}^{-2\beta d\ell} \binom{|V|-k}{\ell} \bigg/ \binom{k+\ell}{\ell} \le \frac{Z_{G,k+\ell}^{\beta,0}}{Z_{G,k}^{\beta,0}} \le \mathrm{e}^{2\beta d\ell} \binom{|V|-k}{\ell} \bigg/ \binom{k+\ell}{\ell}. \tag{4.1}
$$

This shows that the fixed-spin pressures cannot change too much when changing the value of the total spin, a fact that will prove to be useful when deriving properties of the limiting fixed-spin pressure.

4.2. *Concentration of finite-graph pressure: Proof of Proposition 1.3(i).* In this section, we will use a vertex-revealing process, combined with the Azuma-Hoeffding inequality, to show that the quenched finite graph fixed-spin pressure is whp close to its mean. We start by setting up the necessary notation. Let $G' = (V', E')$ be a graph obtained from $G = (V, E)$ by adding one vertex v with at most d edges between v and V. Then,

$$
\begin{aligned}
Z_{G',k}^{\beta,0} &= \mathrm{e}^{\beta|E'|} \sum_{\substack{U' \subset V' \\ |U'|=k}} \mathrm{e}^{-2\beta e(U',V' \setminus U')} \\
&= \mathrm{e}^{\beta|E'|} \bigg[\sum_{\substack{U \subset V \\ |U|=k}} \mathrm{e}^{-2\beta e(U,(V \setminus U) \cup \{v\})} + \sum_{\substack{U \subset V \\ |U|=k-1}} \mathrm{e}^{-2\beta e(U \cup \{v\},V \setminus U)} \bigg].
\end{aligned}
$$

We observe that $0 \le |E'| - |E| \le d$, and

$$
\begin{aligned}
0 &\le e(U,(V \setminus U) \cup \{v\}) - e(U,V \setminus U) \le d, \\
0 &\le e(U \cup \{v\}, V \setminus U) - e(U, V \setminus U) \le d.
\end{aligned}
$$

Therefore,

$$
\mathrm{e}^{-2\beta d} Z_{G,k}^{\beta,0} \le Z_{G',k}^{\beta,0} \le \mathrm{e}^{\beta d}(Z_{G,k}^{\beta,0} + Z_{G,k-1}^{\beta,0}). \tag{4.2}
$$

It follows from (4.1) that

$$
\frac{Z_{G,k}^{\beta,0}}{Z_{G,k-1}^{\beta,0}}, \ \frac{Z_{G',k+1}^{\beta,0}}{Z_{G',k}^{\beta,0}} \in \left[\mathrm{e}^{-2\beta d} \frac{|V|-k+1}{k+1}, \mathrm{e}^{2\beta d} \frac{|V|-k+1}{k} \right]. \tag{4.3}
$$

Combining (4.2) and (4.3) we obtain

$$
\begin{aligned}
e^{-2\beta d}\Big(1 \wedge e^{-2\beta d}\frac{|V|-k+1}{k+1}\Big) \;&\leq\; \frac{Z^{\beta,0}_{G',k}}{Z^{\beta,0}_{G,k}},\; \frac{Z^{\beta,0}_{G',k+1}}{Z^{\beta,0}_{G,k}} \\
&\leq\; e^{\beta d}\Big(1 + e^{2\beta d}\frac{k+1}{|V|-k+1}\Big)\Big(1 \vee e^{2\beta d}\frac{|V|-k+1}{k}\Big).
\end{aligned}
\tag{4.4}
$$

For $t \in (0,1)$, we define

$$
Z^{\beta,0}_G(t) = \sum_{\sigma:\, |\sigma_+|=\lceil t|V|\rceil} \exp(-H^{\beta,0}_G(\sigma)).
$$

Since $0 \leq \lceil t|V'|\rceil - \lceil t|V|\rceil \leq 1$, the inequality (4.4) implies that

$$
e^{-2\beta d}\Big(1 \wedge e^{-2\beta d}\frac{1-t}{t+1}\Big) \leq \frac{Z^{\beta,0}_{G'}(t)}{Z^{\beta,0}_G(t)} \leq e^{\beta d}\Big(1 + e^{2\beta d}\frac{t}{1-t}\Big)\Big(1 \vee e^{2\beta d}\frac{2-t}{t}\Big).
\tag{4.5}
$$

For any random graph G_n, we may consider $\log Z^{\beta,0}_n(t)$ as a martingale with respect to the vertex-revealing filtration (see e.g. Bollobás, 2001, Chapter 11.4 or Dembo et al., 2014 for the construction of the filtration). Indeed, for $k \in [n]$, add vertex $k \in [n]$ along with its edges to the graph, and let $G_{k,n}$ denote the graph with vertex set $[k]$ and edge set the restriction of G_n to $[k]$. Let $\mathscr{F}_k$ denote the sigma-algebra generated by $(G_{l,n})_{l \leq k}$, and define

$$
M_k = \mathbb{E}[\log Z^{\beta,0}_n(t) \mid \mathscr{F}_{k-1}],
\tag{4.6}
$$

Then, $M_n = \log Z^{\beta,0}_n(t)$ and $M_0 = \mathbb{E}[\log Z^{\beta,0}_n(t)]$. Due to (4.5), the increments of the martingale are uniformly bounded by

$$
|M_{k+1} - M_k| \leq \triangle = \beta d + \log(1 + e^{2\beta}\frac{t}{1-t}) + \log(1 \vee e^{2\beta d}\frac{2-t}{t}).
$$

Hence, it follows from the Azuma-Hoeffding inequality that

$$
\mathbb{P}\Big(|n^{-1}\log Z^{\beta,0}_n(t) - n^{-1}\mathbb{E}(\log Z^{\beta,0}_n(t))| \geq \varepsilon\Big) \leq 2\exp(-\varepsilon^2 n/2\,\triangle^2).
$$

In other words,

$$
\mathbb{P}\Big(|\varphi^{\beta,0}_n(t) - \tilde{\varphi}^{\beta,0}_n(t)| \geq \varepsilon\Big) \leq 2\exp(-\varepsilon^2 n/2\,\triangle^2),
\tag{4.7}
$$

which proves that

$$
\tilde{\varphi}^{\beta,0}_n(t) - \varphi^{\beta,0}_n(t) \xrightarrow{a.s.} 0.
$$

The same statement holds for the sequence $\varphi^{\beta,B}_n(t)$ since

$$
\varphi^{\beta,B}_n(t) = \varphi^{\beta,0}_n(t) + B\Big(2\frac{\lceil nt\rceil}{n} - 1\Big).
\tag{4.8}
$$

This completes the proof of Proposition 1.3(i). $\square$

4.3. *Convergence of the mean pressure: Proof of Proposition 1.3(ii).* In this section, we prove that there exists a subsequence $\{n_k\}_{k \geq 0}$ along which $t \mapsto \tilde{\varphi}^{\beta,0}_n(t)$ converges uniformly to a continuous function $t \mapsto \varphi_{\beta,0}(t)$. By (4.1), for $0 \leq t < t+s \leq \frac{1}{2}$,

$$
e^{-2\beta d\ell}\binom{n-k}{\ell}\Big/\binom{k+\ell}{\ell} \leq \frac{Z^{\beta,0}_n(t+s)}{Z^{\beta,0}_n(t)} \leq e^{2\beta d\ell}\binom{n-k}{\ell}\Big/\binom{k+\ell}{\ell},
\tag{4.9}
$$

with $k = \lceil nt\rceil$ and $\ell = \lceil ns\rceil$. By Stirling's formula,

$$
\binom{n-k}{\ell}\Big/\binom{k+\ell}{\ell} = \exp\Big[(1-t)nI\Big(\frac{s}{1-t}\Big) - (t+s)nI\Big(\frac{s}{t+s}\Big) + \mathcal{O}(\log n)\Big].
$$

Therefore, with $\varphi_n^{\beta,0}(t) = \frac{1}{n}\log Z_n^{\beta,0}(t)$, there exist non-random constants $c_1, c_2 > 0$ such that

$$-2\beta ds + J(t,s) - \frac{c_1(\log n)}{n} \leq \varphi_n^{\beta,0}(t+s) - \varphi_n^{\beta,0}(t) \leq 2\beta ds + J(t,s) + \frac{c_2(\log n)}{n}, \qquad (4.10)$$

where

$$
\begin{aligned}
J(t,s) &= (1-t)I\left(\tfrac{s}{1-t}\right) - (t+s)I\left(\tfrac{s}{t+s}\right) \\
&= s\log(\tfrac{1}{t+s}-1) - (1-t)\log(1-\tfrac{s}{1-t}) + t\log(1-\tfrac{s}{t+s}).
\end{aligned}
$$

This inequality holds with probability 1, so the same inequality also holds for $\tilde\varphi_n^{\beta,0}$, i.e.,

$$-2\beta ds + J(t,s) - \frac{c_1(\log n)}{n} \leq \tilde\varphi_n^{\beta,0}(t+s) - \tilde\varphi_n^{\beta,0}(t) \leq 2\beta ds + J(t,s) + \frac{c_2(\log n)}{n}, \qquad (4.11)$$

We observe the following:

(O1) $J(t,s) \geq 0$ for all $0 < t < t+s \leq \frac{1}{2}$, since the function $xI(1/x)$ is increasing in $x > 1$ and $1 - t \geq \frac{1}{2} \geq t + s$,

(O2) $(t,s) \mapsto J(t,s)$ is continuous on $\{(t,s)\colon t,s \in [0,\frac{1}{2}], s+t \leq \frac{1}{2}\}$ (and hence uniformly continuous on this compact set).

(O3) For any $t \in (0,1)$ and for s small, $J(t,s) = s\log\left(\frac{1-t}{t}\right) + \mathcal{O}(s^2)$.

By (O2), we conclude that the collection of functions $\{\tilde\varphi_n^{\beta,0}(t), 0 \leq t \leq \frac{1}{2}\}$ is equicontinuous. Besides, $\tilde\varphi_n^{\beta,0}(0) = \beta dn/(2n) = d\beta/2$ for all $n \in \mathbb{N}$. Hence, using Ascoli-Arzelà's theorem, we can extract a subsequence $(n_k)_{k\geq 1}$ along which $\{\tilde\varphi_n^{\beta,0}, n \geq 1\}$ converge uniformly in $[0,\frac{1}{2}]$ to a continuous function $\varphi_{\beta,0}$. By symmetry, $\tilde\varphi_n^{\beta,0}(t) = \tilde\varphi_n^{\beta,0}(1-t) + \mathcal{O}(1/n)$. Moreover, by (4.8), $\tilde\varphi_n^{\beta,B}(t) = \tilde\varphi_n^{\beta,0}(t) + B(2t-1) + \mathcal{O}(1/n)$. Therefore, $\{\tilde\varphi_{n_k}^{\beta,B}, k \geq 1\}$ converges uniformly in $[0,1]$ to the continuous function $\varphi_{\beta,B}$ defined by

$$\varphi_{\beta,B}(t) = \varphi_{\beta,0}(t) + B(2t-1),$$

where $\varphi_{\beta,0}$ is extended to all the interval $[0,1]$ by $\varphi_{\beta,0}(t) = \varphi_n^{\beta,0}(1-t)$ for $t \in [\frac{1}{2},1]$.

Recall that

$$B_c := \sup_{0<t<u\leq\frac{1}{2}} \frac{\varphi_{\beta,0}(t) - \varphi_{\beta,0}(u)}{2(u-t)}.$$

By (O1), we have $\tilde\varphi_n^{\beta,0}(t) - \tilde\varphi_n^{\beta,0}(t+s) \leq 2\beta ds + c_2(\log n/n)$ on $0 < t < t+s \leq \frac{1}{2}$, and thus

$$B_c \leq \beta d. \qquad (4.12)$$

By Jensen's inequality, $\mathbb{E}[\log Z_n^{\beta,0}(t)] \leq \log \mathbb{E}[Z_n^{\beta,0}(t)]$ for all $t \in (0,1)$. Therefore,

$$\varphi_{\beta,0}(t) \leq \hat\varphi_{\beta,0}(t). \qquad (4.13)$$

We claim that

$$\sup_{0\leq t\leq\frac{1}{2}} \varphi_{\beta,0}(t) = \sup_{0\leq t\leq\frac{1}{2}} \hat\varphi_{\beta,0}(t) = \hat\varphi_{\beta,0}(t_3^0), \qquad (4.14)$$

where $t_3^0 \in (0,\frac{1}{2})$ is one of two global maximizers of $\hat\varphi_{\beta,0}$ as stated in Lemma 2.3.

Suppose for the moment (4.14) holds, then, for all $\beta > \beta_c$,

$$B_c \geq \sup_{t_3^0<u\leq\frac{1}{2}} \frac{\hat\varphi_{\beta,0}(t_3^0) - \hat\varphi_{\beta,0}(u)}{2(u-t_3^0)} > 0. \qquad (4.15)$$

We now show that B_c grows linearly in β when β tends to infinity. We first show that there exists a positive constant $c > 0$, such that, for $\beta \geq 12$,

$$\lim_{n\to\infty} \mathbb{P}\left(Z_n^{\beta,0}(0) \geq \mathrm{e}^{c\beta n} Z_n^{\beta,0}(\tfrac{1}{2})\right) = 1. \qquad (4.16)$$

Indeed,

$$
\begin{aligned}
\frac{Z_n^{\beta,0}(0)}{Z_n^{\beta,0}(\frac{1}{2})} &= \left(\sum_{\sigma \,:\, |\sigma_+| = \lceil n/2 \rceil} \exp(-2\beta e(\sigma_+, \sigma_-)) \right)^{-1} \\
&\geq \left(\binom{n}{\lceil n/2 \rceil} \exp\left(-2\beta i_{n,d} \lceil n/2 \rceil \right) \right)^{-1},
\end{aligned}
$$

where $i_{n,d}$ is the isoperimetric number of the random regular graph defined as

$$
i_{n,d} := \inf \left\{ \frac{e(A, A^c)}{|A|} : A \subset [n], \, |A| \leq n/2 \right\},
$$

where $[n]$ is the set of vertices in G_n. Bollobás (1988) showed that

$$
\lim_{n \to \infty} \mathbb{P}\left(i_{n,d} \geq \frac{d}{2} - \sqrt{d \log 2} \right) = 1. \tag{4.17}
$$

Therefore, whp,

$$
\frac{Z_n^{\beta,0}(0)}{Z_n^{\beta,0}(\frac{1}{2})} \geq \exp\left(\left(\beta(\tfrac{d}{2} - \sqrt{d \log 2}) - \log 2 \right) n \right) \geq \exp(c\beta n), \tag{4.18}
$$

for some $c > 0$. Notice that for the second inequality we have used that $\beta \geq 12$ and $d \geq 3$.

Using (4.16), we thus conclude that

$$
B_c \geq \liminf_{n \to \infty} \frac{\tilde{\varphi}_n^{\beta,0}(0) - \tilde{\varphi}_n^{\beta,0}(\frac{1}{2})}{1} = \liminf_{n \to \infty} \frac{\varphi_n^{\beta,0}(0) - \varphi_n^{\beta,0}(\frac{1}{2})}{1} \geq c\beta. \tag{4.19}
$$

Combining this with (4.12) and (4.15), we obtain the desired result in Proposition 1.3(ii). We are thus left to prove (4.14).

Now we prove (4.14), which is a direct consequence of Proposition 1.3(i) and the fact that

$$
\varphi_n^{\beta,0}(t_3^0) \xrightarrow{\text{a.s.}} \hat{\varphi}_{\beta,0}(t_3^0) = \sup_{t \in [0, \frac{1}{2}]} \hat{\varphi}_{\beta,0}(t). \tag{4.20}
$$

Furthermore, (4.20) follows from the following two claims

$$
\mathbb{P}\left(\lim_{n \to \infty} \frac{\log Z_n^{\beta,0}}{n} = \hat{\varphi}_{\beta,0}(t_3^0) = \sup_{t \in [0, \frac{1}{2}]} \hat{\varphi}_{\beta,0}(t) \right) = 1, \tag{4.21}
$$

and, for any $\delta > 0$ and for all n large enough,

$$
\mathbb{P}\left(\left| \frac{\log Z_n^{\beta,0}(t_3^0)}{n} - \frac{\log Z_n^{\beta,0}}{n} \right| \geq \delta \right) \leq n^{-2}. \tag{4.22}
$$

The first claim (4.21) follows from (1.7), so it remains to prove (4.22). By (1.3),

$$
\mathbb{E}(Z_n^{\beta,0}(t)) \leq \exp(n\hat{\varphi}_{\beta,0}(t) + \mathcal{O}(1)).
$$

Hence, Markov's inequality implies that for any $\varepsilon > 0$

$$
\mathbb{P}\left(Z_n^{\beta,0}(t) \geq \exp(n(\hat{\varphi}_{\beta,0}(t) + \varepsilon)) \right) \leq \exp(-\varepsilon n/2). \tag{4.23}
$$

For $k \in [n]$, define $A_k = \{\sigma : |\sigma_+| = k\}$ and write

$$
Z_{n,k}^{\beta,0} = \sum_{\sigma \in A_k} \exp(-H_n^{\beta,0}(\sigma)).
$$

Then, by symmetry,

$$\sum_{k \leq \lceil n/2 \rceil} Z_{n,k}^{\beta,0} \geq \frac{Z_n^{\beta,0}}{2}. \tag{4.24}$$

For any $\delta > 0$, we define

$$Z_n^{\beta,0}(t_3^0, \delta^+) = \sum_{k=0}^{\lceil (t_3^0 - \delta)n \rceil} Z_{n,k}^{\beta,0} + \sum_{k=\lfloor (t_3^0 + \delta)n \rfloor}^{\lceil n/2 \rceil} Z_{n,k}^{\beta,0},$$

and

$$Z_n^{\beta,0}(t_3^0, \delta^-) = \sum_{k=\lfloor (t_3^0 - \delta)n \rfloor}^{\lceil (t_3^0 + \delta)n \rceil} Z_{n,k}^{\beta,0}.$$

As for (4.23), using Markov's inequality, we see that for any $\varepsilon > 0$

$$\mathbb{P}\left(Z_n^{\beta,0}(t_3^0, \delta^+) \geq \exp\left(n \sup\{\hat\varphi_{\beta,0}(t) \colon 0 \leq t_3^0 \leq \tfrac{1}{2}, |t - t_3^0| \geq \delta\} + n\varepsilon \right) \right) \leq \exp(-\varepsilon n/3).$$

Combining this with (4.21), for some $\varepsilon = \varepsilon(\delta) > 0$,

$$\mathbb{P}(Z_n^{\beta,0}(t_3^0, \delta^+) \geq Z_n^{\beta,0}/4) \leq \exp(-\varepsilon n/4).$$

Thus, by (4.24),

$$\mathbb{P}(Z_n^{\beta,0}(t_3^0, \delta^-) \geq Z_n^{\beta,0}/4) \geq 1 - \mathbb{P}(Z_n^{\beta,0}(t_3^0, \delta^+) \geq Z_n^{\beta,0}/4) \geq 1 - \exp(-\varepsilon n/4). \tag{4.25}$$

On the other hand, using (4.1), (4.10) and (O3), it follows that for given small $\delta > 0$ for all n large enough and for k such that $|k - [t_3^0 n]| \leq \delta n$ and $(\log n)/n \leq \delta$, one has, for some $\tilde c > 0$,

$$\left| \frac{\log Z_n^{\beta,0}(t_3^0)}{n} - \frac{\log Z_{n,k}^{\beta,0}}{n} \right| \leq 2\beta\delta d + J(t_3^0, \delta) + \tilde c\left(\frac{\log n}{n} \right) \leq 2C\delta,$$

where $C = \beta d + \log\left(\frac{1-t_3^0}{t_3^0} \right) + \tilde c$. Therefore,

$$\left| \frac{\log Z_n^{\beta,0}(t_3^0)}{n} - \frac{\log Z_n^{\beta,0}(t_3^0, \delta^-)}{n} \right| \leq 3C\delta.$$

Combining this with (4.25) and the fact that $Z_n^{\beta,0}(t_3^0, \delta^-) \leq Z_n^{\beta,0}$, we obtain (4.22). This completes the proof of (4.14), and thus of Proposition 1.3(ii). $\quad\square$

4.4. *Relating the quenched and annealed pressures.* In this section, we relate the pressure in the quenched and annealed settings in more detail. Using Dembo et al. (2014, Proposition 1.2), we have $\frac{1}{n} \log Z_n^{\beta,B} - \frac{1}{n}\mathbb{E}[\log Z_n^{\beta,B}] \to 0$ a.s., and thus

$$\varphi(\beta, B) = \lim_{n \to \infty} \frac{1}{n}\mathbb{E}[\log Z_n^{\beta,B}] \text{ a.s.} \tag{4.26}$$

On the other hand, since $Z_n^{\beta,B} = \sum_{i=0}^n Z_n^{\beta,B}(\frac{i}{n})$,

$$\frac{1}{n} \max_{0 \leq i \leq n} \log Z_n^{\beta,B}\left(\frac{i}{n} \right) \leq \frac{1}{n} \log Z_n^{\beta,B} \leq \frac{1}{n}\left[\log(n+1) + \max_{0 \leq i \leq n} \log Z_n^{\beta,B}\left(\frac{i}{n} \right) \right],$$

and hence

$$\frac{1}{n}\mathbb{E}[\log Z_n^{\beta,B}] = \max_{0 \leq i \leq n} \tilde\varphi_n^{\beta,B}\left(\frac{i}{n} \right) + o(1), \tag{4.27}$$
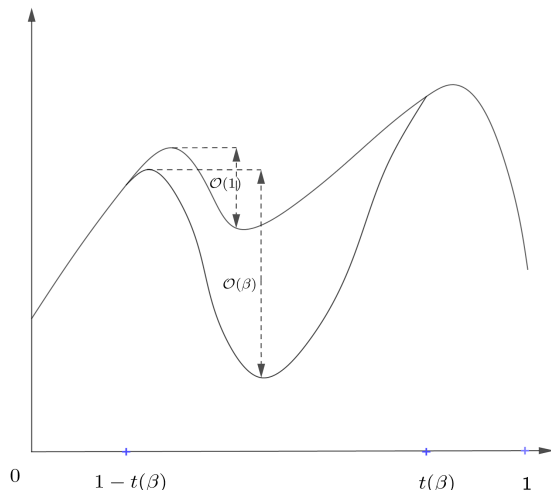
FIGURE 4.2. Graphs of $\hat{\varphi}_{\beta,B}$ and $\varphi_{\beta,B}$. The two functions are equal in $[0, 1-t(\beta)] \cup [t(\beta), 1]$, and in the remaining interval, $\hat{\varphi}_{\beta,B}$ is above $\varphi_{\beta,B}$.

where we recall that $\tilde{\varphi}_n^{\beta,B}(t) = \frac{1}{n}\mathbb{E}[\log Z_n^{\beta,B}(t)]$. By Proposition 1.3 (ii), there is a subsequence $(n_k)_{k\geq 1}$, such that $(\tilde{\varphi}_{n_k}^{\beta,B})_{k\geq 1}$ converges uniformly to a continuous function $\varphi_{\beta,B}$. Hence

$$\lim_{k\to\infty} \max_{0\leq i\leq n_k} \tilde{\varphi}_{n_k}^{\beta,B}\left(\frac{i}{n_k}\right) = \max_{0\leq t\leq 1} \varphi_{\beta,B}(t).$$

Combining this with (4.26) and (4.27), we obtain that

$$\varphi(\beta, B) = \max_{0\leq t\leq 1} \varphi_{\beta,B}(t) \text{ a.s.}$$

In addition, recall from (1.7) that $\hat{\varphi}(\beta, B) = \lim_{n\to\infty} \frac{1}{n}\log \mathbb{E}[Z_n^{\beta,B}]$ exists, and satisfies

$$\varphi(\beta, B) \stackrel{\text{a.s.}}{=} \hat{\varphi}(\beta, B) = \max_{0\leq t\leq 1} \hat{\varphi}_{\beta,B}(t).$$

Then, $\varphi_{\beta,B}(t) \leq \hat{\varphi}_{\beta,B}(t)$ for all t by (4.13), but $Z_n^{\beta,B}(t) = e^{n\varphi_{\beta,B}(t)(1+o_\mathbb{P}(1))}$ has expected value $\mathbb{E}[Z_n^{\beta,B}(t)] = e^{n\hat{\varphi}_{\beta,B}(t)(1+o(1))}$. Thus, it is impossible that $\varphi_{\beta,B}(t) < \hat{\varphi}_{\beta,B}(t)$ for all $t \in [0,1]$, while at the same time and instead, these two functions must have the same optimizer and value in their optimum. Since the optimizer equals t_1^B, we thus arrive at the fact that $\varphi_{\beta,B}(t_1^B) = \hat{\varphi}_{\beta,B}(t_1^B)$ holds for all β, B.

However, we also know that, for every b,

$$\varphi_{\beta,b}(t_1^b) = \varphi_{\beta,B}(t_1^b) + 2(b-B)t_1^b, \qquad \hat{\varphi}_{\beta,b}(t_1^b) = \hat{\varphi}_{\beta,B}(t_1^b) + 2(b-B)t_1^b.$$

As a result, for all b, $\varphi_{\beta,B}(t_1^b) = \hat{\varphi}_{\beta,B}(t_1^b)$ for every $b \in \mathbb{R}$. Now, by varying b, we vary t_1^b in a neighborhood of t_1^B, which shows that $\varphi_n^{\beta,B}(t) \xrightarrow{\text{a.s.}} \hat{\varphi}_{\beta,B}(t)$ for all t in a neighborhood of t_1^B. Further, since $t_1^b \to 1$ as $b \to \infty$, it follows that $\varphi_{\beta,B}(t) = \hat{\varphi}_{\beta,B}(t)$ for all $t \in [t_1^B - \delta, 1]$ for some $\delta > 0$.

Next take $\beta > \beta_c$. Then, since $t_1^b \to t(\beta) \equiv \frac{1}{2}(m(\beta) + 1)$ as $b \to 0$, where $m(\beta)$ is the spontaneous magnetization, we conclude that $\varphi_{\beta,B}(t) = \hat{\varphi}_{\beta,B}(t)$ for all $t \in [t(\beta), 1]$. Next, we start from $\varphi_{\beta,B}(t_1^b) = \hat{\varphi}_{\beta,B}(t_1^b)$ for all b, but now take $b < -B$, so that $t_1^b < \frac{1}{2}$. Again, we can take any $b < -B$ in this inequality, and letting $b \searrow -B$ gives $t_1^b \to 1 - t(\beta)$, while for $b \to -\infty$, $t_1^b \to 0$. Thus, we

conclude also that $\varphi_{\beta,B}(t) = \hat{\varphi}_{\beta,B}(t)$ for all $t \in [0, 1 - t(\beta)]$. Unfortunately, however, this does not imply that $\varphi_{\beta,B}(t_3^B) = \hat{\varphi}_{\beta,B}(t_3^B)$, since $t_3^B \in (1 - t(\beta), \frac{1}{2})$.

Since $t(\beta) > \frac{1}{2}$ for $\beta > \beta_c$, we cannot conclude that $\varphi_{\beta,B}(t) = \hat{\varphi}_{\beta,B}(t)$ for all $t \in (\frac{1}{2} - t(\beta), t(\beta))$. In fact, this is certainly not true, as we next show.

Indeed, we claim that $\varphi^{\beta,B}(\frac{1}{2}) \neq \hat{\varphi}^{\beta,B}(\frac{1}{2})$ for β very large. Indeed, by Lemma 2.2(iia),

$$\hat{\varphi}_{\beta,B}(0) - \hat{\varphi}_{\beta,B}(\tfrac{1}{2}) \leq \hat{\varphi}_{\beta,B}(t_3^B) - \hat{\varphi}_{\beta,B}(t_2^B) \leq \lambda,$$

which remains finite for large β. Instead, (4.18) shows that

$$\varphi_{\beta,B}(0) - \varphi_{\beta,B}(\tfrac{1}{2}) \geq c\beta.$$

Together with $\varphi_{\beta,B}(0) = \hat{\varphi}_{\beta,B}(0)$ this completes the proof of the claim. $\qquad\square$

5. Mixing of quenched Ising models: Proof of Theorem 1.4

In this section, we investigate the mixing of the quenched Ising model on the d-regular random graph. We start in Section 5.1 to investigate the slow mixing for $\beta > \beta_c$, $B < B_c(\beta)$. We continue in Section 5.2 to prove the rapid mixing for $\beta > \beta_c$ and $B > B_c^G$.

5.1. *Slow mixing in the quenched setting: Proof of Theorem 1.4(i).* First, we consider the case $\beta \geq 12$. As we have shown in (4.16), whp $Z_n^{\beta,0}(0) \geq \exp(c\beta n)Z_n^{\beta,0}(\frac{1}{2})$ for some positive constant $c > 0$. Therefore, whp $Z_n^{\beta,B}(0) \geq \exp((c\beta - B)n)Z_n^{\beta,B}(\frac{1}{2})$. Hence, for $B \leq c\beta/2$, whp

$$Z_n^{\beta,B}(0) \geq \exp(c\beta n/2)Z_n^{\beta,B}(\tfrac{1}{2}). \tag{5.1}$$

By Lemma 2.3, we have

$$t_{\text{mix}} \geq \left(\frac{1}{2\Phi(U)} - 1\right)\log 2, \qquad \Phi(U) = \frac{Q(U, U^c)}{\mu_n^{\beta,B}(U)}, \tag{5.2}$$

for any $U \subset \Omega_n$ with $\mu_n^{\beta,B}(U) \leq \frac{1}{2} + o(1)$.

We start by proving Theorem 1.4(ib), for which we assume that $\beta > 12$ and consider the bottleneck ratio $\Phi(S)$ of the following set

$$S = \{\sigma : |\sigma_+| \leq \lceil n/2 \rceil\}.$$

Observe that in each step of the Glauber dynamics, we flip at most one spin. Therefore, after one step, the number of positive spin increases at most by one. Hence,

$$Q(S, S^c) = \sum_{\sigma \in S, \sigma' \in S^c} \mu_n^{\beta,B}(\sigma)P(\sigma, \sigma') = \sum_{\sigma \in A_{\lceil n/2 \rceil}, \sigma' \in A_{\lceil n/2 \rceil+1}} \mu_n^{\beta,B}(\sigma)P(\sigma, \sigma') \leq \mu_n^{\beta,B}(A_{\lceil n/2 \rceil}), \tag{5.3}$$

where, for $k \geq 0$,

$$A_k = \{\sigma \in \Omega_n : |\sigma_+| = k\}.$$

Combining this estimate and (5.1) with the fact that S contains A_0, we get that, whp,

$$\Phi(S) = \frac{Q(S, S^c)}{\mu_n^{\beta,B}(S)} \leq \frac{\mu_n^{\beta,B}(\sigma : |\sigma_+| = \lceil n/2 \rceil)}{\mu_n^{\beta,B}(\sigma : |\sigma_+| = 0)} = \frac{Z_n^{\beta,B}(\frac{1}{2})}{Z_n^{\beta,B}(0)} \leq \exp(-c\beta n/2). \tag{5.4}$$

To apply (5.2), we need to show $\mu_n^{\beta,B}(S) \leq \frac{1}{2} + o(1)$ whp. Observe that

$$\sum_{\sigma : |\sigma_+| \leq \lceil n/2 \rceil} \exp(-H_n^{\beta,B}(\sigma)) \leq \sum_{\sigma : |\sigma_+| \leq \lceil n/2 \rceil} \exp(-H_n^{\beta,0}(\sigma))$$

$$\leq \sum_{\sigma : |\sigma_+| > \lceil n/2 \rceil} \exp(-H_n^{\beta,0}(\sigma)) \leq \sum_{\sigma : |\sigma_+| > \lceil n/2 \rceil} \exp(-H_n^{\beta,B}(\sigma)).$$

Hence,

$$\mu_n^{\beta,B}(\sigma \colon |\sigma_+| \le [n/2]) \le \tfrac{1}{2}.$$

In addition, by (5.1),

$$\mu_n^{\beta,B}(\sigma \colon |\sigma_+| = \lceil n/2 \rceil) = o(1).$$

Combining the last two estimates we get $\mu_n^{\beta,B}(S) \le \tfrac{1}{2} + o(1)$ whp. In conclusion, by (5.2) and (5.4), $t_{\mathrm{mix}} \ge \exp(c\beta n/4)$ whp if $\beta \ge 12$ and $B \le c\beta/4$.

We now continue to prove Theorem 1.4(ia). Let us fix $\beta > \beta_c$ and $B < B_c$. Then there exists $0 < t < u \le \tfrac{1}{2}$, such that

$$\varphi_{\beta,0}(t) - \varphi_{\beta,0}(u) \ge (u-t)(B_c + B),$$

which implies that

$$\varphi_{\beta,B}(t) - \varphi_{\beta,B}(u) \ge (u-t)(B_c - B),$$

since $\varphi_{\beta,B}(t) = \varphi_{\beta,0}(t) + B(2t-1)$. Let (n_k) be the subsequence along which $\{\varphi_n^{\beta,0}\}_{n \ge 0}$ converges almost surely to $\varphi_{\beta,0}$. (We can take such a subsequence thanks to Proposition 1.3.) Therefore, whp

$$\frac{Z_{n_k}^{\beta,B}(t)}{Z_{n_k}^{\beta,B}(u)} \ge \exp(n_k(u-t)(B_c - B)/2). \tag{5.5}$$

Now, by repeating the same arguments for Theorem 1.4(ia) (here we use $S = \{\sigma \colon |\sigma_+| = [un]\}$, for which it is obvious that $\mu_n^{\beta,B}(S) \le \tfrac{1}{2} + o(1)$), we obtain that

$$\lim_{k \to \infty} \mathbb{P}(t_{\mathrm{mix}}^{n_k} \ge \exp(cn_k)) = 1,$$

for some $c > 0$, where we write $t_{\mathrm{mix}}^{n_k}$ for the mixing time considered on the graph G_{n_k}. The above estimate proves Theorem 1.4(ia). □

5.2. *Rapid mixing in quenched setting: Proof of Theorem 1.4(ii).* We first recall a result by Mossel and Sly (2013) that gives a sufficient condition for the rapid mixing of Glauber dynamics. For a graph $G = (V, E)$ and vertex $v \in V$, we write $B(v, R)$ for the ball of radius R around v and we write $S(v, R) = B(v, R) \setminus B(v, R-1)$ for the sphere of radius R around v.

Theorem 5.1 (Fast mixing of Gibbs samplers, Mossel and Sly, 2013, Theorem 3). *Let G be a graph on $n \ge 2$ vertices such that there exist constants $R, T, \chi \ge 1$ such that the following three conditions hold for all $v \in V$:*

    ▷ **Volume:** *The volume of the ball $B(v, R)$ satisfies $|B(v, R)| \le \chi$.*
    ▷ **Local mixing:** *For any configuration η on $S(v, R)$, the mixing time of the Gibbs sampler on $B(v, R-1)$ with fixed boundary condition η is bounded above by T.*
    ▷ **Spatial mixing:** *For each vertex $u \in S := S(v, R)$, define*

$$a_u := \sup_{\eta^+, \eta^-} P\left(\sigma_v = + \mid \sigma_S = \eta^+\right) - P\left(\sigma_v = + \mid \sigma_S = \eta^-\right),$$

where the supremum is over configurations η^+, η^- on $S(v, R)$ differing only at u with $\eta_u^+ \equiv +, \eta_u^- \equiv -$. Then the spatial mixing assumption states that

$$\sum_{u \in S(v,R)} a_u \le \tfrac{1}{4}.$$

Under the above three conditions, there exists a positive constant C, such that the mixing time of the Gibbs sampler satisfies $t_{\mathrm{mix}} \le CT \log(\chi) n \log n$.

Remark 5.2 (Continuous vs. discrete time). In fact, Mossel and Sly (2013, Theorem 3) applies to the *continuous-time* Gibbs sampler and has a factor n less in the bound. However, by Mossel and Sly (2013, Lemma 15), the above bound with the extra factor n holds for the discrete-time system.

The volume and local mixing time conditions are verified for any graph with maximum degree at most d by Mossel and Sly (2013), as stated in the following two lemmas:

Lemma 5.3 (Volume bounds, Mossel and Sly, 2013, Lemma 1). *Let $G = (V, E)$ be a graph of maximal degree d. Then the volume of $B(v, R)$ is less than*

$$\chi = 1 + d \sum_{i=1}^{R} (d-1)^{i-1}.$$

Lemma 5.4 (Local mixing bounds, Mossel and Sly, 2013, Lemma 2). *Let $G = (V, E)$ be a graph of maximal degree d, and consider the ferromagnetic Ising model on G. Then, local mixing holds with*

$$T = 80 d^3 \chi^3 e^{5\beta d(\chi+1)}, \quad \chi = 1 + d \sum_{i=1}^{R} (d-1)^{i-1}.$$

By Lemmas 5.3–5.4, to apply Theorem 5.1, we only need to prove spatial mixing estimates, as we claim in the following proposition:

Proposition 5.5 (Spatial mixing bounds). *Suppose that $\beta > \beta_c$ and $B > B_c^G(\beta)$. Then there exists a large R such that the spatial mixing condition holds whp.*

Proof of Theorem 1.4(ii). Theorem 1.4(ii) follows immediately from Theorem 5.1 and Lemma 5.3, 5.4, and Proposition 5.5.   □

It remains to prove Proposition 5.5. We follow the approach by Mossel and Sly (2013) using Weitz's result to compare the Gibbs measure on graphs to the one on the tree of self-avoiding paths. For any $v \in V$, we denote the tree of paths in G starting at v that do not intersect themselves, except possibly at the terminal vertex of the path, by $T_{\mathrm{saw}} := T_{\mathrm{saw}}(G, v)$. By the construction, each path in T_{saw} can be naturally mapped to a vertex in G which is the terminal vertex. For any set $\Lambda \subset V$, let $\mathrm{W}(\Lambda) \subset T_{\mathrm{saw}}(G, v)$ be the pullback of this natural map. Using this, we relate configurations η_Λ to the corresponding configurations $\eta_{\mathrm{W}(\Lambda)}$ on $\mathrm{W}(\Lambda)$. Furthermore, it holds that $d_G(A, B) = d_{T_{\mathrm{saw}}}(\mathrm{W}(A), \mathrm{W}(B))$. Here, $d_G(A, B)$ is the graph distance in G between two vertex sets A and B, while with a slight abuse of notation, we denote the graph distance in T_{saw} between the vertex sets of the two sub-trees $\mathrm{W}(A)$ and $\mathrm{W}(B)$ by $d_{T_{\mathrm{saw}}}(\mathrm{W}(A), \mathrm{W}(B))$. The following lemma describes a relation between the Glauber dynamics for the Ising model on subsets of $V(G)$ and that on T_{saw}:

Lemma 5.6 (Glauber dynamics on G and on T_{saw}, Weitz, 2006, Theorem 3.1). *For a graph G and $v \in V(G)$, there exists $A \subseteq T_{\mathrm{saw}}(G, v)$ and a configuration ν_A on A such that for any $\Lambda \subset V$ and configuration η_Λ on Λ,*

$$P_G(\sigma_v = + \mid \sigma_\Lambda) = P_{T_{\mathrm{saw}}}(\sigma_v = + \mid \sigma_{\mathrm{W}(\Lambda) \setminus A} = \eta_{\mathrm{W}(\Lambda) \setminus A}, \sigma_A = \nu_A).$$

Here, the set A is the set of leaves in T_{saw} corresponding to the terminal vertices of paths which return to a vertex already visited by the path.

We note that the construction of ν_A is described in the proof of Mossel and Sly (2013, Lemma 13) and Weitz (2006, Theorem 3.1).

When G is d-regular then $T_{\mathrm{saw}}(G, v)$ is a d-regular tree, denoted by $\mathbb{T}_d$. We now prove the spatial mixing of the Glauber dynamics on $\mathbb{T}_d$ under the uniqueness regime of Gibbs measure, which is the main innovation of this paper in the rapid mixing regime:

Lemma 5.7 (Spatial mixing in the uniqueness regime)**.** *Suppose that $\beta > \beta_c$ and $B > B_c^G(\beta)$. Let o be the root of $\mathbb{T}_d$, let $U \subseteq \Lambda \subset \mathbb{T}_d$, and let η^+, η^- be two configurations that differ only on U with $\eta_U^+ \equiv +, \eta_U^- \equiv -$. Then there exist positive constants $\varkappa, C, K$ depending only on β and B, such that $\varkappa < 1/(d-1)$ and*

$$|P(\sigma_o = + \mid \sigma_\Lambda = \eta^+) - P(\sigma_o = + \mid \sigma_\Lambda = \eta^-)| \le C|U|\varkappa^{d(o,\Lambda)-K}.$$

Proof: For any $v \in \mathbb{T}_d$ and a boundary condition η, let us define

$$R_v^\eta = \frac{P(\sigma_v = - \mid \eta)}{P(\sigma_v = + \mid \eta)},$$

where $P(\cdot \mid \eta)$ is the Gibbs measure on the subtree $T_v \subseteq \mathbb{T}_d$ with boundary condition η outside T_v. It is well known that

$$R_v^\eta = e^{-2B} \prod_{w \prec v} L_\beta(R_w^\eta), \tag{5.6}$$

where $w \prec v$ means w is a child of v in $\mathbb{T}_d$, and

$$L_\beta(x) = \frac{e^{2\beta}x + 1}{e^{2\beta} + x}.$$

Equation (5.6) can be obtained by taking the conditional expectation of σ_v conditionally on σ_w for $w \prec v$, and noticing that on a tree these expectations factorize.

A boundary condition η can be characterized by its *support*, which is defined as $A_\eta = \{v \colon \eta_v \text{ is fixed}\}$, and the values of η on A_η. Then for any vertex v and boundary condition η, we set

$$d(v,\eta) = d(v, A_\eta). \tag{5.7}$$

For any ℓ, let us denote $R_{v,\ell}^+ = R_v^{\eta_{v,\ell}^+}$ (respectively, $R_{v,\ell}^- = R_v^{\eta_{v,\ell}^-}$), where $\eta_{v,\ell}^+$ (resp. $\eta_{v,\ell}^-$) is the plus (respectively minus) boundary condition on $T(v,\ell)$, the subtree of height ℓ starting at v. Using (5.6) and the fact that $L_\beta(\cdot)$ is an increasing function, we obtain that, for any boundary condition η,

$$R_{v,\ell}^- \le R_v^\eta \le R_{v,\ell}^+, \tag{5.8}$$

where $\ell = d(v,\eta)$. By (5.6), we have $\log R_{v,\ell+1}^+ = -2B + (d-1)\log L_\beta(R_{v,\ell}^+)$ and $\log R_{v,\ell+1}^- = -2B + (d-1)\log L_\beta(R_{v,\ell}^-)$. It is well known that if $\beta > \beta_c$ and $B > B_c^G(\beta)$, then the fixed point equation

$$r = e^{-2B}L_\beta(r)^{d-1} \tag{5.9}$$

has a unique solution, denoted by $r^\star$, and thus the Gibbs measure is unique (see the proof of Martinelli et al., 2004, Proposition 4.5). In particular, as $\ell \to \infty$

$$R_{v,\ell}^-, R_{v,\ell}^+ \to r^\star.$$

Hence, it follows from (5.8) that

$$R_v^\eta \to r^\star \qquad \text{as} \qquad d(v,\eta) \to \infty.$$

In particular, for any $\delta > 0$, there exists $K = K(\delta)$, such that if $d(v,\eta) \ge K$, then

$$|R_v^\eta - r^\star| \le \delta. \tag{5.10}$$

We are now ready to prove the lemma. First we consider the case that $U = \{u\}$. Then η^+ differs from η^- only at u. Let $o = u_0, \ldots, u = u_\ell$ be the geodesic path from o to u. We observe that for

$\delta > 0$ and K as in (5.10), if $d(u_1, \eta) \geq K$, then

$$|P(\sigma_o = + \mid \eta^+) - P(\sigma_o = + \mid \eta^-)| = \left| \frac{1}{1 + R_o^{\eta^+}} - \frac{1}{1 + R_o^{\eta^-}} \right|$$

$$\leq |R_o^{\eta^+} - R_o^{\eta^-}| = \mathrm{e}^{-2B} \prod_{w \prec o, w \neq u_1} L_\beta(R_w^\eta) |L_\beta(R_{u_1}^{\eta^+}) - L_\beta(R_{u_1}^{\eta^-})|$$

$$\leq \mathrm{e}^{-2B} \prod_{w \prec o, w \neq u_1} L_\beta(R_w^\eta) \max_{x \in [R_{u_1}^{\eta^-}, R_{u_1}^{\eta^+}]} L_\beta'(x) |R_{u_1}^{\eta^+} - R_{u_1}^{\eta^-}|$$

$$\leq \mathrm{e}^{-2B} \left(\max_{|x - r^\star| \leq \delta} L_\beta(x) \right)^{d-1} \left(\max_{|x - r^\star| \leq \delta} L_\beta'(x) \right) |R_{u_1}^{\eta^+} - R_{u_1}^{\eta^-}|. \tag{5.11}$$

Here, in the second line, we have used (5.6), and the fact that the boundary conditions η^+ and η^- constrained on T_w are the same if $w \neq u_1$, while in the third line we have used the mean-value theorem and (5.10) and in the last line we have used the fact that $L_\beta'(x) \geq 0$ for all x. Similarly, if $d(u_i, \eta) \geq K$ then

$$|R_{u_i}^{\eta^+} - R_{u_i}^{\eta^-}| \leq \mathrm{e}^{-2B} \prod_{w \prec u_i, w \neq u_{i+1}} L_\beta(R_w^\eta) \max_{x \in [R_{u_{i+1}}^{\eta^-}, R_{u_{i+1}}^{\eta^+}]} L_\beta'(x) |R_{u_{i+1}}^{\eta^+} - R_{u_{i+1}}^{\eta^-}|$$

$$\leq \mathrm{e}^{-2B} \left(\max_{|x - r^\star| \leq \delta} L_\beta(x) \right)^{d-2} \left(\max_{|x - r^\star| \leq \delta} L_\beta'(x) \right) |R_{u_{i+1}}^{\eta^+} - R_{u_{i+1}}^{\eta^-}|. \tag{5.12}$$

Note that u_i for $i \geq 1$ has only $(d-2)$ children instead of $(d-1)$ as the root.

Next, we claim that there is $\varkappa < 1/(d-1)$, such that

$$\mathrm{e}^{-2B} L_\beta(r^\star)^{d-2} L_\beta'(r^\star) < \varkappa. \tag{5.13}$$

We postpone the proof of this claim to Section 6, since we will use some properties of the equation (5.9) proved in Section 6.

We turn to prove the proposition. Using (5.11)–(5.13), for K large enough and $\ell = d(o, \eta) \geq K$,

$$|P(\sigma_o = + \mid \eta^+) - P(\sigma_o = + \mid \eta^-)| \leq \sup_x L_\beta(x) \varkappa^{\ell - K} |R_{u_{\ell - K}}^{\eta^+} - R_{u_{\ell - K}}^{\eta^-}|$$

$$\leq \sup_x L_\beta(x) \varkappa^{\ell - K} \sup_x L_\beta(x)^{d-1}$$

$$\leq \mathrm{e}^{2\beta d} \varkappa^{\ell - K}. \tag{5.14}$$

Note that in the first line, the term $\sup_x L_\beta(x)$ appears since in (5.11) the exponent of $\max_{|x-r^\star| \leq \delta} L_\beta(x)$ is $(d-1)$ instead of $(d-2)$ as in (5.12), and in the second line, we use (5.6) to obtain the uniform upper bound for R_u^η.

Now we consider the general case $U = \{u_1, \ldots, u_k\}$. Let $\eta^0, \ldots, \eta^k$ be the interpolation sequence of configurations on Λ with $\eta^0 = \eta^-$ and $\eta^k = \eta^+$, where η^i and η^{i-1} differ only at u_i with $\eta_{u_i}^i = +$ and $\eta_{u_i}^{i-1} = -$. By (5.14), for all i,

$$|P(\sigma_o = + \mid \eta^i) - P(\sigma_o = + \mid \eta^{i-1})| \leq \mathrm{e}^{2\beta d} \varkappa^{\ell - K},$$

and thus

$$|P(\sigma_o = + \mid \eta^+) - P(\sigma_o = + \mid \eta^-)| \leq k \mathrm{e}^{2\beta d} \varkappa^{\ell - K},$$

which proves Lemma 5.7 subject to (5.13). $\qquad \square$

With Lemma 5.7 in hand, we are now ready to complete the proof of Proposition 5.5:

Proof of Proposition 5.5. For any v, let T be the tree of self-avoiding paths starting at v. Let $\mathrm{W}(S(v, R))$ denote the vertices in T that correspond to vertices in $S(v, R)$, and, for each $u \in S(v, R)$,

let $\mathrm{W}(u)$ denote the set of vertices in T that correspond to u. Then, by Lemmas 5.6 and 5.7,

$$a_u = \sup_{\eta^+,\eta^-} \left| P\left(\sigma_v = + \mid \sigma_{\mathrm{W}(S(v,R))\backslash A} = \eta^+_{\mathrm{W}(S(v,R))\backslash A}, \sigma_A = \nu_A\right) \right.$$

$$\left. -P\left(\sigma_v = + \mid \sigma_{\mathrm{W}(S(v,R))\backslash A} = \eta^-_{\mathrm{W}(S(v,R))\backslash A}, \sigma_A = \nu_A\right) \right|$$

$$\leq C|\mathrm{W}(u)|\varkappa^{R-K}, \tag{5.15}$$

since $d(v,\eta) \geq R$. We observe that for any fixed R,

$$\lim_{n\to\infty} \mathbb{P}\left(B(v,R) \text{ has at most 1 cycle for all } v \in [n]\right) = 1,$$

since by the construction of d-regular configuration models using the random pairing of half-edges as described in Section 1.1, for each $v \in [n]$, the probability that $B(v,R)$ has at least 2 cycles is of order $(d^R/n)^2 = \mathcal{O}(n^{-2})$. Assume that $B(v,R)$ has at most 1 cycle. Then every $u \in S(v,R)$ appears at most twice in the tree of self-avoiding paths, which gives $|\mathrm{W}(u)| \leq 2$. Therefore, (5.15) implies that

$$\sum_{u\in S(v,R)} a_u \leq 2C(d-1)^R \varkappa^{R-K} \leq \tfrac{1}{4}, \tag{5.16}$$

for some $R = R(K)$ large enough, since $\varkappa < 1/(d-1)$. This completes the proof of Proposition 5.5. $\square$

Remark 5.8 (Spatial mixing and cycles). *The inequality* (5.16) *still holds if* $|\mathrm{W}(u)| \leq \varkappa_1^R$, *with some* $\varkappa_1 < 1/((d-1)\varkappa)$ *and* R *large. Therefore, spatial mixing holds for graphs satisfying that the ball* $B(v,R)$ *has at most* $\varkappa_1^R$ *cycles for every* $v \in V$.

6. Identification of B_c^G: Proof of Proposition 1.5

We recall from (2.3) that the equation of annealed critical points is $\hat{\varphi}'_{\beta,B}(t) = 0$, or

$$\log\left(\frac{1-t}{t}\right) + 2B + d\log f_\beta(t) = 0, \tag{6.1}$$

where f_β is as in (1.5).

On the other hand, the equation (5.9) can be rewritten as

$$\log(rL_\beta(r)) + 2B - d\log L_\beta(r) = 0. \tag{6.2}$$

We aim to show that the two equations (6.1) and (6.2) are equivalent. We consider the following change of variable $t = p(r)$, where

$$p(r) = \frac{1}{rL_\beta(r)+1} = \frac{e^{2\beta}+r}{e^{2\beta}r^2+2r+e^{2\beta}}.$$

We remark that the function $p(r)$ is strictly decreasing and is a bijective map from $(0,\infty)$ to $(0,1)$. Then

$$r = p^{-1}(t) = \frac{e^{-2\beta}(1-2t) + \sqrt{1+(e^{-4\beta}-1)(1-2t)^2}}{2t} = \frac{f_\beta(t)(1-t)}{t}. \tag{6.3}$$

Thus

$$\frac{1-t}{t} = rL_\beta(r), \quad f_\beta(t) = \frac{tr}{1-t} = \frac{1}{L_\beta(r)}. \tag{6.4}$$

Combining (6.1), (6.2) and (6.4), we obtain that the two equations (6.1) and (6.2) are equivalent. In particular,

$$\hat{B}_c = B_c^G,$$

since

$$\hat{B}_c = \inf\{B > 0: \text{ equation (6.1) has a unique solution}\},$$
$$B_c^G = \inf\{B > 0: \text{ equation (6.2) has a unique solution}\}.$$

This completes the proof of Proposition 1.5. $\qquad\square$

Proof of (5.13). We need to prove that if $B > B_c^G$ then

$$e^{-2B} L_\beta(r^\star)^{d-2} L_\beta'(r^\star) < \frac{1}{d-1}, \tag{6.5}$$

where $r^\star$ is the unique solution to (5.9). Let us consider the function

$$g(r) = r - e^{-2B} L_\beta(r)^{d-1}.$$

Then $g(0) < 0$ and $r^\star$ is the unique solution to the equation $g(r) = 0$. Therefore $g'(r^\star) \geq 0$, since otherwise the equation $g(r) = 0$ has at least two solutions. In other words,

$$0 \leq g'(r^\star) = 1 - (d-1)e^{-2B} L_\beta(r^\star)^{d-2} L_\beta'(r^\star). \tag{6.6}$$

Hence, to prove (6.5), we only need to show that $g'(r^\star) \neq 0$.

Assume that $g'(r^\star) = 0$. We shall show that this assumption leads to a contradiction. Since $B > B_c^G = \hat{B}_c$, the equation $\hat{\varphi}_{\beta,B}(t) = 0$ has a unique solution $t^\star \in (\frac{1}{2}, 1)$ (which is indeed t_1^B). As we have shown in (6.3), the unique solution $r^\star$ of (6.2) satisfies

$$r^\star = p^{-1}(t^\star) \in (0, 1), \tag{6.7}$$

since $t^\star \in (\frac{1}{2}, 1)$.

Suppose that $g''(r^\star) < 0$. Then there exists $\epsilon > 0$ such that $r^\star + \epsilon < 1$ and $g''(x) < 0$ for all $x \in [r^\star, r^\star + \epsilon]$. Hence, by the Taylor expansion, we have $g(r^\star + \epsilon) = g(r^\star) + \epsilon g'(r^\star) + \frac{1}{2} g''(r^\star) x^2 < 0$ for some $x \in (r^\star, r^\star + \epsilon)$. Moreover, $g(1) = 1 - e^{-2B} > 0$. Thus the equation $g(r) = 0$ has a solution in $(r^\star + \epsilon, 1)$, so this equation has at least two solutions in $(0, 1)$, which is a contradiction. Using the same argument and the fact that $g(0) < 0$, we can also prove that $g''(r^\star) > 0$ leads to a contradiction. Therefore, we have $g''(r^\star) = 0$.

Since $g'(r^\star) = g(r^\star) = 0$,

$$1 = e^{-2B}(d-1) L_\beta(r^\star)^{d-2} L_\beta'(r^\star), \qquad r^\star = e^{-2B} L_\beta(r^\star)^{d-1}.$$

Thus

$$r^\star = \frac{L_\beta(r^\star)}{(d-1) L_\beta'(r^\star)} = \frac{(e^{2\beta} r^\star + 1)(e^{2\beta} + r^\star)}{(d-1)(e^{4\beta} - 1)}. \tag{6.8}$$

On the other hand, $g''(r^\star) = 0$ implies that

$$(d-2) L'(r^\star)^2 + L(r^\star) L''(r^\star) = 0,$$

or equivalently

$$(d-2)(e^{4\beta} - 1) - 2(e^{2\beta} r^\star + 1) = 0. \tag{6.9}$$

Combining (6.8) and (6.9) yields that

$$r^\star = \frac{(d-2)(e^{2\beta} + r^\star)}{2(d-1)} \quad \Rightarrow \quad r^\star = \frac{(d-2)e^{2\beta}}{d} > 1,$$

since $\beta > \beta_c = \operatorname{atanh}(1/(d-1))$. This is contrary to (6.7).

In conclusion, $g'(r^\star) > 0$ and thus (6.5) (or equivalently (5.13)) follows. $\qquad\square$

Remark 6.1 (Relation constrained annealed measure and d-ary tree with minus boundary condition).
Fix $\beta > \beta_c$ and $0 \leq B < \hat{B}_c$. Then the curve of $\hat{\varphi}_{\beta,B}$ has three critical values $t_3^B < t_2^B < t_1^B$. Let $\hat{\mu}_n^-$ be the annealed measure restricted to $\Omega_n^- = \{\sigma \colon |\sigma_+| \leq \lceil t_2^B n \rceil\}$. More precisely,

$$\hat{\mu}_n^-(\sigma) = \frac{\mathbb{E}[\exp(-H_n(\sigma))]}{\mathbb{E}[Z_n^-]} \mathbb{1}_{\{\sigma \in \Omega_n^-\}}, \tag{6.10}$$

where

$$Z_n^- = \sum_{\sigma \in \Omega_n^-} \exp(-H_n(\sigma)).$$

Here we omit the superscript β, B to simplify the notation. Using the same arguments in Can (2019, Theorem 1.1), we can prove that

$$\lim_{n \to \infty} \mathbb{E}_{\hat{\mu}_n^-}\left(\frac{\sigma_1 + \cdots + \sigma_n}{n}\right) = 2t_3^B - 1. \tag{6.11}$$

Now we consider the magnetization of the root, denoted by ρ, in $\mathbb{T}_d$ with minus boundary conditions at level ℓ. Recall from (5.6) that, as $\ell \to \infty$,

$$R_{\rho,\ell}^\eta = \frac{P(\sigma_\rho = - \mid \eta_{\rho,\ell}^-)}{P(\sigma_\rho = + \mid \eta_{\rho,\ell}^-)} \to \mathrm{e}^{-2B} L_\beta(r_-)^d,$$

where $\eta_{\rho,\ell}^-$ is the minus boundary condition on the level ℓ of $\mathbb{T}_d$ and r_- is a solution to (5.9), or equivalently $r_- = \mathrm{e}^{-2B} L_\beta(r_-)^{d-1}$. Therefore,

$$\lim_{\ell \to \infty} R_{\rho,\ell}^\eta = r_- L_\beta(r_-),$$

and thus

$$\lim_{\ell \to \infty} \langle \sigma_\rho \rangle_{\eta_{\rho,\ell}^-} = \lim_{\ell \to \infty} \frac{1 - R_{\rho,\ell}^\eta}{1 + R_{\rho,\ell}^\eta} = \frac{1 - r_- L_\beta(r_-)}{1 + r_- L_\beta(r_-)}. \tag{6.12}$$

As we have shown above, t_3^B and r_- are related by

$$t_3^B = \frac{1}{r_- L_\beta(r_-) + 1}.$$

Combining this with (6.11) and (6.12), we obtain that

$$\lim_{n \to \infty} \langle \sigma_{U_n} \rangle_{\hat{\mu}_n^-} = \lim_{\ell \to \infty} \langle \sigma_\rho \rangle_{\eta_{\rho,\ell}^-}, \tag{6.13}$$

where U_n is a random vertex chosen uniformly in $[n]$. We conclude that restricting the annealed measure to Ω_n^- has the same effect in the large n limit as the minus boundary conditions for the Ising model on the d-ary tree.

7. Annealed cut-off behavior: Proof of Proposition 3.4

Let us recall the statement of Proposition 3.4. For any $\gamma > 0$, we define

$$T_n^+(\gamma) = c_\star n \log n + \gamma n, \qquad T_n^-(\gamma) = c_\star n \log n - \gamma n,$$

where

$$c_\star = -[s^\star(1 - s^\star)G''(s^\star)]^{-1} > 0.$$

The aim of Proposition 3.4 is to prove that, under Conditions (C1) and (C3),

$$\lim_{\gamma \to \infty} \liminf_{n \to \infty} d_n(T_n^-(\gamma)) = 1, \tag{7.1}$$

and

$$\lim_{\gamma \to \infty} \limsup_{n \to \infty} d_n(T_n^+(\gamma)) = 0. \tag{7.2}$$

We give the proof in three steps. In the first step in Section 7.1, we analyse the projection chain $(X_t)_{t \geq 0}$ defined in Section 3.2. In the second step in Section 7.2, we prove (7.1) and in the final step in Section 7.3, we prove (7.2).

7.1. *A careful analysis of the projection chain.* The projection chain $(X_t)_{t \geq 0}$ has been defined as

$$X_t = |\{i \colon \xi_t(i) = 1\}|,$$

where $(\xi_t)_{t \geq 0}$ is the Glauber dynamics. Then $(X_t)_{t \geq 0}$ is a birth-death process on $\{0, \dots, n\}$ with probability transitions given in (3.4). The drift R_t of $(X_t)_{t \geq 0}$ satisfies

$$
\begin{aligned}
R_t \quad &:= \quad \mathbb{E}\left[X_{t+1} - X_t \mid X_t \right] \\
&= \quad \frac{n - X_t}{n} \times \frac{\mathrm{e}^{nF_n(X_t+1)}}{\mathrm{e}^{nF_n(X_t+1)} + \mathrm{e}^{nF_n(X_t)}} - \frac{X_t}{n} \times \frac{\mathrm{e}^{nF_n(X_t-1)}}{\mathrm{e}^{nF_n(X_t-1)} + \mathrm{e}^{nF_n(X_t)}}.
\end{aligned}
\tag{7.3}
$$

Then, by (3.7),

$$\left| R_t - R\left(X_t/n \right) \right| \leq \frac{C}{n}, \tag{7.4}$$

where

$$R(s) = (1-s)\frac{\mathrm{e}^{F'(s)}}{\mathrm{e}^{F'(s)} + 1} - s\frac{1}{\mathrm{e}^{F'(s)} + 1} = \frac{\mathrm{e}^{F'(s)}}{\mathrm{e}^{F'(s)} + 1} - s. \tag{7.5}$$

Combining (7.3) and (7.4), we get

$$X_t + R(X_t/n) - C/n \leq \mathbb{E}[X_{t+1} \mid X_t] \leq X_t + R(X_t/n) + C/n. \tag{7.6}$$

We recall that $G''(s^\star) < 0$ and $G'(s^\star) = 0$. Therefore, $I'(s^\star) + F'(s^\star) = 0$, or equivalently $s^\star = \mathrm{e}^{F'(s^\star)}/(\mathrm{e}^{F'(s^\star)} + 1)$. Thus, $R(s^\star) = 0$ and, using that $G''(s) = F''(s) - [s(1-s)]^{-1}$,

$$R'(s^\star) = \frac{\mathrm{e}^{F'(s^\star)}F''(s^\star)}{(\mathrm{e}^{F'(s^\star)} + 1)^2} - 1 = s^\star(1 - s^\star)F''(s^\star) - 1 = s^\star(1 - s^\star)G''(s^\star) = -\frac{1}{c_\star} < 0. \tag{7.7}$$

By (7.7) and $R(s^\star) = 0$, we can find a positive constant δ such that

$$R'(s) < 0 \text{ for all } s \in (s^\star - 2\delta, s^\star + 2\delta) \subset (0, 1). \tag{7.8}$$

We define, for $\delta \in (0, 1)$

$$\Omega_n^\delta = \left\{ \sigma \in \Omega_n \colon s^\star - \delta \leq \frac{|\sigma_+|}{n} \leq s^\star + \delta \right\}. \tag{7.9}$$

Before going into the full details, let us start by summarizing the main steps in the analysis of the projection chain $(X_t)_{t \geq 0}$. Thanks to the assumption that the function G only has one local maximizer at $s^\star$, the measure of $(X_t)_{t \geq 0}$ is highly concentrated around $ns^\star$. Moreover, we will show in Lemma 7.1 that the chain $(X_t)_{t \geq 0}$ accesses Ω_n^δ very quickly, namely, after a time of order n. Then we consider the chain after entering Ω_n^δ. Thus, next assume that $X_0 \in \Omega_n^\delta$. Then we will show that the drift function $R(s)$ plays a central role to describe the time for X_t to reach $ns^\star$. Roughly speaking, in Lemma 7.4 below, we show that

$$Y_{t+1} - Y_t \approx \left(1 + \frac{R''(s^\star)}{n} \right) Y_t,$$

where $Y_t = X_t - ns^\star$. Notice that $R''(s^\star) < 0$, and, therefore, when $Y_0 \in [-\delta n, \delta n]$, the time for Y_t to reach $[-\mathcal{O}(\sqrt{n}), \mathcal{O}(\sqrt{n})]$ is $(-R''(s^\star))^{-1} n \log n \pm \mathcal{O}(n)$. Finally, we prove in Lemma 7.7 that if the starting points of $(X_t)_{t \geq 0}$ and $(\tilde{X}_t)_{t \geq 0}$ are both in $[ns^\star - \mathcal{O}(\sqrt{n}), ns^\star + \mathcal{O}(\sqrt{n})]$, then the two chains mix after a time of order n. Combining the above steps, we conclude that the mixing time of $(X_t)_{t \geq 0}$ concentrates around $c_\star n \log n$ (note here that $c_\star = (-R''(s^\star))^{-1}$) with a window of order n.

We now present the details of the argument. We start by proving that Ω_n^δ is hit quickly:

Lemma 7.1 (Ω_n^δ is hit quickly). *For δ as in (7.8), there exists a positive constant $C = C(\delta)$, such that*

$$\sup_{\sigma \in \Omega_n} \mathbb{P}_\sigma \left(\inf\{t \colon \xi_t \in \Omega_n^\delta\} \geq Cn \right) = \mathcal{O}(1/n).$$

Below, we give a sketch of the proof, the full proof can be found in Can et al. (2019, Appendix A.1) of the extended version.

Proof: Since G is strictly increasing in $(0, s^\star)$ and strictly decreasing in $(s^\star, 1)$, there exists a positive constant ε, such that

$$G'(s) \geq \varepsilon \quad \text{if} \quad s \in (0, s^\star - \delta) \quad \text{and} \quad G'(s) \leq -\varepsilon \quad \text{if} \quad s \in (s^\star + \delta, 1). \tag{7.10}$$

As in (3.9) and (3.10), the stationary measure of $(X)_{t \geq 0}$ satisfies

$$\nu_n(k) \sim \pi_n(k) \asymp \exp\left(n\big[G(k/n) - G(0) \big] \right). \tag{7.11}$$

For each k and ℓ, we define the waiting time for (X_t) going from k to ℓ as

$$\tau_{k \to \ell} = \inf\{t : X_t = \ell \mid X_0 = k\}.$$

To prove Lemma 7.1, it suffices to show that

$$\mathbb{P}\left(\max\{\tau_{n \to [n(s^\star + \delta)]}, \tau_{0 \to [n(s^\star - \delta)]}\} \geq Cn \right) = \mathcal{O}(1/n), \tag{7.12}$$

for some positive constant $C > 0$. A standard computation for the birth-death chain (see e.g. Barrera et al., 2009, Proposition 2) gives that

$$\mathbb{E}(\tau_{k \to k-1}) = \frac{1}{q_n(k)} \sum_{j=k}^{n} \frac{\pi_n(j)}{\pi_n(k)},$$

$$\mathbb{E}(\tau_{k \to k-1}^2) = \frac{2}{q_n(k)\pi_n(k)} \sum_{j=k}^{n} \mathbb{E}(\tau_{j \to k-1})\pi_n(j) - \mathbb{E}(\tau_{k \to k-1}).$$

By (3.4), $q_n(k) \asymp (k/n)$. Thus using (7.10) and (7.11), we obtain that

$$\sum_{j=k}^{n} \frac{\pi_n(j)}{\pi_n(k)} \asymp \sum_{j=k}^{n} \exp\left(n\big[G(j/n) - G(k/n) \big] \right) = \mathcal{O}(1). \tag{7.13}$$

Therefore,

$$\mathbb{E}(\tau_{k \to k-1}) = \mathcal{O}(n/k), \quad \mathbb{E}(\tau_{k \to k-1}^2) = \mathcal{O}(1).$$

Hence,

$$\max\{\mathbb{E}(\tau_{n \to [n(s^\star + \delta)]}), \mathrm{Var}(\tau_{n \to [n(s^\star + \delta)]})\} = \mathcal{O}(n),$$

which implies by using Chebyshev's inequality that

$$\mathbb{P}\left(\tau_{n \to [n(s^\star + \delta)]} \geq Cn \right) = \mathcal{O}(1/n),$$

for some $C > 0$. A similar estimate holds for $\tau_{0 \to [n(s^\star - \delta)]}$, and then we get (7.12). □

Lemma 7.2 (It takes long to leave $\Omega_n^{2\delta}$). *For each $\gamma > 0$, there exists a positive constant $C = C(\gamma)$, such that*

$$\sup_{\sigma \in \Omega_n^\delta} \mathbb{P}_\sigma \left(\inf\{t \colon \xi_t \notin \Omega_n^{2\delta}\} \leq T_n^+(\gamma) \right) \leq \frac{C(\log n)^2}{n^2}.$$

Again, we give a sketch of the proof below, the full proof can be found in Can et al. (2019, Appendix A.1) of the extended version.

Proof: By (7.7), it holds that

$$a := -\inf\{R'(s) : |s - s^\star| \le 2\delta\} > 0.$$

Therefore there exists a function $R_{2\delta} : [0,1] \mapsto \mathbb{R}$, such that

$$R_{2\delta}(s) = R(s) \text{ for } |s - s^\star| \le 2\delta \text{ and } R'_{2\delta}(s) \le -a \text{ for all } s \in [0,1]. \qquad (7.14)$$

Define also

$$T_1 = \inf\{t : |X_t - ns^\star| \ge 2n\delta\}.$$

Observe that if $|X_0 - ns^\star| \le 2n\delta$ then the waiting time T_1 depends only on the transition probabilities $\{(p_n(k), q_n(k), r_n(k)) : |k - ns^\star| \le 2n\delta\}$. We define an auxiliary birth-death chain (X'_t) with transition probabilities $\{(p'_n(k), q'_n(k), r'_n(k)) : k \in \mathbb{Z}\}$ defined as

$$(p'_n(k), q'_n(k), r'_n(k)) = \begin{cases} (p_n(k), q_n(k), r_n(k)) & \text{if} \quad |k - ns^\star| \le 2n\delta, \\ (R_{2\delta}(k/n), 0, 1 - R_{2\delta}(k/n)) & \text{otherwise.} \end{cases} \qquad (7.15)$$

Then $R_{2\delta}$ is related to the drift of $(X'_t)_{t \ge 0}$ by

$$\mathbb{E}[X'_{t+1} - X'_t \mid X'_t] = R_{2\delta}(X'_t/n) + \mathcal{O}(\tfrac{1}{n}).$$

Moreover,

$$\sup_{|x_0 - ns^\star| \le \delta n} \mathbb{P}_{x_0}(T_1 \le T_n^+(\gamma)) = \sup_{|x'_0 - ns^\star| \le \delta n} \mathbb{P}_{x'_0}(T'_1 \le T_n^+(\gamma)),$$

where

$$T'_1 = \inf\{t : |X'_t - ns^\star| \ge 2n\delta\}.$$

By the strong Markov property,

$$\sup_{|x'_0 - ns^\star| \le \delta n} \mathbb{P}_{x'_0}(T'_1 \le T_n^+(\gamma))$$

$$\le \sup_{|x'_0 - ns^\star| \le \delta n} \mathbb{P}_{x'_0}(T'_2 \le T_n^+(\gamma)) \times \sup_{|x'_1 - ns^\star| \le 3n\delta/2} \mathbb{P}_{x'_1}(T'_1 \le T_n^+(\gamma)), \qquad (7.16)$$

where

$$T'_2 = \inf\{t : |X'_t - ns^\star| > 3n\delta/2\}.$$

By using the same arguments as in Ding et al. (2009a, Lemma 4.12), we can prove that

$$\sup_{0 \le t \le T_n^+(2\gamma)} |\mathbb{E}(X'_t) - ns^\star| \le 9n\delta/8, \quad \mathrm{Var}(X'_t) = \mathcal{O}(n). \qquad (7.17)$$

The crucial point in the proof of (7.17) is the contraction property of $\mathrm{Var}(X'_t)$ that $\mathrm{Var}(X'_{t+1}) \le \mathrm{Var}(X'_t)(1 - \frac{\varepsilon}{n})$ for some $\varepsilon > 0$. As for Ding et al. (2009a, Lemma 4.12), this contraction property can be shown when the drift function satisfies $f(x) = R_{2\delta}(x) - ax$ is decreasing for some $a > 0$ (this holds by (7.14)). Note that in Ding et al. (2009a, Lemma 4.12) the authors proved (7.17) directly for $(X_t)_{t \ge 0}$ with the function $f(x) = \tanh(\beta x) - (1 - \varepsilon)x$, for some $\varepsilon > 0$. In our case, we do not know the behavior of the drift function outside the interval $(s^\star - 2\delta, s^\star + 2\delta)$ well, so we need to consider the modification chain $(X'_t)_{t \ge 0}$.

It follows from (7.17) and Chebyshev's inequality that

$$\sup_{0 \le t \le T_n^+(2\gamma)} \mathbb{P}\left(|X'_t - ns^\star| \ge 5n\delta/4\right) \le \sup_{0 \le t \le T_n^+(2\gamma)} \mathbb{P}\left(|X'_t - \mathbb{E}(X'_t)| \ge n\delta/8\right) = \mathcal{O}(1/n).$$

As a consequence,

$$\mathbb{E}(N') \le C_1 \log n, \qquad (7.18)$$

where $C_1 = C_1(a, \delta)$ is a large constant and

$$N' = \#\{t \le T_n^+(2\gamma) : |X'_t - ns^\star| \ge 5n\delta/4\}.$$

Since $|X'_{t+1} - X'_t| \leq 1$ for all t, if $|X'_{t_0} - ns^\star| \geq 3n\delta/2$ for some t_0 then $|X'_t - ns^\star| \geq 5n\delta/4$ for all $t_0 \leq t \leq t_0 + \delta n/4$. Hence,

$$\mathbb{E}[N' \mid T'_2 \leq T_n^+(\gamma)] \geq \delta n/4. \tag{7.19}$$

Combining (7.18) and (7.19), we arrive at

$$\sup_{|x'_0 - ns^\star| \leq \delta n} \mathbb{P}_{x'_0}(T'_2 \leq T_n^+(\gamma)) \leq \frac{\mathbb{E}[N']}{\mathbb{E}[N \mid T'_2 \leq T_n^+(\gamma)]} = \frac{\mathcal{O}(\log n)}{n}.$$

Similarly, we also have

$$\sup_{|x'_0 - ns^\star| \leq 3\delta n/2} \mathbb{P}_{x'_0}(T'_1 \leq T_n^+(\gamma)) = \frac{\mathcal{O}(\log n)}{n}.$$

Now combining the last two estimates and (7.16), we arrive at the claim in Lemma 7.2. □

Lemma 7.3 (A variance bound on X_t). *There exists a positive constant C, such that for all $\gamma > 1$ and $n \geq n_0 = n_0(\gamma)$, with n_0 a large number,*

$$\sup_{\sigma \in \Omega_n} \mathrm{Var}_\sigma(X_t) \leq Cn \qquad \text{for all} \qquad Cn \leq t \leq T_n^+(\gamma).$$

Proof: Fix $C > 0$ sufficiently large. Define

$$\mathcal{E} = \{|X_t - ns^\star| \leq 2\delta n \quad \text{for all} \quad Cn \leq t \leq T_n^+(\gamma)\}.$$

By Lemmas 7.1 and 7.2,

$$\mathbb{P}(\mathcal{E}^c) = \mathcal{O}(1/n).$$

Therefore,

$$\mathrm{Var}(X_t) \leq \mathrm{Var}(X_t \mathbb{1}_{\mathcal{E}}) + \mathbb{E}(X_t^2 \mathbb{1}_{\mathcal{E}^c}) = \mathrm{Var}(X_t \mathbb{1}_{\mathcal{E}}) + \mathcal{O}(n). \tag{7.20}$$

Suppose that $\mathcal{E}$ happens. Then $(X_{t+Cn})_{t \geq 0}$ has the same law as the auxiliary chain $(X'_t)_{t \geq 0}$ defined in (7.15) in the proof of Lemma 7.2 for all $Cn \leq t \leq T_n^+(\gamma)$. Therefore,

$$\sup_{\sigma \in \Omega_n} \mathrm{Var}_\sigma(X_t \mathbb{1}_{\mathcal{E}}) \leq \sup_{\sigma \in \Omega_n^{2\delta}} \mathrm{Var}_\sigma(X'_{t-Cn} \mathbb{1}_{\mathcal{E}'}), \tag{7.21}$$

where

$$\mathcal{E}' = \{|X'_s - ns^\star| \leq 2\delta n \quad \text{for all} \quad s \leq T_n^+(\gamma - C)\}.$$

By Lemma 7.2,

$$\mathbb{P}(\mathcal{E}'^c) = \mathcal{O}((\log n)^2/n^2).$$

Hence, by (7.17) for all $0 \leq s \leq T_n^+(\gamma - C)$,

$$\begin{aligned}
\mathrm{Var}(X'_s \mathbb{1}_{\mathcal{E}'}) &\leq \mathrm{Var}(X'_s) + \mathbb{E}(X'^2_s \mathbb{1}_{\mathcal{E}'^c}) + 2\mathbb{E}(X'_s)\mathbb{E}(X'_s \mathbb{1}_{\mathcal{E}'^c}) \\
&\leq \mathrm{Var}(X'_s) + 3n^2 \mathbb{P}(\mathcal{E}'^c) = \mathcal{O}(n).
\end{aligned}$$

Combining this with (7.20) and (7.21) we arrive at the claim in Lemma 7.3. □

Lemma 7.4 (Relating $X_{T_n^\pm(\gamma)}$ to $ns^\star$). *The following assertions hold:*

    (i) *For any $\alpha > 0$,*

$$\lim_{\gamma \to \infty} \limsup_{n \to \infty} \mathbb{P}_{\mathbf{1}}(X_{T_n^-(\gamma)} \leq ns^\star + \alpha\sqrt{n}) = 0.$$

    (ii)

$$\lim_{\ell \to \infty} \lim_{\gamma \to \infty} \limsup_{n \to \infty} \sup_{\sigma \in \Omega_n} \mathbb{P}_\sigma(|X_{T_n^+(\gamma)} - ns^\star| \geq \ell\sqrt{n}) = 0.$$

Proof: By Lemma 7.1, the chain $(X_t)_{t\geq 0}$ is in Ω_n^δ after $\mathcal{O}(n)$ steps with probability $1 - \mathcal{O}(1/n)$. Therefore, in the proof of Lemma 7.4, we can assume that $X_0 \in \Omega_n^\delta$ by replacing $T_n^\pm(\gamma)$ by $T_n^\pm(\gamma - C)$.

By Lemma 7.2,

$$\mathbb{P}(T_1 \leq T_n^+(\gamma)) = \frac{\mathcal{O}((\log n)^2)}{n^2}, \tag{7.22}$$

where

$$T_1 = \inf\{t \colon |X_t - ns^\star| \geq 2n\delta\}.$$

Define

$$Y_t = X_t - ns^\star \qquad \text{and} \qquad \tilde{Y}_t = Y_t \mathbb{1}_{\{t \leq T_1\}}.$$

Then,

$$\begin{aligned}
\mathbb{E}[\tilde{Y}_{t+1} - \tilde{Y}_t \mid X_t] &= \mathbb{E}[Y_{t+1}\mathbb{1}_{\{t+1\leq T_1\}} - Y_t\mathbb{1}_{\{t\leq T_1\}} \mid X_t] \\
&= \mathbb{E}[(Y_{t+1} - Y_t)\mathbb{1}_{\{t\leq T_1\}} - Y_{t+1}\mathbb{1}_{\{T_1=t+1\}} \mid X_t] \\
&= \mathbb{1}_{\{t\leq T_1\}}\mathbb{E}[X_{t+1} - X_t \mid X_t] - \mathbb{E}[Y_{t+1}\mathbb{1}_{\{T_1=t+1\}} \mid X_t].
\end{aligned}$$

Combining this with (7.6) and the fact that $Y_{t+1} \leq n$, we get

$$\left|\mathbb{E}[\tilde{Y}_{t+1} - \tilde{Y}_t \mid X_t] - R(X_t/n)\mathbb{1}_{\{t\leq T_1\}}\right| \leq \frac{C}{n} + n\mathbb{P}(T_1 = t + 1 \mid X_t). \tag{7.23}$$

If $t \leq T_1$ then $X_t/n \in (s^\star - 2\delta, s^\star + 2\delta)$, and thus using $R(s^\star) = 0$ we get

$$\left|R(X_t/n)\mathbb{1}_{\{t\leq T_1\}} - R'(s^\star)\tilde{Y}_t/n\right| = \left|R\left(s^\star + \tilde{Y}_t/n\right) - R'(s^\star)\tilde{Y}_t/n\right| \leq C_1\left(\tilde{Y}_t/n\right)^2,$$

where $C_1 = \sup_{s\in(s^\star-2\delta,s^\star+2\delta)}|R''(s)|/2$. Combining this estimate with (7.23) leads to

$$\left|\mathbb{E}[\tilde{Y}_{t+1} - \tilde{Y}_t \mid X_t] - R'(s^\star)\tilde{Y}_t/n\right| \leq C_1\left(\tilde{Y}_t/n\right)^2 + n\mathbb{P}(T_1 = t + 1 \mid X_t) + \frac{C}{n},$$

and thus by taking the expectation and using (7.22), we get, for $t \leq T_n^+(\gamma)$,

$$\left|\mathbb{E}[\tilde{Y}_{t+1}] - \left(1 + \frac{R'(s^\star)}{n}\right)\mathbb{E}[\tilde{Y}_t]\right| \leq \frac{C_1\mathbb{E}[(\tilde{Y}_t)^2]}{n^2} + \frac{2C}{n}. \tag{7.24}$$

By Lemma 7.3, $\mathrm{Var}(\tilde{Y}_t) = \mathcal{O}(n)$. Thus it follows from (7.24) that

$$\left|\mathbb{E}[\tilde{Y}_{t+1}] - \left(1 + \frac{R'(s^\star)}{n}\right)\mathbb{E}[\tilde{Y}_t]\right| \leq C_1\left(\frac{\mathbb{E}[\tilde{Y}_t]}{n}\right)^2 + \frac{3C}{n}.$$

Let us denote $b_t = \mathbb{E}(\tilde{Y}_t)$. Then $|b_{t+1} - b_t| \leq 2$ since $|\tilde{Y}_{t+1} - \tilde{Y}_t| \leq 2$. Moreover, the above estimate gives that, for all $t \geq 1$,

$$\left(1 - \frac{a_0}{n}\right)b_t - a_1\left(\frac{b_t}{n}\right)^2 - \frac{a_2}{n} \leq b_{t+1} \leq \left(1 - \frac{a_0}{n}\right)b_t + a_1\left(\frac{b_t}{n}\right)^2 + \frac{a_2}{n},$$

with $|b_0| \leq \delta n$ and $a_0 = -R''(s^\star) > 0$ and $a_1 = C_1 = \sup_{s\in(s^\star-2\delta,s^\star+2\delta)}|R''(s)|/2$.

A straightforward analysis using recursion (which is quite similar to the proof of estimates for a process $(Z_t)_{t\geq 0}$ in Ding et al. (2009a, Sections 4.3.2 and 4.4.1); see Can et al. (2019, Appendix A.1) of the extended version for the full proof) gives that if $b_0 = \delta n$, then

$$b_{T_n^-(\gamma)} \geq \sqrt{n}e^{\gamma a_0/2}, \tag{7.25}$$

and

$$\sqrt{n}e^{-2\gamma a_0} \leq b_{T_n^+(\gamma)} \leq \sqrt{n}. \tag{7.26}$$

By Lemma 7.3,

$$\mathrm{Var}(\tilde{Y}_{T_n^-(\gamma)}), \mathrm{Var}(\tilde{Y}_{T_n^+(\gamma)}) = \mathcal{O}(n). \tag{7.27}$$

Using Chebyshev's inequality, $(7.25) - (7.27)$,

$$\mathbb{P}\left(\tilde{Y}_{T_n^-(\gamma)} \le \alpha\sqrt{n}\right) \le \mathbb{P}\left(\tilde{Y}_{T_n^-(\gamma)} - b_{T_n^-(\gamma)} \le (\alpha - e^{\gamma a_0/2})\sqrt{n}\right)$$

$$\le \frac{\operatorname{Var}(\tilde{Y}_{T_n^-(\gamma)})}{n(\alpha - e^{\gamma a_0/2})^2} = \mathcal{O}(e^{-\gamma a_0}),$$

and

$$\mathbb{P}\left(\tilde{Y}_{T_n^+(\gamma)} \ge \ell\sqrt{n}\right) \le \mathbb{P}\left(\tilde{Y}_{T_n^+(\gamma)} - b_{T_n^+(\gamma)} \ge (\ell - 1)\sqrt{n}\right) \le \frac{\operatorname{Var}(\tilde{Y}_{T_n^+(\gamma)})}{n(\ell - 1)^2} = \mathcal{O}(\ell^{-2}).$$

The first inequality implies Lemma 7.4(i), while the second proves Lemma 7.4(ii) by taking $\ell \to \infty$. □

7.2. The left side of the critical window: Proof of (7.1).

Lemma 7.5 (Concentration of the number of plus spins). *We have*

$$\lim_{\alpha \to \infty} \lim_{n \to \infty} \mu_n\left(\sigma \colon \big||\sigma_+| - ns^\star\big| > \alpha\sqrt{n}\right) = 0.$$

Proof: For any k, let

$$\mu_n^k = \mu_n(\sigma \colon |\sigma_+| = k).$$

Then, by the definition of μ_n,

$$\frac{\mu_n^k}{\mu_n^\ell} = \frac{\binom{n}{k}\exp(nF_n(k))}{\binom{n}{\ell}\exp(nF_n(\ell))} \asymp \exp\left(n[G(k/n) - G(\ell/n)]\right). \tag{7.28}$$

Let $\delta > 0$ be any positive constant such that $(s^\star - 2\delta, s + 2\delta) \subset (0, 1)$. Since $s^\star$ is the unique global maximizer of G, there exists $\varepsilon > 0$, such that $G(k/n) - G(s^\star) \le -\varepsilon$ for all k satisfying $|k - ns^\star| \ge \delta n$. Therefore, by (7.28),

$$\mu_n(\sigma \colon \big||\sigma_+| - ns^\star\big| \ge \delta n) \lesssim n\exp(-\varepsilon n). \tag{7.29}$$

For the case $|k - ns^\star| \le \delta n$, we use $G'(s^\star) = 0$ and a Taylor expansion around $s^\star$ to obtain that

$$G(k/n) - G(s^\star) \le -\frac{a}{2}\left(\frac{k}{n} - s^\star\right)^2,$$

where

$$-a = \inf_{s \colon |s - s^\star| \le \delta} G''(s) < 0.$$

Here the fact that $G''(s^\star) < 0$ follows from condition (C3), which implies that $G''(s) < 0$ in a neighborhood of $s^\star$. Therefore, as $\alpha \to \infty$,

$$\mu_n(\sigma \colon \alpha\sqrt{n} \le \big||\sigma_+| - ns^\star\big| \le \delta n) \lesssim \sum_{\alpha\sqrt{n} \le \ell \le \delta n} \exp(-an\ell^2/2)$$

$$\lesssim \int_\alpha^\infty \exp(-at^2/2)dt \to 0,$$

since $a > 0$. Combining this estimate with (7.29) we get the desired result. □

We are now in the position to prove (7.1):

Proof of (7.1). For any $\alpha > 0$, let

$$A_\alpha = \{\sigma \colon |\sigma_+| \le ns^\star + \alpha\sqrt{n}\}.$$

By definition of the total variation distance, and the fact that μ_n is the stationary distribution,

$$d_n(T_n^-(\gamma)) = \sup_{A \subset \Omega_n} \sup_{\sigma \in \Omega_n} |\mu_n(A) - \mathbb{P}_\sigma(\xi_{T_n^-(\gamma)} \in A)| \qquad (7.30)$$

$$\geq \mathbb{P}_1(\xi_{T_n^-(\gamma)} \in A_\alpha) - \mu_n(A_\alpha).$$

Therefore, for all $\alpha > 0$,

$$\lim_{\gamma \to \infty} \liminf_{n \to \infty} d_n(T_n^-(\gamma)) \geq \lim_{\gamma \to \infty} \liminf_{n \to \infty} \mathbb{P}_1(\xi_{T_n^-(\gamma)} \in A_\alpha) - \lim_{n \to \infty} \mu_n(A_\alpha) \qquad (7.31)$$

$$\geq \lim_{\gamma \to \infty} \liminf_{n \to \infty} \mathbb{P}_1(X_{T_n^-(\gamma)} \leq ns^\star + \alpha\sqrt{n}) - \lim_{n \to \infty} \mu_n(A_\alpha),$$

Hence, by taking $\alpha \to \infty$ and using Lemma 7.5 as well as Lemma 7.4(i), we obtain that

$$\lim_{\gamma \to \infty} \liminf_{n \to \infty} d_n(T_n^-(\gamma)) \geq \lim_{\alpha \to \infty} \lim_{\gamma \to \infty} \liminf_{n \to \infty} \mathbb{P}_1(X_{T_n^-(\gamma)} \leq ns^\star + \alpha\sqrt{n}) = 1, \qquad (7.32)$$

which proves (7.1). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

7.3. *The right side of the critical window: Proof of* (7.2). Recall Ω_n^δ from (7.9). Let $(\xi_t)_{t \geq 0}$ and $(\bar\xi_t)_{t \geq 0}$ be two Glauber dynamics, and let $(X_t)_{t \geq 0}$ and $(\bar X_t)_{t \geq 0}$ be the corresponding projection chains, i.e.,

$$X_t = |\{i \colon \xi_t(i) = 1\}|, \qquad \bar X_t = |\{i \colon \bar\xi_t(i) = 1\}|. \qquad (7.33)$$

For $\sigma_0 \in \Omega_n^\delta$ and $\sigma \in \Omega_n$, we define

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

$$\Theta = \{\sigma \colon \min\{U(\sigma), V(\sigma), U(\sigma_0) - U(\sigma), V(\sigma_0) - V(\sigma)\} \geq \delta n/8\}.$$

We also define

$$D(t) = |U(\xi_t) - U(\bar\xi_t)|, \qquad H_1(t) = \{\inf\{s \colon X_s = \bar X_s\} \leq t\},$$

$$H_2(t_1, t_2) = \bigcap_{t=t_1}^{t_2} \{\xi_t \in \Theta, \bar\xi_t \in \Theta\}.$$

We will crucially rely on the following result from Ding et al. (2009a):

Proposition 7.6 (General mixing time bound, Ding et al., 2009a, Theorem 5.1). *There exists a positive constant c, such that for any possible couplings of $(\xi_t)_{t \geq 0}$ and $(\bar\xi_t)_{t \geq 0}$, for any $r_1 < r_2$ and $\alpha > 0$ and all large n,*

$$\max_{\sigma_0 \in \Omega_n^\delta} \|\mathbb{P}_{\sigma_0}(\xi_{r_2} \in \cdot) - \mu(\cdot)\|_{\mathrm{TV}} \leq \max_{\sigma_0, \sigma \in \Omega_n^\delta} \left[\mathbb{P}_{\sigma_0, \sigma}\left(D(r_1) > \alpha\sqrt{n}\right) + \mathbb{P}_{\sigma_0, \sigma}(H_1(r_1)^c) \right.$$

$$\left. + \mathbb{P}_{\sigma_0, \sigma}(H_2(r_1, r_2)^c) + c\alpha \sqrt{\frac{n}{r_2 - r_1}} \right].$$

We next investigate the different terms appearing on the right hand side in Proposition 7.6, where we choose $r_1 = r_1(\gamma) = T_n^+(\gamma) - Cn = T_n^+(\gamma - C)$, $r_2 = r_2(\gamma) = T_n^+(2\gamma) - Cn = T_n^+(2\gamma - C)$, with C as in Lemma 7.1 and $\alpha = \gamma^{1/4}$ for some γ large enough:

Lemma 7.7 (Unlikely that projection chains remain uncoupled for a long time). *We have*

$$\lim_{\gamma \to \infty} \limsup_{n \to \infty} \max_{\sigma, \sigma'} \mathbb{P}_{\sigma, \sigma'}(H_1(T_n^+(\gamma))^c) = 0.$$

As a consequence,

$$\lim_{\gamma \to \infty} \limsup_{n \to \infty} \max_{\sigma, \sigma'} \mathbb{P}_{\sigma, \sigma'}(H_1(r_1(\gamma))^c) = 0.$$

Proof: By Lemma 7.4(ii)

$$\lim_{\ell\to\infty}\lim_{\gamma\to\infty}\limsup_{n\to\infty}\max_{\sigma\in\Omega}\mathbb{P}_\sigma(|X_{T_n^+(\gamma)} - ns^\star| \geq \ell\sqrt{n}) = 0.$$

Hence, it suffices to show that

$$\lim_{\ell\to\infty}\lim_{\gamma\to\infty}\limsup_{n\to\infty}\max_{\sigma,\sigma'\in\Omega^\ell}\mathbb{P}_{\sigma,\sigma'}(\inf\{t: X_t = \bar{X}_t\} \geq \gamma n) = 0, \tag{7.34}$$

where

$$\Omega^\ell = \{\sigma\in\Omega_n\colon |\sigma_+ - ns^\star| \leq \ell\sqrt{n}\}.$$

By monotonicity, it is sufficient to prove (7.34) for the cases that $X_0 = ns^\star + \ell\sqrt{n}$ and $\bar{X}_0 = ns^\star - \ell\sqrt{n}$.

By (7.6),

$$\max\left\{|\mathbb{E}[X_{t+1} - X_t \mid X_t] - R(X_t/n)|, |\mathbb{E}[\bar{X}_{t+1} - \bar{X}_t \mid \bar{X}_t] - R(\bar{X}_t/n)|\right\} \leq \frac{C}{n}, \tag{7.35}$$

where C is a universal constant. Moreover, if $ns^\star - \delta n \leq \bar{X}_t \leq X_t \leq ns^\star + \delta n$, then

$$R(X_t/n) - R(\bar{X}_t/n) \leq -\frac{1}{n}\left[\inf_{|s-s^\star|\leq\delta}|R'(s)|(X_t - \bar{X}_t)\right] \tag{7.36}$$

since $R'(s) < 0$ for all $|s - s^\star| \leq \delta$. Define

$$K = \frac{2C}{\inf_{|s-s^\star|\leq\delta}|R'(s)|} \in (0,\infty), \tag{7.37}$$

and

$$\begin{aligned}\tau_M &= \inf\{t: |X_t - \bar{X}_t| \leq 2K \text{ or } |X_t - ns^\star| \geq \delta n \text{ or } |\bar{X}_t - ns^\star| \geq \delta n\},\\ M_t &= X_t - \bar{X}_t, \qquad \tilde{M}_t = M_t \mathbb{1}_{\{\tau_M\geq t\}}.\end{aligned}$$

By (7.36), $\tilde{M}_t$ is a supermartingale. Indeed, let $\mathcal{F}_t = \sigma(X_s, \bar{X}_s\colon s\leq t)$ and observe that if $\tilde{M}_t = 0$ then $\tilde{M}_{t+1} = 0$, so that

$$\begin{aligned}\mathbb{E}[\tilde{M}_{t+1} - \tilde{M}_t \mid \mathcal{F}_t] &\leq \mathbb{E}[(\tilde{M}_{t+1} - \tilde{M}_t)\mathbb{1}_{\{\tilde{M}_t>0\}} \mid \mathcal{F}_t]\\ &= \mathbb{E}[X_{t+1} - X_t \mid \mathcal{F}_t]\mathbb{1}_{\{\tilde{M}_t>0\}} - \mathbb{E}[\bar{X}_{t+1} - \bar{X}_t \mid \mathcal{F}_t]\mathbb{1}_{\{\tilde{M}_t>0\}}\\ &\leq \left[R(X_t/n) - R(\bar{X}_t/n)\right]\mathbb{1}_{\{\tilde{M}_t>0\}} + 2C/n\\ &\leq -\frac{1}{n}\left[\inf_{|s-s^\star|\leq\delta}|R'(s)|(2K) - 2C\right] \leq 0,\end{aligned}$$

where in the third line, we have used (7.35), while in the last line, we have used (7.36) and (7.37) with a notice that if $\tilde{M}_t > 0$, then the condition $ns^\star - \delta n \leq \bar{X}_t \leq X_t \leq ns^\star + \delta n$ needed to apply (7.36) holds. We claim that, on the event $\{\tau_M > t\}$,

$$\text{Var}(\tilde{M}_{t+1} \mid \mathcal{F}_t) \geq \Delta, \tag{7.38}$$

for some universal constant $\Delta > 0$. Indeed, when $\tau_M > t$,

$$\mathbb{P}(\tilde{M}_{t+1} - \tilde{M}_t = 0 \mid \mathcal{F}_t), \mathbb{P}(\tilde{M}_{t+1} - \tilde{M}_t = 1 \mid \mathcal{F}_t) \geq c,$$

for some positive constant $c > 0$. Thus $\text{Var}(\tilde{M}_{t+1} \mid \mathcal{F}_t) \geq c$. Moreover, as we have shown above, $\tilde{M}_t$ is a supermartingale with increment bounded by 2. Therefore, using Ding et al. (2009b, Lemma 3.5) we obtain that if $M_0 = 2\ell\sqrt{n}$ then

$$\mathbb{P}(\tau_M \geq \ell^3 n) = \mathcal{O}(M_0/\sqrt{\ell^3 n}) = \mathcal{O}(1/\sqrt{\ell}).$$

Hence, by definition of τ_M and Lemma 7.2,

$$
\begin{aligned}
\mathbb{P}(\inf\{t\colon M_t \leq 2K\} \geq \ell^3 n) &\leq \mathbb{P}(\tau_M \geq \ell^3 n) + \mathbb{P}(\inf\{t\colon |X_t - ns^\star| \geq \delta n\} \leq \ell^3 n) \\
&\quad + \mathbb{P}(\inf\{t\colon |\bar{X}_t - ns^\star| \geq \delta n\} \leq \ell^3 n) \\
&= \mathbb{P}(\tau_M \geq \ell^3 n) + \mathcal{O}(n^{-3/2}) = \mathcal{O}(1/\sqrt{\ell}).
\end{aligned}
$$

where we let $\ell^3 = o(\log n)$ and apply Lemma 7.2 in the last line.

Thanks to the above estimate, now we can assume that $M_0 = 2K$. We observe that

$$
\mathbb{P}(M_{t+1} - M_t \in \{-1, -2\}) = 1 - \mathbb{P}(M_{t+1} - M_t \in \{0, 1, 2\}) \geq c, \tag{7.39}
$$

for some universal constant c. Let $\tau_0 = \tilde{\tau}_0 = 0$ and define for $i \geq 1$,

$$
\begin{aligned}
\tau_i &= \inf\{t > \tilde{\tau}_{i-1} : M_t \geq 4K \text{ or } M_t \leq 0\}, \\
\tilde{\tau}_i &= \inf\{t > \tau_i : M_t \leq 2K\}.
\end{aligned}
$$

We observe that $((\tau_{i+1} - \tilde{\tau}_i, \tilde{\tau}_i - \tau_i))_{i \geq 0}$ are independent random variables. Moreover, using (7.38) and the fact that $|M_{t+1} - M_t| \leq 2$, we can show that

$$
\sup_{i \geq 0} \left(\mathbb{E}[(\tau_{i+1} - \tilde{\tau}_i)^2] + \mathbb{E}[(\tilde{\tau}_i - \tau_i)^2] \right) \leq C,
$$

for some universal constant C. Therefore, using Chebyshev's inequality,

$$
\mathbb{P}(\tau_\ell \geq 4C\ell) = \mathcal{O}(1/\ell). \tag{7.40}
$$

By (7.39),

$$
\mathbb{P}(M_{\tau_i} \leq 0 \mid \tilde{\tau}_{i-1}) \geq c^{2K}.
$$

Hence,

$$
\mathbb{P}(\exists i \leq \ell \colon M_{\tau_i} \leq 0) \geq 1 - (1 - c^{2K})^\ell = 1 - \mathcal{O}(1/\ell).
$$

Combining this estimate with (7.40) we obtain that

$$
\mathbb{P}_{2K}(\inf\{t : M_t \leq 0\} \geq 4C\ell) = \mathcal{O}(1/\ell).
$$

This completes the proof of (7.34). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

Lemma 7.8 (($(\xi_t)_{t \geq 0}$ and $(\bar{\xi}_t)_{t \geq 0}$ are about equally far from σ_0). *Consider two Glauber dynamics* $(\xi_t)_{t \geq 0}$ *and* $(\bar{\xi}_t)_{t \geq 0}$ *started at some* $\sigma_0 \in \Omega_n^\delta$ *and* $\sigma \in \Omega_n$, *respectively. There exists a positive constant* C, *such that for all* $\alpha > 0$,

$$
\lim_{\gamma \to \infty} \limsup_{n \to \infty} \mathbb{P}_{\sigma_0, \sigma}(D(T_n^+(\gamma)) \geq \alpha\sqrt{n}) \leq C/\alpha.
$$

Consequently,

$$
\lim_{\gamma \to \infty} \limsup_{n \to \infty} \mathbb{P}_{\sigma_0, \sigma}(D(r_1(\gamma)) \geq \alpha\sqrt{n}) \leq C/\alpha.
$$

Lemma 7.9 (Both chains are in Θ for all large times). *Consider two Glauber dynamics* $(\xi_t)_{t \geq 0}$ *and* $(\bar{\xi}_t)_{t \geq 0}$ *started at some* $\sigma_0 \in \Omega_n^\delta$ *and* $\sigma \in \Omega_n$ *respectively. Then for* $\gamma_2 > \gamma_1$,

$$
\mathbb{P}_{\sigma_0, \sigma}(H_2(T_n^+(\gamma_1), T_n^+(\gamma_2))^c) \leq \frac{C(\gamma_2 - \gamma_1) + C_1}{n}, \tag{7.41}
$$

where C *is a universal constant and* C_1 *is a constant depending on* γ_1. *Consequently,*

$$
\lim_{\gamma \to \infty} \limsup_{n \to \infty} \mathbb{P}_{\sigma_0, \sigma}(H_2(r_1(\gamma), r_2(\gamma))^c) = 0. \tag{7.42}
$$

The proofs of Lemmas 7.8 and 7.9 are essentially the same as those of Ding et al. (2009a, Lemmas 5.5 and 5.6), so we omit them here. Their full proofs will be given in Can et al. (2019, Appendix A.2) of the extended version.

Proof of (7.2). By Lemma 7.1,

$$\limsup_{n \to \infty} \max_{\sigma \in \Omega_n} \mathbb{P}_\sigma(\xi_{Cn} \in \Omega_n^\delta) = 1. \tag{7.43}$$

Combining (7.43) and Lemmas 7.7, 7.8, 7.9 and Proposition 7.6 with $r_1 = T_n^+(\gamma - Cn)$, $r_2 = T_n^+(2\gamma - Cn)$ and $\alpha = \gamma^{1/4}$, we arrive at

$$\limsup_{n \to \infty} \max_{\sigma \in \Omega} \|\mathbb{P}_\sigma(\xi_{T_n^+(2\gamma)} \in \cdot) - \mu_n(\cdot)\|_{\mathrm{TV}} \leq C\gamma^{-1/4},$$

for some $C > 0$. This proves (7.2) by taking $\gamma \to \infty$. $\qquad\square$

Acknowledgements

References

Angel, O., van der Hofstad, R., and Holmgren, C. Limit laws for self-loops and multiple edges in the configuration model. *Ann. Inst. Henri Poincaré Probab. Stat.*, **55** (3), 1509–1530 (2019). MR4010943.

Barrera, J., Bertoncini, O., and Fernández, R. Abrupt convergence and escape behavior for birth and death chains. *J. Stat. Phys.*, **137** (4), 595–623 (2009). MR2565098.

Basak, A. and Dembo, A. Ferromagnetic Ising measures on large locally tree-like graphs. *Ann. Probab.*, **45** (2), 780–823 (2017). MR3630287.

Bollobás, B. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, **1** (4), 311–316 (1980). MR595929.

Bollobás, B. The isoperimetric number of random regular graphs. *European J. Combin.*, **9** (3), 241–244 (1988). MR947025.

Bollobás, B. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition (2001). ISBN 0-521-80920-7; 0-521-79722-5. MR1864966.

Bovier, A. *Statistical mechanics of disordered systems. A mathematical perspective*, volume 18 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge (2006). ISBN 978-0-521-84991-3; 0-521-84991-8. MR2252929.

Bovier, A. and den Hollander, F. *Metastability. A potential-theoretic approach*, volume 351 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham (2015). ISBN 978-3-319-24775-5; 978-3-319-24777-9. MR3445787.

Bovier, A., Marello, S., and Pulvirenti, E. Metastability for the dilute Curie-Weiss model with glauber dynamics. *ArXiv Mathematics e-prints* (2019). arXiv: 1912.10699.

Can, V. H. Critical behavior of the annealed Ising model on random regular graphs. *J. Stat. Phys.*, **169** (3), 480–503 (2017). MR3711608.

Can, V. H. Annealed limit theorems for the Ising model on random regular graphs. *Ann. Appl. Probab.*, **29** (3), 1398–1445 (2019). MR3914548.

Can, V. H., Giardinà, C., Giberti, C., and van der Hofstad, R. Annealed Ising model on configuration models (2021+). To appear in *Ann. Inst. H. Poincaré Probab. Statist.*

Can, V. H., van der Hofstad, R., and Kumagai, T. Glauber dynamics for Ising models on random regular graphs: cut-off and metastability. *ArXiv Mathematics e-prints* (2019). arXiv: 1912.07798.

Chen, G.-Y. and Saloff-Coste, L. On the mixing time and spectral gap for birth and death chains. *ALEA Lat. Am. J. Probab. Math. Stat.*, **10** (1), 293–321 (2013). MR3083928.

Dembo, A. and Montanari, A. Gibbs measures and phase transitions on sparse random graphs. *Braz. J. Probab. Stat.*, **24** (2), 137–211 (2010a). MR2643563.

Dembo, A. and Montanari, A. Ising models on locally tree-like graphs. *Ann. Appl. Probab.*, **20** (2), 565–592 (2010b). MR2650042.

Dembo, A., Montanari, A., Sly, A., and Sun, N. The replica symmetric solution for Potts models on d-regular graphs. *Comm. Math. Phys.*, **327** (2), 551–575 (2014). MR3183409.

Ding, J., Lubetzky, E., and Peres, Y. Censored Glauber dynamics for the mean field Ising model. *J. Stat. Phys.*, **137** (3), 407–458 (2009a). MR2564285.

Ding, J., Lubetzky, E., and Peres, Y. The mixing time evolution of Glauber dynamics for the mean-field Ising model. *Comm. Math. Phys.*, **289** (2), 725–764 (2009b). MR2506768.

Dommers, S. Metastability of the Ising model on random regular graphs at zero temperature. *Probab. Theory Related Fields*, **167** (1-2), 305–324 (2017). MR3602847.

Dommers, S., den Hollander, F., Jovanovski, O., and Nardi, F. R. Metastability for Glauber dynamics on random graphs. *Ann. Appl. Probab.*, **27** (4), 2130–2158 (2017). MR3693522.

Dommers, S., Giardinà, C., Giberti, C., van der Hofstad, R., and Prioriello, M. L. Ising critical behavior of inhomogeneous Curie-Weiss models and annealed random graphs. *Comm. Math. Phys.*, **348** (1), 221–263 (2016). MR3551266.

Dommers, S., Giardinà, C., and van der Hofstad, R. Ising models on power-law random graphs. *J. Stat. Phys.*, **141** (4), 638–660 (2010). MR2733399.

Dommers, S., Giardinà, C., and van der Hofstad, R. Ising critical exponents on random trees and graphs. *Comm. Math. Phys.*, **328** (1), 355–395 (2014). MR3196989.

Dorogovtsev, S. N., Goltsev, A. V., and Mendes, J. F. F. Critical phenomena in complex networks. *Rev. Mod. Phys.*, **80**, 1275–1335 (2008). DOI: 10.1103/RevModPhys.80.1275.

Ellis, R. S. *Entropy, large deviations, and statistical mechanics.* Classics in Mathematics. Springer-Verlag, Berlin (2006). ISBN 978-3-540-29059-9; 3-540-29059-1. MR2189669.

Giardinà, C., Giberti, C., van der Hofstad, R., and Prioriello, M. L. Quenched central limit theorems for the Ising model on random graphs. *J. Stat. Phys.*, **160** (6), 1623–1657 (2015). MR3382761.

Giardinà, C., Giberti, C., van der Hofstad, R., and Prioriello, M. L. Annealed central limit theorems for the Ising model on random graphs. *ALEA Lat. Am. J. Probab. Math. Stat.*, **13** (1), 121–161 (2016). MR3476210.

van der Hofstad, R. *Random graphs and complex networks. Vol. 1.* Cambridge Series in Statistical and Probabilistic Mathematics, [43]. Cambridge University Press, Cambridge (2017). ISBN 978-1-107-17287-6. MR3617364.

den Hollander, F. and Jovanovski, O. Glauber dynamics on the Erdős–Rényi random graph. *In 'In and Out of Equilibrium' 3, Celebrating Vladas Sidoravicius, Progress in Probability*, **77**, 519–589 (2021). DOI: 10.1007/978-3-030-60754-8_24.

Janson, S. The probability that a random multigraph is simple. *Combin. Probab. Comput.*, **18** (1-2), 205–225 (2009). MR2497380.

Janson, S. The probability that a random multigraph is simple. II. *J. Appl. Probab.*, **51A** (Celebrating 50 Years of The Applied Probability Trust), 123–137 (2014). MR3317354.

Levin, D. A., Luczak, M. J., and Peres, Y. Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability. *Probab. Theory Related Fields*, **146** (1-2), 223–265 (2010).

MR2550363.

Levin, D. A. and Peres, Y. *Markov chains and mixing times.* American Mathematical Society, Providence, RI (2017). ISBN 978-1-4704-2962-1. MR3726904.

Lubetzky, E. and Sly, A. Universality of cutoff for the Ising model. *Ann. Probab.*, **45** (6A), 3664–3696 (2017). MR3729612.

Martinelli, F., Sinclair, A., and Weitz, D. Glauber dynamics on trees: boundary conditions and mixing time. *Comm. Math. Phys.*, **250** (2), 301–334 (2004). MR2094519.

Molloy, M. and Reed, B. A critical point for random graphs with a given degree sequence. In *Proceedings of the Sixth International Seminar on Random Graphs and Probabilistic Methods in Combinatorics and Computer Science, "Random Graphs '93" (Poznań, 1993)*, volume 6, pp. 161–179 (1995). MR1370952.

Montanari, A., Mossel, E., and Sly, A. The weak limit of Ising models on locally tree-like graphs. *Probab. Theory Related Fields*, **152** (1-2), 31–51 (2012). MR2875752.

Mossel, E. and Sly, A. Exact thresholds for Ising-Gibbs samplers on general graphs. *Ann. Probab.*, **41** (1), 294–328 (2013). MR3059200.

Niss, M. History of the Lenz-Ising model 1920–1950: from ferromagnetic to cooperative phenomena. *Arch. Hist. Exact Sci.*, **59** (3), 267–318 (2005). MR2124728.

Niss, M. History of the Lenz-Ising model 1950–1965: from irrelevance to relevance. *Arch. Hist. Exact Sci.*, **63** (3), 243–287 (2009). MR2501198.

Weitz, D. Counting independent sets up to the tree threshold. In *STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pp. 140–149. ACM, New York (2006). MR2277139.