AN $\widetilde{O}(n^3)$-TIME SAMPLER FOR ZERO-FIELD FERROMAGNETIC ISING MODELS

WEIMING FENG, HENG GUO, AND YIYAO ZHANG

ABSTRACT. We give an approximate sampler for ferromagnetic Ising models with no field on arbitrary graphs that runs in time $\widetilde{O}(m+n)+\widetilde{O}_\beta\left(n^3\log^3\frac{1}{\varepsilon}\right)$, where n and m are the numbers of vertices and edges, respectively, and ε is the approximation error. Our algorithm utilises the machinery behind the classic cut sparsifier by Benczúr and Karger (2015) and the transport flow idea by Chen, Feng, Ju, Miao, Yin, and Zhang (2025). The main proofs were found using GPT-5.6 Sol Pro.

1. INTRODUCTION

The Ising model is perhaps the most studied and most well-known statistical physics model. It features both a simple description and a rich mathematical structure. The main computational task for the Ising model is to sample from its Gibbs distribution, which we define next.

Let $G := (V, E)$ be a connected undirected (multi)graph. Let $\beta > 0$ be a parameter describing the interaction strength. For any configuration $\sigma \in \{0,1\}^V$, define the number of monochromatic edges as $\mathrm{mono}(\sigma) := |\{uv \in E : \sigma_u = \sigma_v\}|$, where parallel edges are counted with multiplicities. Let $Z_\beta^{\mathrm{Ising}} := \sum_{\tau\in\{0,1\}^V} \beta^{\mathrm{mono}(\tau)}$ be the so-called partition function. The *Gibbs distribution* of the zero-field Ising model on G with interaction parameter β is defined by

$$\forall \sigma \in \{0,1\}^V, \quad \mu_{\beta,G}^{\mathrm{Ising}}(\sigma) := \frac{\beta^{\mathrm{mono}(\sigma)}}{Z_\beta^{\mathrm{Ising}}}.$$

When $\beta > 1$, the system is called ferromagnetic, and anti-ferromagnetic otherwise.

The computational complexity of sampling from $\mu_{\beta,G}^{\mathrm{Ising}}$ has been a classic topic in theoretical computer science. The pioneering work of Jerrum and Sinclair [JS93] showed that a polynomial-time approximate sampler[1] exists for ferromagnetic Ising models, even with consistent external fields,[2] whereas the problem becomes **NP**-hard for anti-ferromagnetic Ising models.

For anti-ferromagnetic Ising models, we have by now fully understood the computational phase transition regarding the parameter β and the maximum degree [LLY13, SST14, SS14, GŠV16, CJM+26]. Below the critical threshold, the running time of the sampler has been improved to near-linear [CLV23, AJK+22, CFYZ22]. However, on the ferromagnetic side, algorithmic progress is relatively sparse. The only running time improvement is obtained for the non-zero field case, where Chen and Zhang [CZ23] introduced a near-linear time sampling algorithm. For the zero-field case, prior to this work, the fastest sampler was still the Jerrum and Sinclair algorithm [JS93], which runs in time $O(m^2n^4(m+\log\varepsilon^{-1}))$, where n denotes the number of vertices, m the number of edges, and ε the error. This run-time can be as high as $O(n^{10})$ for dense graphs. In this paper, we bring it down to near-cubic. For two probability

(Weiming Feng) SCHOOL OF COMPUTING AND DATA SCIENCE, THE UNIVERSITY OF HONG KONG, HONG KONG, CHINA. EMAIL: WFENG@HKU.HK. WEIMING FENG ACKNOWLEDGES THE SUPPORT OF ECS GRANT 27202725 FROM HONG KONG RGC.

(Heng Guo) SCHOOL OF INFORMATICS, UNIVERSITY OF EDINBURGH, EDINBURGH, UNITED KINGDOM. EMAIL: HGUO@INF.ED.AC.UK.

(Yiyao Zhang) STATE KEY LABORATORY FOR NOVEL SOFTWARE TECHNOLOGY, NEW CORNERSTONE SCIENCE LABORATORY, NANJING UNIVERSITY, NANJING, CHINA. EMAIL: ZHANGYIYAO@SMAIL.NJU.EDU.CN.

[1]The original paper [JS93] does not give a sampler explicitly, but it is not hard to obtain one, as shown by Randall and Wilson [RW99].

[2]Roughly speaking, external fields are vertex weights. In this paper we focus on the zero-field case, and thus do not introduce them formally.

distributions μ and ν on the same finite state space Ω, define their *total variation distance* by $d_{\mathrm{TV}}(\mu,\nu) := \frac{1}{2}\sum_{x\in\Omega}|\mu(x)-\nu(x)|$.

Theorem 1. *For any constant $\beta > 1$, there is an algorithm such that, given $G = (V, E)$ and an error bound $\varepsilon \in (0, 1)$, it outputs a random configuration with law ν satisfying $d_{\mathrm{TV}}(\nu, \mu^{\mathrm{Ising}}_{\beta,G}) \le \varepsilon$ in time $\widetilde{O}(m + n) + \widetilde{O}_\beta\left(n^3 \log^3 \frac{1}{\varepsilon}\right)$, where $\widetilde{O}$ hides a* polylog(n) *factor, $n := |V|$, and $m := |E|$. In particular, if G is simple, the total running time of the algorithm is $\widetilde{O}_\beta\left(n^3 \log^3 \frac{1}{\varepsilon}\right)$.*

1.1. **Algorithm overview.** The two main ingredients in the algorithm of Theorem 1 are sparsification and an improved mixing time analysis for the random-cluster model. Although the authors provided some high-level guidance, the detailed proofs of both ingredients were found by GPT-5.6 Sol Pro. Subsequently, the authors simplified, streamlined, and wrote all of the proofs. The authors also take full responsibility for the correctness of the paper. For consistency with standard mathematical exposition, the words "we" and "our" are used throughout the paper, including when presenting arguments originating from the model.

The first ingredient of the algorithm is a sparsification argument. It states that the Gibbs distribution can be approximated in total variation distance by another Gibbs distribution on a sparsified graph.

Theorem 2 (Ising sparsification)**.** *For every constant $\beta > 1$, there is an algorithm that, given $G := (V, E)$ and an error bound $\eta \in (0, 1)$, explicitly constructs a sparsified multigraph $G' := (V', E')$ and an implicit representation of a lifting map* Lift $: \{0, 1\}^{V'} \to \{0, 1\}^V$ *in time $O(m \log^2 n + n)$, such that $|V'| \le |V| = n$ and $|E'| = O_\beta(n \log(n/\eta))$, where edges are counted with multiplicity. Moreover:*

- *If μ' is the zero-field Ising Gibbs distribution on G' with interaction parameter β, then $d_{\mathrm{TV}}(\mathrm{Lift}(\mu'), \mu^{\mathrm{Ising}}_{\beta,G}) \le \eta$, where $\mathrm{Lift}(\mu')$ denotes the law of $\mathrm{Lift}(X)$ for $X \sim \mu'$.*
- *Given any $\sigma \in \{0, 1\}^{V'}$, the configuration $\mathrm{Lift}(\sigma)$ can be computed in $O(n)$ time.*

The main technique we use is the machinery behind the classic cut sparsifier by Benczúr and Karger [BK15] and Karger's cut counting bounds [Kar93, Kar99]. Indeed, the Ising partition function can be viewed as a generating function for cuts, and it perhaps is not a surprise that these tools are useful here. However, this argument appears to be new in the context of partition function approximation, and is the first sparsification argument in this context as far as we are aware. In fact, we notice that the same technique works for all-terminal reliability [GJ19] as well, and this improves its FPRAS's running time from $\widetilde{O}(mn/\varepsilon^2)$ [CGZZ24] to $\widetilde{O}(m + n^2/\varepsilon^2)$. Details are given in Section A.

The intuition here is simple. For any subgraph with sufficiently high edge connectivity, with high probability, all vertices have the same spin, and thus can be contracted while approximating the original Gibbs distribution. This crucially relies on the fact that β is a constant. If $\beta = 1 + o(1)$, such as $\beta = 1 + c/n$, then the argument is no longer valid. The machinery in [BK15] allows us to correctly and efficiently identify what subgraphs need to be contracted. With the sparsification in hand, we only need to consider Ising models on graphs with $m = O_\beta(n \log(n/\varepsilon))$ edges.

On the sparsified graph, we use Glauber dynamics for the random-cluster model, introduced by Fortuin and Kasteleyn [FK72], which can be coupled with the ferromagnetic Ising model. This model is defined as follows. Again, suppose the graph is $G = (V, E)$. For $S \subseteq E$, let $\kappa(S)$ denote the number of connected components of (V, S), including isolated vertices. Given $p \in (0, 1)$, the random-cluster model on G defines the probability distribution

$$\forall S \subseteq E, \quad \mu^{\mathrm{RC}}_{p,G}(S) := \frac{p^{|S|}(1-p)^{m-|S|}2^{\kappa(S)}}{Z_{\mathrm{RC}}},$$

where Z_{RC} is the corresponding partition function. Edges in S are called open, and edges in $E \setminus S$ are called closed. Parallel edges are treated as distinct edges. Edwards and Sokal [ES88] showed that the ferromagnetic Ising model and the random-cluster model can be tightly coupled.

Proposition 3 (Edwards–Sokal coupling [ES88])**.** *Let $p := 1 - 1/\beta$. Sample $S \sim \mu^{\mathrm{RC}}_{p,G}$ and assign an independent uniformly random spin in $\{0,1\}$ to each connected component of (V,S), giving every vertex the spin of its component. The resulting configuration has distribution $\mu^{\mathrm{Ising}}_{\beta,G}$. Conversely, sample $\sigma \sim \mu^{\mathrm{Ising}}_{\beta,G}$ and, conditional on σ, include each monochromatic edge in S independently with probability p and close every bichromatic edge. Then $S \sim \mu^{\mathrm{RC}}_{p,G}$.*

Thus, it suffices to sample from $\mu^{\mathrm{RC}}_{p,G}$. For this, we use the single-edge heat-bath *Glauber dynamics*. Start from an arbitrary subset $X_0 \subseteq E$. At each step $t \geq 0$, choose an edge $e \in E$ uniformly at random, let u and v be its endpoints, and set $X_t^- := X_t \setminus \{e\}$. The Glauber dynamics resamples the membership of e conditional on X_t^-. Formally, the next state is $X_{t+1} := X_t^- \cup \{e\}$ with probability

$$p' := \begin{cases} p, & \text{if } u \text{ and } v \text{ lie in the same connected component of } (V, X_t^-), \\ \dfrac{p}{p + 2(1-p)} & \text{otherwise.} \end{cases}$$

Otherwise, with probability $1-p'$, set $X_{t+1} := X_t^-$. Parallel edges are treated as distinct edges. It is well known that the Glauber dynamics converges to the stationary distribution $\mu^{\mathrm{RC}}_{p,G}$.

Let P denote the transition kernel of the Glauber dynamics above. For $\varepsilon \in (0,1)$, define its mixing time by

$$T_{\mathrm{mix}}(P,\varepsilon) := \min\left\{ t \geq 0 : \max_{X_0 \subseteq E} d_{\mathrm{TV}}\big(P^t(X_0,\cdot), \mu^{\mathrm{RC}}_{p,G}\big) \leq \varepsilon \right\}.$$

We prove the following mixing time bound for the Glauber dynamics.

Theorem 4 (Mixing time of the Glauber dynamics)**.** *Let $p \in (0,1)$ be a constant. For every connected multigraph $G := (V,E)$ with $n := |V|$ and $m := |E| > 0$ and every $\varepsilon \in (0,1)$,*

$$T_{\mathrm{mix}}(P,\varepsilon) = O_p\left(nm^2 \left(\log(n+m) + \log \frac{1}{\varepsilon} \right) \right).$$

Previously, the best (and only) mixing-time upper bound for this dynamics was the result of Guo and Jerrum [GJ18], which is an $O(m^2 n^4 (m + \log \varepsilon^{-1}))$ bound. The improvement of Theorem 4 comes in two ways. The bound in [GJ18] is obtained via a lifting argument from the canonical path argument for the so-called subgraph-world model [JS93], and the key is to bound the congestion of the lifted multicommodity flow [Sin92]. There is in fact some slack in the congestion analysis of [GJ18]. The flow going through a particular edge can be restricted to flows from a much smaller subset of subgraph-world configurations, instead of all states. This improves the bound by a factor of roughly $O(n^3)$. Then, to further sharpen the mixing time, we use the transport flow framework of Chen, Feng, Ju, Miao, Yin, and Zhang [CFJ+25]. Roughly speaking, this argument "localises" the multicommodity-flow construction, allowing us to bound the log-Sobolev constant of the Markov chain rather than only its spectral gap, as in [GJ18]. This yields a further $O(m)$-factor improvement in the mixing time. The key step to apply this argument is new local couplings for the random-cluster model.

We expect that the mixing time bound in Theorem 4 can be improved, but we also conjecture that the truth is not near-linear. We discuss this further in Section 6.

Each transition of the Glauber dynamics on the random-cluster model can be implemented in $O(\log^2 n)$ amortized time using a fully dynamic connectivity data structure [WN13, CZ23]. Therefore, one can use Glauber dynamics to sample a random configuration from $\mu^{\mathrm{RC}}_{p,G}$ in time $\widetilde{O}_p\left(nm^2 \log \frac{1}{\varepsilon}\right)$. The algorithm in Theorem 1 is obtained by combining the Ising sparsification theorem in Theorem 2, the Edwards–Sokal coupling in Proposition 3, and the mixing-time bound for the Glauber dynamics in Theorem 4. The full algorithm is given in Section 5.

2. A Near-Linear-Time Ising Sparsification

In this section, we prove Theorem 2. Self-loops contribute the same factor β to every spin configuration and therefore do not affect the Ising Gibbs distribution. We may delete all self-loops and assume throughout this section that G is loopless. Our main tools are the edge-strength estimates underlying the Benczúr–Karger cut sparsifier [BK15] and Karger's bounds on the number of cuts [Kar93, Kar99]. We first recall the relevant definitions.

Definition 5 (k-edge-connected)**.** *Let $G := (V, E)$ be a graph. For a subset $S \subseteq V$, let $\delta_G(S)$ denote the multiset of edges with exactly one endpoint in S; this multiset is called a* cut *of G. The cut is* nontrivial *if S is a nonempty proper subset of V.*

We say that G is k-edge-connected if all nontrivial cuts of G have size at least k, i.e., $|\delta_G(S)| \geq k$ for every nonempty proper subset $S \subset V$.

In particular, a singleton vertex is considered k-edge-connected for any $k > 0$.

Definition 6 (Edge strength)**.** *Fix a graph $G := (V, E)$. For an edge $e = uv \in E$, its strength s_e is the largest integer k for which u and v are contained in a common induced subgraph that is k-edge-connected.*

In [BK15], Benczúr and Karger introduced the notion of *tight strength* $\widetilde{s}_e$ for every edge $e \in E$, which is a computable lower bound on s_e. The following is the result for the unweighted case, which is sufficient for our purpose.

Proposition 7 ([BK15, Theorem 6.5])**.** *Given an unweighted multigraph $G := (V, E)$ with $n := |V|$ and $m := |E|$, one can compute in $O(m \log^2 n)$ time positive values $(\widetilde{s}_e)_{e \in E}$ such that $\widetilde{s}_e \leq s_e$ for every $e \in E$ and $\sum_{e \in E} 1/\widetilde{s}_e \leq c_0 n$, where $c_0 > 0$ is an absolute constant.*

We also use the following cut-counting bound of Karger [Kar93, Kar99].

Proposition 8 ([Kar99, Corollary A.7])**.** *Let $G := (V, E)$ be a multigraph on $n \geq 2$ vertices such that every nontrivial cut of G has size at least K. For every real $\alpha \geq 1$, the number of distinct cuts of G with size at most αK is at most $n^{2\alpha}$.*

We are now ready to prove the Ising sparsification theorem.

Proof of Theorem 2. Apply Proposition 7 to G to obtain $(\widetilde{s}_e)_{e \in E}$ satisfying

$$\widetilde{s}_e \leq s_e \quad \text{for every } e \in E, \qquad \sum_{e \in E} \frac{1}{\widetilde{s}_e} \leq c_0 n, \tag{1}$$

and set

$$K := \left\lceil \frac{\log(2n^5/\eta)}{\log((\beta + 1)/2)} \right\rceil \tag{2}$$

and let $E_{\geq K} := \{e \in E : \widetilde{s}_e \geq K\}$. Let $C_1, \ldots, C_N$ be the connected components of $(V, E_{\geq K})$. Construct G' by contracting each component C_i into a distinct single vertex, deleting all resulting self-loops, and retaining every edge between distinct components with its multiplicity. Formally, $V' := \{C_1, \ldots, C_N\}$. For any two distinct $C_i, C_j \in V'$, the number of parallel edges between C_i and C_j equals the number of edges in E with endpoints in C_i and C_j.

Since all edges in $E_{\geq K}$ are contracted, every edge e of G' satisfies $\widetilde{s}_e < K$ in G. Hence

$$|E'| = \sum_{e \in E'} 1 < \sum_{e \in E'} \frac{K}{\widetilde{s}_e} \leq K \cdot c_0 n = O_\beta\big(n \log(n/\eta)\big),$$

where the last inequality follows from (1), and $|V'| = N \leq n$.

We can lift an Ising configuration on G' back to a configuration on G by giving each contracted component the same spin as in the corresponding vertex in G'. We next bound the total variation distance between the Ising distribution on G and the distribution lifted from G'. For $U \subseteq V$, let $G[U]$ denote the subgraph of G induced by U, and call U a K-strong vertex set if $G[U]$ is K-edge-connected. Let $B_1, \ldots, B_L$ be the inclusion-maximal K-strong vertex sets, and call the induced graphs $G[B_1], \ldots, G[B_L]$ the maximal K-strong components.

If two K-strong vertex sets B_i and B_j intersect, then their union is again K-strong. To see this, any nonempty proper subset $S \subset B_i \cup B_j$ restricts to a nonempty proper subset of at least one of the two sets and hence the cut has size at least K. Consequently, the maximal sets $B_1, \ldots, B_L$ are pairwise disjoint. Every vertex belongs to one of these maximal sets because a singleton is K-strong. Now consider an edge $e = uv \in E_{\geq K}$. Since $\widetilde{s}_e \geq K$ and $\widetilde{s}_e \leq s_e$, the definition of s_e implies that u and v belong to a common K-strong vertex set, and hence to the same maximal set B_ℓ. Following an $E_{\geq K}$-path shows that every connected component C_i of $(V, E_{\geq K})$ is contained in a unique B_ℓ.

Consider the random-cluster model with parameter $p := 1 - 1/\beta$ on G through the Edwards–Sokal coupling, Proposition 3. For any edge $e \in E$, the conditional probability that e is open given the states of all other edges is at least

$$x := \min\left\{p, \frac{p}{p + 2(1-p)}\right\} = \frac{p}{p + 2(1-p)} = \frac{\beta - 1}{\beta + 1}.$$

Thus the random-cluster configuration stochastically dominates independent Bernoulli-x bond percolation, where every edge is open independently with probability x.

Fix a maximal K-strong component $H = G[B_\ell] = (V_H, E_H)$ with at least two vertices, and let $n_H := |V_H| \leq n$. For $j \geq 1$, let $\mathcal{C}_j$ be the collection of cuts of H whose sizes belong to the interval $[jK, (j+1)K)$. By Proposition 8,

$$|\mathcal{C}_j| \leq n_H^{2(j+1)} \leq n^{2(j+1)}.$$

The percolated subgraph of H can be disconnected only if all the edges of some nontrivial cut of H are closed. For a fixed cut of H with size at least jK, under Bernoulli-x bond percolation, the probability that all edges in the cut are closed is at most $(1-x)^{jK}$. By a union bound,

$$\Pr(H \text{ is disconnected}) \leq \sum_{j \geq 1} |\mathcal{C}_j|(1-x)^{jK} \leq \sum_{j \geq 1} n^{2(j+1)}(1-x)^{jK} = \frac{n^4(1-x)^K}{1 - n^2(1-x)^K}. \tag{3}$$

Moreover, $1-x = 2/(\beta+1)$, and the choice of K in (2) implies $(1-x)^K = (2/(\beta+1))^K \leq \eta/(2n^5)$. In particular, $n^2(1-x)^K \leq \eta/(2n^3) \leq 1/2$. Substituting this estimate into (3) yields

$$\Pr(H \text{ is disconnected}) \leq \frac{n^4(1-x)^K}{1 - n^2(1-x)^K} \leq 2n^4(1-x)^K \leq \frac{\eta}{n},$$

where singleton components are trivially connected. Since connectivity is an increasing event, stochastic domination gives the same upper bound under the random-cluster measure. There are at most n maximal K-strong components, so another union bound shows that all of them are connected in the random-cluster configuration with probability at least $1 - \eta$. Under the Edwards–Sokal coupling, each open component receives the same spin. Therefore, for the Ising model on graph G, it holds that

$$\mu_{\beta,G}^{\text{Ising}}(\mathcal{E}) \geq 1 - \eta,$$

where $\mathcal{E}$ is the event that all vertices in each block C_i are assigned the same spin.

Store for each $v \in V$ the index of the block C_i containing it. This is an implicit representation of the lifting map: for $\sigma' \in \{0,1\}^{V'}$, set $\mathrm{Lift}(\sigma')_v := \sigma'_{C_i}$ whenever $v \in C_i$.

Let $m_{\text{in}} := m - |E'|$. For every $\sigma' \in \{0,1\}^{V'}$, $\mathrm{mono}(\mathrm{Lift}(\sigma')) = m_{\text{in}} + \mathrm{mono}_{G'}(\sigma')$. The first term is independent of σ', and hence $\mathrm{Lift}(\mu') = \mu_{\beta,G}^{\text{Ising}}(\cdot \mid \mathcal{E})$. Consequently,

$$d_{\mathrm{TV}}(\mathrm{Lift}(\mu'), \mu_{\beta,G}^{\text{Ising}}) = 1 - \mu_{\beta,G}^{\text{Ising}}(\mathcal{E}) \leq \eta.$$

The components of $(V, E_{\geq K})$, the graph G', and the block-index representation of Lift require only $O(m+n)$ additional time after the edge strength estimation. This proves the theorem. $\square$

3. Preliminaries for Markov chain analysis

Our main algorithm is a Markov chain, so we review some basics of Markov chains in this section. Let Ω be a finite state space. A Markov chain $(X_t)_{t\geq 0}$ on Ω is specified by a transition matrix P, where $P(x,y)$ is the probability of moving from x to y in one step. Let $P^t(x,y)$ denote the probability of moving from x to y in t steps. A probability distribution π on Ω is stationary if $\pi P = \pi$. The chain is irreducible if, for every $x, y \in \Omega$, there exists $t \geq 0$ such that $P^t(x,y) > 0$. The period of x is the greatest common divisor of the set $\{t \geq 1 : P^t(x,x) > 0\}$, and the chain is aperiodic if every state has period one. Every finite irreducible and aperiodic Markov chain has a unique stationary distribution π, and $P^t(x,\cdot)$ converges to π from every initial state $x \in \Omega$. The chain is reversible with respect to π if, for every pair $x, y \in \Omega$, the detailed balance condition $\pi(x)P(x,y) = \pi(y)P(y,x)$ holds. Reversibility implies that π is stationary.

3.1. **Log-Sobolev inequality.** The so-called log-Sobolev constant provides a very sharp bound on the mixing time. Suppose that P is reversible with respect to π. For a function $f : \Omega \to \mathbb{R}$, let $\mathbb{E}_\pi[f] := \sum_{x\in\Omega} \pi(x)f(x)$. When $\mathbb{E}_\pi[f^2] > 0$, define the entropy of f^2 and the Dirichlet form of f as follows:

$$\mathrm{Ent}_\pi(f^2) := \sum_{x\in\Omega} \pi(x) f(x)^2 \log\left(\frac{f(x)^2}{\mathbb{E}_\pi[f^2]}\right),$$

$$\mathcal{E}_P(f,f) := \frac{1}{2}\sum_{x,y\in\Omega} \pi(x)P(x,y)\big(f(x)-f(y)\big)^2.$$

Here $0\log 0$ is interpreted as zero. The log-Sobolev inequality with constant $\rho > 0$ is

$$\mathrm{Ent}_\pi(f^2) \leq \frac{1}{\rho}\cdot \mathcal{E}_P(f,f) \tag{4}$$

for every $f : \Omega \to \mathbb{R}$. The largest admissible ρ is denoted by $\rho(P)$ and is called the log-Sobolev constant of P. For an irreducible, aperiodic, and reversible Markov chain, the log-Sobolev constant controls the mixing time. Let $\pi_{\min} := \min_{x\in\Omega} \pi(x)$. We mainly consider Glauber dynamics in this paper, which has no negative eigenvalues [DGU14]. For these chains, a standard log-Sobolev mixing bound is given by [DS96]:

$$T_{\mathrm{mix}}(P,\varepsilon) \leq O\left(\frac{1}{\rho(P)}\left(\log\log\frac{1}{\pi_{\min}} + \log\frac{1}{\varepsilon}\right)\right). \tag{5}$$

3.2. **Transport flow of Glauber dynamics.** We consider the random cluster dynamics over edge configurations. Let $G = (V,E)$ be a graph. Let π be a strictly positive probability distribution on $\{0,1\}^E$ and $m := |E|$. Let P be the single-site heat-bath Glauber dynamics for π. Thus, at each step, P chooses a uniformly random edge $e \in E$ and resamples its value from the conditional distribution given the values on $E \setminus \{e\}$. For the random-cluster model, we identify an edge set $S \subseteq E$ with its indicator vector $\mathbf{1}_S \in \{0,1\}^E$. We will use the transport flow introduced in [CFJ+25] to bound the log-Sobolev constant for Glauber dynamics.

Definition 9 (Transport flow in P)**.** *A path in the transition graph of the Markov chain P is a sequence $\gamma := (x_0, x_1, \ldots, x_\ell)$ such that $P(x_i, x_{i+1}) > 0$ for every $i < \ell$. Let $s(\gamma) := x_0$, $t(\gamma) := x_\ell$, and $\ell(\gamma) := \ell$. For two distributions ν_0 and ν_1 on $\{0,1\}^E$, a transport flow from ν_0 to ν_1 is a distribution Γ over paths in the transition graph of P such that, for $\gamma \sim \Gamma$, the starting point $s(\gamma)$ has distribution ν_0 and the endpoint $t(\gamma)$ has distribution ν_1.*

The transport-flow framework requires us to control not only π, but also every conditional distribution obtained from π by fixing some coordinates. For a nonempty set $\Lambda \subseteq E$ and an assignment $\tau \in \{0,1\}^{E\setminus\Lambda}$, let π^τ denote the distribution of $X \sim \pi$ conditional on $X_{E\setminus\Lambda} = \tau$. Thus, τ pins the coordinates outside Λ, while the coordinates in Λ remain free. We regard π^τ as a distribution on $\{0,1\}^E$ supported on $\{x \in \{0,1\}^E : x_{E\setminus\Lambda} = \tau\}$. Let P^τ be the single-site heat-bath Glauber dynamics for π^τ: at each step, it chooses a uniformly random $e \in \Lambda$ and resamples X_e conditional on $X_{\Lambda\setminus\{e\}}$. We use π_e^τ to denote the marginal distribution on e conditional on

τ. We regard P^τ as a Markov chain on $\{0,1\}^E$: it only updates the coordinates in Λ and keeps the coordinates outside Λ fixed as τ. We use $\pi^{\tau\wedge e\leftarrow a}$ to denote π conditioned on τ together with the event $X_e = a$. Equivalently, $\tau \wedge e \leftarrow a$ is the assignment obtained by extending τ and fixing e to a. For heat-bath Glauber dynamics, the following result was observed in [CFJ+25].

Proposition 10. *Let $\Lambda \subseteq E$ with $k := |\Lambda| \geq 1$, let $\tau \in \{0,1\}^{E\setminus\Lambda}$. For any $f : \{0,1\}^E \to \mathbb{R}$,*

$$\frac{1}{k}\sum_{e\in\Lambda}\sum_{a\in\{0,1\}} \pi_e^\tau(a)\,\mathcal{E}_{P^{\tau\wedge e\leftarrow a}}(f,f) \leq \mathcal{E}_{P^\tau}(f,f).$$

Proof. For $k = 1$, the left-hand side is zero, so the inequality is immediate. Assume that $k \geq 2$. By the definition of the Dirichlet form,

$$\begin{aligned}
&\frac{1}{k}\sum_{e\in\Lambda}\sum_{a\in\{0,1\}} \pi_e^\tau(a)\mathcal{E}_{P^{\tau\wedge e\leftarrow a}}(f,f)\\
&\quad= \frac{1}{2k}\sum_{e\in\Lambda}\sum_{a\in\{0,1\}}\sum_{x,y\in\{0,1\}^E} \pi_e^\tau(a)\pi^{\tau\wedge e\leftarrow a}(x)P^{\tau\wedge e\leftarrow a}(x,y)\big(f(x)-f(y)\big)^2\\
&\quad= \frac{1}{2(k-1)}\sum_{e\in\Lambda}\sum_{\substack{x,y\in\{0,1\}^E\\ x_e=y_e}} \pi^\tau(x)P^\tau(x,y)\big(f(x)-f(y)\big)^2 \qquad (6)\\
&\quad= \frac{1}{2}\sum_{x,y\in\{0,1\}^E} \pi^\tau(x)P^\tau(x,y)\big(f(x)-f(y)\big)^2 = \mathcal{E}_{P^\tau}(f,f). \qquad (7)
\end{aligned}$$

We verify (6). We only need to consider terms for which $x_{E\setminus\Lambda} = y_{E\setminus\Lambda} = \tau$, $x_e = y_e = a$, and x, y differ at exactly one coordinate in $\Lambda \setminus \{e\}$; every other term is zero. For such x, y,

$$\pi_e^\tau(a)\pi^{\tau\wedge e\leftarrow a}(x) = \pi^\tau(x), \qquad\qquad P^{\tau\wedge e\leftarrow a}(x,y) = \frac{k}{k-1}P^\tau(x,y),$$

where the second identity follows from the heat-bath definitions. This verifies (6).

To verify (7), it suffices to consider x, y that differ at exactly one coordinate in Λ, say e'. This pair is counted $k-1$ times in the sum in (6), once for each $e \in \Lambda \setminus \{e'\}$. □

Together with this proposition, [CFJ+25, Theorems 6 and 11] give the following result, which we use to prove the log-Sobolev inequality for the Glauber dynamics. For a transition $x \to y$ of P, let $(x \to y) \in \gamma$, where $\gamma := (x_0, x_1, \ldots, x_\ell)$, mean that $x = x_i$ and $y = x_{i+1}$ for some $i < \ell$.

Theorem 11 (Log-Sobolev inequality via transport flows [CFJ+25])**.** *Suppose that there is a constant $\phi \in (0, 1/2)$ such that $\pi_e^\tau(a) \geq \phi$ for every nonempty Λ, τ, $e \in \Lambda$, and $a \in \{0,1\}$. Let $c_1, \ldots, c_m > 0$. Suppose that, whenever $|\Lambda| = k$, for every pinning $\tau \in \{0,1\}^{E\setminus\Lambda}$ and every $e \in \Lambda$ there is a transport flow Γ_e^τ from $\pi^{\tau\wedge e\leftarrow 0}$ to $\pi^{\tau\wedge e\leftarrow 1}$ in the transition graph of P^τ, and that for every transition $x \to y$ of P^τ,*

$$\sum_{e\in\Lambda} \pi_e^\tau(0)\pi_e^\tau(1)\mathbb{E}_{\gamma\sim\Gamma_e^\tau}\left[\ell(\gamma)\mathbf{1}\{(x\to y)\in\gamma\}\right] \leq c_k\pi^\tau(x)P^\tau(x,y).$$

Then the log-Sobolev constant of the Glauber dynamics P satisfies

$$\rho(P) \geq \frac{1-2\phi}{8\log((1-\phi)/\phi)}\left(\sum_{k=1}^m \frac{c_k}{k}\right)^{-1}.$$

4. Mixing time of random-cluster Glauber dynamics

In this section, we prove Theorem 4. We use the transport-flow framework in Theorem 11 to prove the log-Sobolev inequality for the Glauber dynamics. Although Theorem 11 requires us to consider all conditional distributions, the random-cluster model is closed under conditioning: fixing an edge to be closed amounts to deleting it, while fixing an edge to be open amounts to contracting it. Thus, after any pinning, the conditional distribution on the remaining edges is a random-cluster model with the same parameter p on a graph obtained from G by edge deletions and contractions. Any self-loops created by contraction are independent Bernoulli-p edges. Although algorithmically it is easy to handle them separately, we are analysing the mixing time, so we keep them in the analysis below. Therefore, it suffices to establish the required bound for the random-cluster Glauber dynamics P on an arbitrary multigraph G. The same argument then applies to every conditional distribution.

The main technical lemma in this section is the following.

Lemma 12. *Fix a constant $p \in (0,1)$. Let $G := (V,E)$ be a multigraph with $n := |V|$ and $m := |E| > 0$. Let $\pi := \mu^{\mathrm{RC}}_{p,G}$ and let P be its single-edge heat-bath Glauber dynamics. For every $e \in E$, there is a transport flow Γ_e from $\pi^{e\leftarrow 0}$ to $\pi^{e\leftarrow 1}$ in the transition graph of P. Every path γ in the support of these flows satisfies $\ell(\gamma) \le n$, and, for every transition $\omega \to \omega'$ of P,*

$$\sum_{e\in E} \pi_e(0)\pi_e(1)\mathbb{E}_{\gamma\sim\Gamma_e}\left[\ell(\gamma)\mathbf{1}\{(\omega\to\omega')\in\gamma\}\right] \le C_p n m^2 \pi(\omega)P(\omega,\omega'),$$

where

$$C_p := \max\left\{2, \frac{2p}{(2-p)(1-p)}\right\} \le \frac{2}{1-p}.$$

The construction of Γ_e has two steps. First, we construct a local coupling $\mathcal{C}_e$ of the two conditional distributions $\pi^{e\leftarrow 0}$ and $\pi^{e\leftarrow 1}$. Second, for every pair of configurations $(X,X') \sim \mathcal{C}_e$, we construct a path between them in the transition graph of P. The local coupling is the key part of the construction.

Before describing the local coupling, let us first construct Γ_e for a self-loop e. In this case, the state of e is an independent Bernoulli-p variable under π. Couple all other edges identically to be the set $S \subseteq E\setminus\{e\}$ and use the one-step path $(S, e=0) \to (S, e=1)$, which corresponds to the Glauber update on the self-loop edge e. The congestion through a transition $\omega \to \omega'$ updating e is bounded as follows. We may assume $\omega := (S, e=0)$ and $\omega' := (S, e=1)$. Note that a flow for another self-loop e' only updates e'. Furthermore, our construction for a non-loop edge e'' below guarantees that $\Gamma_{e''}$ uses only Glauber updates on non-loop edges. Hence, among all the flows, only Γ_e uses $\omega \to \omega'$. Since this path has length one and the heat-bath transition opens e with probability p/m, while $\Pr_{\gamma\sim\Gamma_e}(\gamma = (\omega,\omega')) = \pi(\omega)/(1-p)$, we have

$$\sum_{f\in E} \pi_f(0)\pi_f(1)\mathbb{E}_{\gamma\sim\Gamma_f}\left[\ell(\gamma)\mathbf{1}\{(\omega\to\omega')\in\gamma\}\right]$$
$$= \pi_e(0)\pi_e(1)\Pr_{\gamma\sim\Gamma_e}(\gamma = (\omega,\omega')) = (1-p)p\cdot\frac{\pi(\omega)}{1-p} = p\pi(\omega) = m\pi(\omega)P(\omega,\omega').$$

Thus, self-loop transitions satisfy the bound in Lemma 12.

We now fix a non-loop edge $e = uv$. At a high level, the coupling $\mathcal{C}_e$ is constructed so that either X and X' differ only at e, or their other disagreements lie on a simple u-to-v path in the graph $G - e$. Starting from X, we change the status of the disagreeing edges one at a time in their order along the u-to-v path, and change the status of e last. The resulting sequence is a path from X to X' in the transition graph of P, and its length is at most n. The distribution of this random path is Γ_e. The detailed construction is in Section 4.1.

4.1. The transport flow.

We first decompose the two conditional random-cluster distributions. Let $G_{-e} := G - e$. The distribution $\pi^{e\leftarrow 0}_{E-e}$ is the random-cluster distribution with parameter p on G_{-e}. Let q be the probability that u and v are connected in $\pi^{e\leftarrow 0}_{E-e}$, i.e.,

$$q := \Pr_{R\sim\pi^{e\leftarrow 0}_{E-e}}\big(u \text{ and } v \text{ are connected in } (V,R)\big). \tag{8}$$

We have the following decomposition lemma.

Lemma 13. *If $q > 0$, then*

$$\pi^{e\leftarrow 1}_{E-e} = \frac{1}{1+q}\pi^{e\leftarrow 0}_{E-e} + \frac{q}{1+q}\pi^{e\leftarrow 0}_{E-e}(\,\cdot\mid u \text{ and } v \text{ are connected in } (V,R)). \tag{9}$$

If $q = 0$, then $\pi^{e\leftarrow 1}_{E-e} = \pi^{e\leftarrow 0}_{E-e}$.

Proof. Note that the number of edges in G_{-e} is $m-1$. For $R \subseteq E(G_{-e})$, define its random-cluster weight on G_{-e} by

$$w(R) := p^{|R|}(1-p)^{m-1-|R|}2^{\kappa(R)}.$$

Thus, $\pi_{E-e}^{e\leftarrow 0}(R)$ is proportional to $w(R)$. Now condition on e being open and ignore the common factor p in the weight contributed by e. Adding e leaves the number of connected components unchanged when u and v are connected in (V,R), and decreases it by one otherwise. Hence, the weight of R under $\pi_{E-e}^{e\leftarrow 1}$ is

$$p^{|R|}(1-p)^{m-1-|R|}2^{\kappa(R\cup\{e\})} = \begin{cases} w(R), & u \text{ and } v \text{ are connected in } (V,R), \\ \dfrac{1}{2}w(R), & \text{otherwise.} \end{cases}$$

Define the normalising factor of $\pi_{E-e}^{e\leftarrow 1}$ by

$$Z_1 := \sum_{R\subseteq E(G_{-e})} p^{|R|}(1-p)^{m-1-|R|}2^{\kappa(R\cup\{e\})}.$$

Similarly, the normalising factor of $\pi_{E-e}^{e\leftarrow 0}$ is $Z_0 := \sum_{R\subseteq E(G_{-e})} w(R)$. Since the event of u and v being connected in (V,R) has probability q under $\pi_{E-e}^{e\leftarrow 0}$, we have

$$Z_1 = \left(q + \frac{1-q}{2}\right) \sum_{R\subseteq E(G_{-e})} w(R) = \frac{1+q}{2} \sum_{R\subseteq E(G_{-e})} w(R) = \frac{1+q}{2} Z_0. \tag{10}$$

It follows that

$$\pi_{E-e}^{e\leftarrow 1}(R) = \begin{cases} \dfrac{2}{1+q}\pi_{E-e}^{e\leftarrow 0}(R), & u \text{ and } v \text{ are connected in } (V,R), \\ \dfrac{1}{1+q}\pi_{E-e}^{e\leftarrow 0}(R), & \text{otherwise.} \end{cases}$$

When $q>0$, this is exactly the claimed mixture. Indeed, let $\mathcal{F}$ be the event that u and v are connected in (V,R). If $R\in\mathcal{F}$, then $\pi_{E-e}^{e\leftarrow 0}(R \mid \mathcal{F}) = \pi_{E-e}^{e\leftarrow 0}(R)/q$, so the right-hand side of (9) evaluated at R is $\frac{1}{1+q}\pi_{E-e}^{e\leftarrow 0}(R) + \frac{q}{1+q}\pi_{E-e}^{e\leftarrow 0}(R\mid\mathcal{F}) = \frac{2}{1+q}\pi_{E-e}^{e\leftarrow 0}(R)$. On the other hand, if $R\notin\mathcal{F}$, then $\pi_{E-e}^{e\leftarrow 0}(R\mid\mathcal{F}) = 0$, so the same right-hand side is $\frac{1}{1+q}\pi_{E-e}^{e\leftarrow 0}(R)$. These are precisely the two cases above. If $q=0$, then u and v are disconnected in every configuration of $\pi_{E-e}^{e\leftarrow 0}$, and hence the equality holds directly. This concludes the proof of Lemma 13. □

By Lemma 13, when $q>0$, the only nontrivial ingredient needed to construct $\mathcal{C}_e$ is a coupling

$$R_0 \sim \pi_{E-e}^{e\leftarrow 0}, \qquad R_1 \sim \pi_{E-e}^{e\leftarrow 0}(\cdot \mid u \text{ and } v \text{ are connected in } (V,R)). \tag{11}$$

We construct this coupling below and give the complete definition of $\mathcal{C}_e$ at the end of the subsection. For this purpose, we introduce another probability distribution, called the subgraph-world model [JS93]. Let $G := (V,E)$ be a multigraph. For any $S\subseteq E$, we define $\mathrm{Odd}(S) := \{v\in V : \deg_S(v) \text{ is odd}\}$, where degrees count edge multiplicity. We call S an even subgraph if $\mathrm{Odd}(S)=\emptyset$. The subgraph-world model with parameter $p/2\in(0,1)$ is the probability distribution on even subgraphs of G given by

$$\mu_{p/2}^{\mathrm{SW}}(S) \propto \left(\frac{p}{2}\right)^{|S|}\left(1-\frac{p}{2}\right)^{|E|-|S|}, \qquad S\subseteq E \text{ with } \mathrm{Odd}(S)=\emptyset.$$

The subgraph-world model and the random-cluster model are related via the following coupling by Grimmett and Janson [GJ09].

Proposition 14. *Let $p\in(0,1)$ and sample $S\sim\mu_{p/2}^{\mathrm{SW}}$. We form $R\subseteq E$ by including every edge of S and including each edge of $E\setminus S$ independently with probability $p/(2-p)$. Then $R\sim\mu_{p,G}^{\mathrm{RC}}$.*

For the first distribution $\pi_{E-e}^{e\leftarrow 0}$ in Lemma 13, we may apply Proposition 14 to G_{-e} to get a lifting from the corresponding subgraph-world model. For the second distribution, namely $\pi_{E-e}^{e\leftarrow 0}$ conditioned on u and v being connected, we introduce the uv-modified subgraph-world model on G_{-e}. This distribution is supported on edge sets $S \subseteq E(G_{-e})$ satisfying $\mathrm{Odd}(S) = \{u, v\}$ and is given by

$$\mu_{p/2}^{uv\text{-mSW}}(S) \propto \left(\frac{p}{2}\right)^{|S|}\left(1-\frac{p}{2}\right)^{|E(G_{-e})|-|S|}, \qquad \mathrm{Odd}(S) = \{u, v\}.$$

Note that the distribution is well-defined if and only if u and v are connected in G_{-e}. Otherwise, there is no such subgraph S with $\mathrm{Odd}(S) = \{u, v\}$. To see this, every connected component has an even number of odd-degree vertices, and it is impossible to have a subgraph with $\mathrm{Odd}(S) = \{u, v\}$ if u and v are disconnected. We have the following lemma.

Lemma 15. *Suppose that $\mu_{p/2}^{uv\text{-mSW}}$ is well-defined. Sample $S \sim \mu_{p/2}^{uv\text{-mSW}}$ and form $R \subseteq E(G_{-e})$ by including every edge of S and including each edge of $E(G_{-e})\setminus S$ independently with probability $p/(2-p)$. Then R has the random-cluster distribution with parameter p on G_{-e} conditioned on the event that u and v are connected.*

Proof. Note that $|E(G_{-e})| = m-1$. For a fixed $R \subseteq E(G_{-e})$, its probability is proportional to

$$\sum_{\substack{S\subseteq R\\ \mathrm{Odd}(S)=\{u,v\}}} \left(\frac{p}{2}\right)^{|S|}\left(1-\frac{p}{2}\right)^{m-1-|S|}\left(\frac{p}{2-p}\right)^{|R|-|S|}\left(1-\frac{p}{2-p}\right)^{m-1-|R|}$$
$$= |\{S \subseteq R : \mathrm{Odd}(S) = \{u, v\}\}| \left(\frac{p}{2}\right)^{|R|}(1-p)^{m-1-|R|}.$$

There is no $S \subseteq R$ with $\mathrm{Odd}(S) = \{u, v\}$ unless u and v are connected in (V, R). Suppose that is the case and fix a u-to-v path $L \subseteq R$. For any even subgraph $T \subseteq R$, consider $T \triangle L$. Here, $\triangle$ denotes symmetric difference. Then $\mathrm{Odd}(T \triangle L) = \{u, v\}$. Conversely, if $\mathrm{Odd}(S) = \{u, v\}$, then $S \triangle L$ is even. Thus, $T \mapsto T \triangle L$ is a bijection between the even subgraphs of (V, R) and the sets $S \subseteq R$ satisfying $\mathrm{Odd}(S) = \{u, v\}$. It is classical that a graph with n vertices, $|R|$ edges, and $\kappa(R)$ connected components has exactly $2^{|R|-n+\kappa(R)}$ even subgraphs; see, e.g., [GJ09, Proposition 2.3]. Therefore, the probability of R is proportional to

$$\mathbf{1}\{u \text{ and } v \text{ are connected in } (V, R)\} \cdot p^{|R|}(1-p)^{m-1-|R|}2^{\kappa(R)},$$

which is precisely the random-cluster weight on the subgraph G_{-e} restricted to the event that u and v are connected. This completes the proof. ◻

Consider the multigraph $G := (V, E)$ and fix a non-loop edge $e = uv \in E$. Assume that u and v are connected in G_{-e}. Then $\mu_{p/2}^{uv\text{-mSW}}$ is well-defined. Moreover, if a subset of edges $S \subseteq E(G_{-e})$ satisfies $\mathrm{Odd}(S) = \{u, v\}$, then S contains a simple u-to-v path. We use L_S to denote a simple u-to-v path in S chosen according to a fixed ordering. We now couple the ordinary and modified subgraph-world distributions on G_{-e}.

Definition 16 (Subgraph-world switching coupling)**.** *The subgraph-world distribution $\mu_{p/2}^{\mathrm{SW}}$ and the modified subgraph-world distribution $\mu_{p/2}^{uv\text{-mSW}}$ are coupled as follows:*

1. *Independently sample $S_0 \sim \mu_{p/2}^{\mathrm{SW}}$ and $S' \sim \mu_{p/2}^{uv\text{-mSW}}$.*
2. *Since S_0 is an even subgraph and $\mathrm{Odd}(S') = \{u, v\}$, their symmetric difference satisfies $\mathrm{Odd}(S_0 \triangle S') = \{u, v\}$. Thus, $S_0 \triangle S'$ contains a simple u-to-v path $L := L_{S_0\triangle S'}$.*
3. *Set $S_1 := S_0 \triangle L$.*

We call the resulting joint distribution of (S_0, S_1) the subgraph-world switching coupling.

Lemma 17. *The subgraph-world switching coupling in Definition 16 is a valid coupling of the two distributions on G_{-e} and $S_0 \triangle S_1 = L$ is a simple u-to-v path in G_{-e}.*

Proof. The marginal of S_0 and the identity $S_0 \triangle S_1 = L$ follow directly from Definition 16. It remains to verify the marginal of S_1.

Let $A, B \subseteq E(G_{-e})$ satisfy $\mathrm{Odd}(A) = \emptyset$ and $\mathrm{Odd}(B) = \{u, v\}$. Thus, $\mathrm{Odd}(A \triangle B) = \{u, v\}$, and the path $L := L_{A \triangle B}$ is well-defined. Define the switching map Φ by

$$\Phi(A, B) := (A \triangle L, B \triangle L).$$

The symmetric difference of the two output sets satisfies $(A \triangle L) \triangle (B \triangle L) = A \triangle B$. Thus, applying Φ to the output selects the same path L and gives

$$\Phi(A \triangle L, B \triangle L) = (A, B).$$

Say that (A, B) has parity type (X, Y) if $(\mathrm{Odd}(A), \mathrm{Odd}(B)) = (X, Y)$. Hence Φ is a bijection between pairs of parity types $(\emptyset, \{u, v\})$ and $(\{u, v\}, \emptyset)$. Moreover, every edge of $L \subseteq A \triangle B$ belongs to exactly one of A and B, so switching L merely exchanges its membership between the two sets. In particular,

$$|A| + |B| = |A \triangle L| + |B \triangle L|.$$

Thus, the switching map preserves the product of the two subgraph-world weights and sends $\mu_{p/2}^{\mathrm{SW}} \otimes \mu_{p/2}^{uv\text{-mSW}}$ to $\mu_{p/2}^{uv\text{-mSW}} \otimes \mu_{p/2}^{\mathrm{SW}}$. Applying Φ to (S_0, S') with $S_0 \sim \mu_{p/2}^{\mathrm{SW}}$ and $S' \sim \mu_{p/2}^{uv\text{-mSW}}$ gives $(S_1, S' \triangle L)$, and its first output is $S_0 \triangle L = S_1$, proving that $S_1 \sim \mu_{p/2}^{uv\text{-mSW}}$. □

We next construct the coupling $\mathcal{C}_e$ in (11). The subgraph-world switching coupling in Definition 16 produces (S_0, S_1). Set $L := S_0 \triangle S_1$. Independently sample $B_f \sim \mathrm{Bernoulli}(p/(2-p))$ for every $f \in E(G_{-e})$, and, for $i \in \{0, 1\}$, set

$$R_i := S_i \cup \{f \in E(G_{-e}) \setminus S_i : B_f = 1\}. \tag{12}$$

By Proposition 14, $R_0 \sim \pi_{E-e}^{e\leftarrow 0}$, and by Lemma 15, R_1 has distribution $\pi_{E-e}^{e\leftarrow 0}$ conditioned on u and v being connected in (V, R_1). Since the same variables $(B_f)_{f \in E(G_{-e})}$ are used for both completions and S_0 and S_1 agree outside L, we have $R_0 \triangle R_1 \subseteq L$. Recall the definition of q in (8). We now give the complete definition of the local coupling.

Definition 18 (Local coupling $\mathcal{C}_e$)**.** *For the fixed non-loop edge $e = uv$, the coupling $\mathcal{C}_e$ is the joint distribution of two full configurations (X, X') generated by the following construction.*

1. *If $q = 0$, sample $R \sim \pi_{E-e}^{e\leftarrow 0}$ and set $X := R$ and $X' := R \cup \{e\}$.*
2. *Suppose that $q > 0$.*
   - (a) *With probability $1/(1+q)$, sample $R \sim \pi_{E-e}^{e\leftarrow 0}$ and set $X := R$ and $X' := R \cup \{e\}$.*
   - (b) *With probability $q/(1+q)$, sample (S_0, S_1) according to Definition 16 and construct (R_0, R_1) according to (12). Set $X := R_0$ and $X' := R_1 \cup \{e\}$.*

By (9) and the marginals in (11), $X \sim \pi^{e\leftarrow 0}$ and $X' \sim \pi^{e\leftarrow 1}$, so Definition 18 is a valid coupling. When $q = 0$, or in the first subcase when $q > 0$, $X \triangle X' = \{e\}$. For the second subcase when $q > 0$, it holds that

$$(X \triangle X') \setminus \{e\} = R_0 \triangle R_1 \subseteq S_0 \triangle S_1 = L_{S_0 \triangle S_1}, \tag{13}$$

where $S_0 \triangle S_1$ is a simple u-to-v path by Lemma 17.

Finally, given a random pair (X, X') from $\mathcal{C}_e$, we construct the transport flow Γ_e as follows. In Case (2b) of Definition 18, let $L = L_{S_0 \triangle S_1}$ be the simple path, orient it from u to v, and let

$$\{f_1, \ldots, f_k\} := (X \triangle X') \setminus \{e\},$$

where the edges are listed in the order in which they occur along L. In all other cases set $k := 0$. Starting from $X_0 := X$, define

$$X_i := X_{i-1} \triangle \{f_i\}, \qquad i = 1, \ldots, k,$$

and finally set

$$X_{k+1} := X_k \triangle \{e\} = X'.$$

Let $\gamma := (X_0, X_1, \ldots, X_{k+1})$. Every two consecutive configurations differ at exactly one edge and form a transition of P. The distribution of γ induced by $(X, X') \sim \mathcal{C}_e$ is Γ_e. Its starting point has distribution $\pi^{e\leftarrow 0}$ and its endpoint has distribution $\pi^{e\leftarrow 1}$, so Γ_e is a transport flow between these two distributions. Moreover, $k = 0$ in Case (1) and Case (2a). In Case (2b), $(X \triangle X') \setminus \{e\}$ is contained in a simple u-to-v path in (13). Hence $k \le n-1$ and every path in the support of Γ_e has length at most n.

4.2. **Analysing the congestion.** Now we can show the key result of the section.

Proof of Lemma 12. It remains to prove the congestion bound for the flows for non-loop edges. For a non-loop edge $e = uv$, let $D_e := 2 - p + pq$ where q is the probability that u and v are connected in (V, R) for $R \sim \pi^{e\leftarrow 0}_{E-e}$, as in Lemma 13. Let Z_0 and Z_1 be as in (10). Using $Z_1 = (1+q)Z_0/2$, we have

$$\pi_e(0) = \frac{(1-p)Z_0}{pZ_1 + (1-p)Z_0} = \frac{2(1-p)}{D_e}, \quad \text{and} \quad \pi_e(1) = \frac{pZ_1}{pZ_1 + (1-p)Z_0} = \frac{p(1+q)}{D_e}. \tag{14}$$

For every transition $\omega \to \omega'$, the congestion estimate required in Lemma 12 is

$$\sum_{e\in E} \pi_e(0)\pi_e(1)\mathbb{E}_{\gamma\sim\Gamma_e}\left[\ell(\gamma)\mathbf{1}\{(\omega\to\omega')\in\gamma\}\right] \le C_p n m^2 \pi(\omega)P(\omega,\omega'). \tag{15}$$

The case in which $\omega \to \omega'$ updates a self-loop was handled before Section 4.1. To control a non-loop transition, we first prove the following intermediate estimate separately for every fixed non-loop edge e:

$$\pi_e(0)\pi_e(1)\mathbb{E}_{\gamma\sim\Gamma_e}\left[\ell(\gamma)\mathbf{1}\{(\omega\to\omega')\in\gamma\}\right] \le C'_p n\pi(\omega), \tag{16}$$

where $C'_p := \frac{2p}{2-p}$. To see why this is sufficient, note that a flow caused by a self-loop does not use a non-loop transition and hence it suffices to sum over the at most m non-loop edges e. This gives a total load at most $C'_p nm\pi(\omega)$. Moreover, the transition probability of Glauber dynamics from ω to ω' can be lower bounded as follows:

$$P(\omega,\omega') \ge \frac{1}{m}\min\left\{\frac{p}{2-p}, 1-p\right\}.$$

Hence we have

$$\frac{C'_p}{\min\{p/(2-p), 1-p\}} = \max\left\{2, \frac{2p}{(2-p)(1-p)}\right\} = C_p.$$

Therefore,

$$C'_p nm\pi(\omega) \le C_p nm^2\pi(\omega)P(\omega,\omega'),$$

which is precisely the required bound in (15).

Every path in the support of Γ_e has length at most n. Therefore, to prove the bound in (16), it suffices to prove the following bound:

$$\pi_e(0)\pi_e(1)\Pr_{\gamma\sim\Gamma_e}\left((\omega\to\omega')\in\gamma\right) \le C'_p\pi(\omega). \tag{17}$$

Recall that by our construction of the transport flow, a path sampled from Γ_e consists of initial transitions, each of which flips an edge $e' \ne e$, followed by the final transition that opens e. We bound these two types of transitions separately.

We first introduce the notation and definitions needed to control the initial transitions. For $S \subseteq E(G_{-e})$, define the weight of S in the subgraph-world model as follows:

$$\mathrm{wt}(S) := \left(\frac{p}{2}\right)^{|S|}\left(1-\frac{p}{2}\right)^{m-1-|S|}.$$

Let $u_0, u_1 \in V$ be a pair of distinct vertices. Define

$$Z_{\{u_0,u_1\}} := \sum_{\substack{S\subseteq E(G_{-e})\\ \mathrm{Odd}(S)=\{u_0,u_1\}}} \mathrm{wt}(S), \quad \text{and} \quad Z_\emptyset := \sum_{\substack{S\subseteq E(G_{-e})\\ \mathrm{Odd}(S)=\emptyset}} \mathrm{wt}(S).$$

We first present basic properties of these quantities. For $S, R \subseteq E(G_{-e})$, the transition from a subgraph-world configuration S to a random-cluster configuration R is given by the following completion kernel:

$$K(S,R) := \mathbf{1}\{S \subseteq R\} \left(\frac{p}{2-p}\right)^{|R\setminus S|} \left(1-\frac{p}{2-p}\right)^{m-1-|R|}. \tag{18}$$

This is the "lifting" probability as in Proposition 14 and Lemma 15.

We make the following claim and prove the claim at the end of the proof.

Claim 19. *For every pair of distinct vertices $u_0, u_1 \in V$ and every $R \subseteq E(G_{-e})$, it holds that*

$$\sum_{\substack{S\subseteq E(G_{-e})\\ \mathrm{Odd}(S)=\{u_0,u_1\}}} \frac{wt(S)}{Z_\emptyset} \cdot K(S,R) = \mathbf{1}\left\{u_0 \text{ and } u_1 \text{ are connected in } (V,R)\right\} \cdot \pi_{E-e}^{e\leftarrow 0}(R). \tag{19}$$

Consequently, summing over $R \subseteq E(G_{-e})$ gives

$$\frac{Z_{\{u_0,u_1\}}}{Z_\emptyset} = \Pr_{R\sim\pi_{E-e}^{e\leftarrow 0}}\big(u_0 \text{ and } u_1 \text{ are connected in } (V,R)\big). \tag{20}$$

Since the probability in (20) is at most one, for every pair of distinct vertices $u_0, u_1 \in V$,

$$Z_{\{u_0,u_1\}} \le Z_\emptyset. \tag{21}$$

For $u_0 = u$ and $u_1 = v$, this probability is q as in Lemma 13. Hence,

$$Z_{\{u,v\}} = qZ_\emptyset. \tag{22}$$

We now prove (17), beginning with an initial transition $\omega \to \omega'$. If $q = 0$, the flow for e has no initial transition. Hence, suppose that $q > 0$. An initial transition can occur only in Case (2b) of Definition 18, which is selected with probability $q/(1+q)$. In this case, Definition 16 independently samples S_0 and S'. Although the switching coupling outputs (S_0, S_1), we first analyse the independent samples (S_0, S') used in its construction, since S_1 is determined by these two samples. Note that S_0 and S' are sampled independently from the subgraph world model and uv-modified subgraph world model respectively. The LHS of (17) can be written as the following sum:

$$\begin{aligned}
&\pi_e(0)\pi_e(1) \Pr_{\gamma\sim\Gamma_e}\big((\omega\to\omega')\in\gamma\big)\\
&= \sum_{\substack{S_0\subseteq E(G_{-e}):\ \mathrm{Odd}(S_0)=\emptyset\\ S'\subseteq E(G_{-e}):\ \mathrm{Odd}(S')=\{u,v\}}} \pi_e(0)\pi_e(1)\cdot\frac{q}{1+q}\cdot\frac{\mathrm{wt}(S_0)}{Z_\emptyset}\cdot\frac{\mathrm{wt}(S')}{Z_{\{u,v\}}}\cdot\Pr\big((\omega\to\omega')\in\gamma \mid S_0,S'\big)\\
&= \frac{2p(1-p)}{D_e^2} \sum_{\substack{S_0\subseteq E(G_{-e}):\ \mathrm{Odd}(S_0)=\emptyset\\ S'\subseteq E(G_{-e}):\ \mathrm{Odd}(S')=\{u,v\}}} \frac{\mathrm{wt}(S_0)\mathrm{wt}(S')}{Z_\emptyset^2}\cdot\Pr\big((\omega\to\omega')\in\gamma \mid S_0,S'\big),
\end{aligned} \tag{23}$$

where the last equality uses (14) and (22). We bound the last sum in two stages: first at the level of the subgraph-world path, and then lifting to random-cluster transitions to finish.

Let e' be the unique edge on which the fixed initial transition $\omega \to \omega'$ differs. For a fixed pair (S_0, S'), let $L := L_{S_0\triangle S'}$ denote the simple path between u and v in G_{-e}, oriented from u to v. Let $(e_1, \dots, e_r)$ denote its edges in this order. Define an auxiliary path $(W_0, W_1, \dots, W_r)$ of subgraph-world configurations by

$$W_0 := S_0, \qquad W_i := W_{i-1} \triangle \{e_i\} = S_0 \triangle \{e_1,\dots,e_i\}, \qquad \text{for } i = 1,\dots,r.$$

Thus, at step i we change the membership of e_i: we add it if it is absent and delete it if it is present. After all edges of L have been changed, the auxiliary path ends at $W_r = S_0 \triangle L = S_1$. Note that (S_0, S_1) is exactly the output of the switching coupling in Definition 16.

Recall that in Case (2b), after obtaining (S_0, S_1), we use the same completion variables $(B_f)_{f\in E(G_{-e})}$ to construct (R_0, R_1) as in (12). The local coupling then outputs $X = R_0$ and

$X' = R_1 \cup \{e\}$. The transport path γ first changes the edges of $R_0 \triangle R_1$, in their order along L, thereby moving from R_0 to R_1, and finally opens e. In particular, its initial part changes only edges in $R_0 \triangle R_1 \subseteq L$. Therefore, if $e' \notin L$, γ never changes e' and cannot contain the fixed transition $\omega \to \omega'$. Otherwise, let j be the index such that $e' = e_j$ and let

$$W := W_{j-1} = S_0 \triangle \{e_1, \dots, e_{j-1}\}.$$

Thus, W is the state of the auxiliary path immediately before it changes e', and it is determined by S_0, S', and the fixed edge e'.

We now define an auxiliary flow on the subgraph-world paths. For a fixed pair (S_0, S'), remove the conditional-probability factor $\Pr((\omega \to \omega') \in \gamma \mid S_0, S')$ from its summand in (23), and send the remaining amount

$$\frac{2p(1-p)}{D_e^2} \cdot \frac{\mathrm{wt}(S_0)\mathrm{wt}(S')}{Z_\emptyset^2}$$

through every transition of its auxiliary path. Fix an auxiliary transition $W \to W \triangle \{e'\}$ and consider all pairs (S_0, S') whose auxiliary paths use this transition. For each such pair, define

$$U := S_0 \triangle S' \triangle W. \tag{24}$$

Since every edge of L belongs to exactly one of S_0 and S', flipping a prefix of L merely exchanges the two edge contributions. Hence

$$\mathrm{wt}(S_0)\mathrm{wt}(S') = \mathrm{wt}(W)\mathrm{wt}(U). \tag{25}$$

Moreover, $U \triangle W = S_0 \triangle S'$ recovers the deterministically selected path L. The location of e' on the oriented path L then recovers the toggled prefix, and hence S_0 and S'. Thus, (24) is an injective encoding for fixed e, W, and e'.

Let $(v_0, v_1, \dots, v_r)$ denote the vertices of the oriented path L, where $v_0 = u$ and $v_r = v$. Since $e' = e_j = v_{j-1}v_j$, let $z := v_{j-1} \neq v$. Thus, z is the endpoint of e' reached before e' is changed along the path from u to v. Since S_0 is even and S' has odd set $\{u, v\}$, changing the prefix from u to z gives

$$\mathrm{Odd}(W) = \begin{cases} \emptyset, & z = u, \\ \{u, z\}, & z \neq u, \end{cases} \qquad \text{and} \qquad \mathrm{Odd}(U) = \{v, z\}.$$

For the fixed state W, the vertex z is determined: if W is even, then $z = u$; otherwise, z is the unique vertex in $\mathrm{Odd}(W) \setminus \{u\}$. Using the injection, (25), and (21), the flow through the fixed auxiliary transition $W \to W \triangle \{e'\}$ is at most the following quantity:

$$\begin{aligned} \sum_{\substack{(S_0,S')\colon \text{their auxiliary path uses} \\ W \to W \triangle \{e'\}}} \frac{2p(1-p)}{D_e^2} \cdot \frac{\mathrm{wt}(S_0)\mathrm{wt}(S')}{Z_\emptyset^2} &\le \frac{2p(1-p)}{D_e^2} \cdot \frac{\mathrm{wt}(W)}{Z_\emptyset^2} \sum_{U\colon \mathrm{Odd}(U) = \{v,z\}} \mathrm{wt}(U) \\ &= \frac{2p(1-p)}{D_e^2} \cdot \frac{\mathrm{wt}(W) Z_{\{v,z\}}}{Z_\emptyset^2} \le \frac{2p(1-p)}{D_e^2} \cdot \frac{\mathrm{wt}(W)}{Z_\emptyset}, \end{aligned} \tag{26}$$

where the first inequality uses the injective encoding and (25), the equality follows from the definition of $Z_{\{v,z\}}$, and the last inequality uses (21).

So far, we have bounded the congestion through a fixed auxiliary transition. We now lift this bound to the actual random-cluster path γ and bound the congestion required in (17). The argument is similar to that in [GJ18]. Recall that a random-cluster configuration is lifted from a subgraph-world configuration using independent Bernoulli random variables $(B_g)_{g \in E(G_{-e})}$ with parameter $p/(2-p)$ as in (12). More precisely, for a subgraph-world configuration S, let

$$\widehat{S} := S \cup \{g \in E(G_{-e}) \setminus S : B_g = 1\}.$$

Conditional on S, the lifted random-cluster configuration $\widehat{S}$ has distribution $K(S, \cdot)$, where K is the lifting kernel defined in (18).

Every auxiliary path $(W_0, W_1, \dots, W_r)$ constructed above can be lifted by using the same variables $(B_g)_{g \in E(G_{-e})}$ for all of its states. Since $W_0 = S_0$ and $W_r = S_1$, the first and last lifted states are respectively R_0 and R_1 from (12). Two consecutive lifted states either coincide or differ only on the edge changed by the corresponding auxiliary transition. After deleting consecutive repeated states, the lifted path changes precisely the edges of $R_0 \triangle R_1$, in their order along L. Appending the final transition that opens e therefore gives exactly the random-cluster path γ used in the construction of Γ_e.

It remains to determine how much traffic an auxiliary transition sends through a fixed random-cluster transition. Consider an auxiliary transition $S \to S \triangle \{e'\}$. If $e' \notin S$, this transition adds e'. For a specified random-cluster configuration R with $e' \notin R$, the transition $S \to S \triangle \{e'\}$ is lifted to $R \to R \cup \{e'\}$ with probability $K(S, R)$. If $e' \in S$, the auxiliary transition deletes e', and the current lifted configuration necessarily contains e'. Fix a random-cluster configuration R with $e' \in R$. The event $\widehat{S} = R$ has probability $K(S, R)$. Since e' already belongs to S, the variable $B_{e'}$ is not used when lifting S and is independent of this event. After e' is deleted from S, the next lifted configuration is $R \setminus \{e'\}$ if $B_{e'} = 0$, and remains R if $B_{e'} = 1$. Therefore, the probability of obtaining the random-cluster transition $R \to R \setminus \{e'\}$ is $(1 - \frac{p}{2-p})K(S, R) \le K(S, R)$. In the latter case, the two lifted configurations coincide, so the repeated state is deleted from the random-cluster path. Thus, in both the addition and deletion cases, the probability that this auxiliary transition produces a specified random-cluster transition whose initial configuration is R is at most $K(S, R)$.

Returning to the fixed transition $\omega \to \omega'$, set $R := \omega \cap E(G_{-e})$. We now sum (26) over the subgraph-world states W that can produce the fixed transition $\omega \to \omega'$. By the lifting bound above, the traffic lifted from a fixed W is at most $\frac{2p(1-p)}{D_e^2} \cdot \frac{\mathrm{wt}(W)}{Z_\emptyset} \cdot K(W, R)$. Recall that, for each subgraph-world state W, the vertex z is the endpoint of e' reached immediately before e' is changed along the oriented path L. As W varies, z can be either endpoint of e'. We partition the subgraph-world states W into at most two groups according to this endpoint. Within one such group, the vertex z is the same for every W. If this common vertex satisfies $z \neq u$, then $\mathrm{Odd}(W) = \{u, z\}$ for every W in the group, and (19) gives

$$\sum_{\substack{W \subseteq E(G_{-e}) \\ \mathrm{Odd}(W) = \{u,z\}}} \frac{\mathrm{wt}(W)}{Z_\emptyset} \cdot K(W, R) \le \pi_{E-e}^{e \leftarrow 0}(R).$$

This bound is applied separately to each endpoint z of e' satisfying $z \neq u$. There are two mutually exclusive cases. If u is not an endpoint of e', then both endpoints give groups with $z \neq u$, so there are two groups of the type just considered and no group with $z = u$. If u is an endpoint of e', then at most one group has $z \neq u$, while the other possible group has $z = u$. In the latter group, W is even, and Proposition 14 gives

$$\sum_{\substack{W \subseteq E(G_{-e}) \\ \mathrm{Odd}(W) = \emptyset}} \frac{\mathrm{wt}(W)}{Z_\emptyset} \cdot K(W, R) = \pi_{E-e}^{e \leftarrow 0}(R).$$

Thus, the sum over the states W in each group is at most $\pi_{E-e}^{e \leftarrow 0}(R)$. Since e' has two endpoints, there are at most two groups in total. For each endpoint z of e', let $\mathcal{W}_z$ denote the group of subgraph-world states that can produce $\omega \to \omega'$ and whose associated endpoint is z. Combining (26) with the lifting bound and then applying the preceding bound to each group gives

$$\begin{aligned}
&\pi_e(0)\pi_e(1) \Pr_{\gamma \sim \Gamma_e}\left((\omega \to \omega') \in \gamma\right) \\
&\le \frac{2p(1-p)}{D_e^2} \sum_{z:\ z \text{ is an endpoint of } e'} \ \sum_{W \in \mathcal{W}_z} \frac{\mathrm{wt}(W)}{Z_\emptyset} \cdot K(W, R) \\
&\le \frac{2p(1-p)}{D_e^2} \sum_{z:\ z \text{ is an endpoint of } e'} \pi_{E-e}^{e \leftarrow 0}(R) \\
&\le \frac{4p(1-p)}{D_e^2} \cdot \pi_{E-e}^{e \leftarrow 0}(R).
\end{aligned}$$

Every state along the initial part has e closed, and hence by (14),

$$\pi(\omega) = \pi_e(0)\pi_{E-e}^{e\leftarrow 0}(R) = \frac{2(1-p)}{D_e} \cdot \pi_{E-e}^{e\leftarrow 0}(R).$$

Since $D_e \geq 2-p$, combining the last two bounds gives

$$\pi_e(0)\pi_e(1) \Pr_{\gamma\sim\Gamma_e}\big((\omega \to \omega') \in \gamma\big) \leq \frac{2p}{D_e} \cdot \pi(\omega) \leq C'_p \pi(\omega).$$

Therefore, (17) holds when $\omega \to \omega'$ is an initial transition of a path sampled from Γ_e.

It remains to consider the final transition of γ, which opens e. Fix a transition $\omega \to \omega'$ of this type. Then $\omega = R$ for some $R \subseteq E(G_{-e})$ and $\omega' = R \cup \{e\}$. The path γ uses this transition exactly when its state immediately before opening e is R. By the construction of the local coupling, the restriction of this pre-final state to $E(G_{-e})$ has distribution $\pi_{E-e}^{e\leftarrow 1}$. Hence,

$$\Pr_{\gamma\sim\Gamma_e}\big((\omega \to \omega') \in \gamma\big) = \pi_{E-e}^{e\leftarrow 1}(R).$$

The left-hand side of (17) is therefore

$$\pi_e(0)\pi_e(1)\pi_{E-e}^{e\leftarrow 1}(R) = \pi_e(0)\pi(\omega').$$

To compare this quantity with $\pi(\omega)$, observe that opening e contributes a factor $p/(1-p)$ to the random-cluster weight. If the endpoints of e are disconnected in (V, R), opening e also decreases the number of connected components by one and contributes an additional factor $1/2$. Consequently,

$$\frac{\pi(\omega')}{\pi(\omega)} = \begin{cases} \dfrac{p}{1-p}, & \text{if the endpoints of } e \text{ are connected in } (V,R), \\ \dfrac{p}{2(1-p)}, & \text{otherwise.} \end{cases}$$

Using $\pi_e(0) = 2(1-p)/D_e$, it follows that

$$\frac{\pi_e(0)\pi(\omega')}{\pi(\omega)} = \begin{cases} 2p/D_e, & \text{if the endpoints of } e \text{ are connected in } (V,R), \\ p/D_e, & \text{otherwise} \end{cases} \leq C'_p.$$

Thus, (17) also holds when $\omega \to \omega'$ is the final transition of γ. This argument uses only the marginal distribution of the endpoint of the local coupling, so it also applies when $q = 0$.

We have now proved (17) for both types of transitions, and hence (16) follows from the path length bound. As explained after (16), summing over the non-loop edges proves the required congestion estimate for every non-loop transition. Together with the self-loop calculation preceding Section 4.1, this proves the claimed congestion estimate, assuming Claim 19. □

Proof of Claim 19. Fix two distinct vertices $u_0, u_1 \in V$ and $R \subseteq E(G_{-e})$. Only $S \subseteq R$ contribute to its left-hand side, and for every such S, direct substitution gives

$$\mathrm{wt}(S)K(S,R) = 2^{-|R|}p^{|R|}(1-p)^{m-1-|R|},$$

which is independent of S. By the handshaking lemma, a set $S \subseteq R$ with $\mathrm{Odd}(S) = \{u_0, u_1\}$ can exist only if u_0 and u_1 are connected in (V, R). If they are not connected, both sides of (19) are zero, and the claim holds. If they are connected, fix a u_0-to-u_1 path $L \subseteq R$. The map $S \mapsto S \triangle L$ is a bijection between the sets satisfying $\mathrm{Odd}(S) = \{u_0, u_1\}$ and the even subgraphs of (V, R). Thus, the sum in (19) has exactly the same number of equal-weight terms as the sum with $\mathrm{Odd}(S) = \emptyset$. For the empty odd-vertex set, the lifting coupling in Proposition 14 shows that this sum equals $\pi_{E-e}^{e\leftarrow 0}(R)$. Therefore,

$$\sum_{\substack{S\subseteq E(G_{-e}) \\ \mathrm{Odd}(S)=\{u_0,u_1\}}} \frac{\mathrm{wt}(S)}{Z_\emptyset} \cdot K(S,R) = \sum_{\substack{S\subseteq E(G_{-e}) \\ \mathrm{Odd}(S)=\emptyset}} \frac{\mathrm{wt}(S)}{Z_\emptyset} \cdot K(S,R) = \pi_{E-e}^{e\leftarrow 0}(R).$$

This concludes the proof of Claim 19. □

4.3. Proof of the mixing bound.

Proof of Theorem 4. We apply Theorem 11. Recall from Section 3 that P^τ denotes the heat-bath Glauber dynamics for the conditional distribution π^τ, regarded as a chain on $\{0,1\}^E$. Since P^τ fixes the pinned coordinates and updates only the k free coordinates, we may identify it with the heat-bath Glauber dynamics for the projection of π^τ onto those coordinates. Under any pinning, this projected distribution on the k unpinned edges is a random-cluster model with parameter p on a multigraph obtained from G by deleting and contracting edges. This multigraph has at most n vertices and exactly k unpinned edges. Conditional on all other edges, the probability that an unpinned edge is open is either p, if its endpoints are already connected, or $p/(2-p)$, otherwise. Hence the two possible values of this edge have conditional marginal probability at least

$$\phi_p := \min\left\{\frac{p}{2-p}, 1-p\right\} > 0.$$

Since p is fixed, ϕ_p is a constant in $(0,1/2)$.

Apply Lemma 12 to the conditional random-cluster model on the k unpinned edges. Let n_τ be the number of vertices in the multigraph obtained after the deletions and contractions. In the notation of Lemma 12, the vertex count is $n_\tau \le n$, while the edge count is k, rather than the original m. Thus, the congestion bound in Lemma 12 is at most $C_p n_\tau k^2 \le C_p n k^2$, and we may take $c_k := C_p n k^2$. Therefore, Theorem 11 gives

$$\rho(P) \ge \frac{1-2\phi_p}{8\log((1-\phi_p)/\phi_p)} \left(\sum_{k=1}^{m} \frac{C_p n k^2}{k}\right)^{-1} = \Omega_p\left(\frac{1}{nm^2}\right).$$

It remains to bound the smallest stationary probability. For every $S \subseteq E$, its unnormalized random-cluster weight is at least $\min\{p, 1-p\}^m$, while

$$Z_{\mathrm{RC}} = \sum_{S\subseteq E} p^{|S|}(1-p)^{m-|S|}2^{\kappa(S)} \le 2^n.$$

Consequently,

$$\pi_{\min} \ge 2^{-n}\min\{p,1-p\}^m,$$

which implies $\log\log\frac{1}{\pi_{\min}} = O_p\big(\log(n+m)\big)$. The Glauber dynamics is irreducible, aperiodic, and reversible with respect to π. Applying the log-Sobolev mixing bound Equation (5) yields

$$T_{\mathrm{mix}}(P,\varepsilon) = O_p\left(nm^2\left(\log(n+m) + \log\frac{1}{\varepsilon}\right)\right),$$

which proves Theorem 4. □

5. Proof of the main theorem

Finally, for completeness, we give the full algorithm and prove the main theorem.

Proof of Theorem 1. We first apply Theorem 2 with error $\varepsilon/2$. This takes $O(m\log^2 n + n)$ time and gives $|V'| \le n$ and $|E'| = O_\beta(n\log(2n/\varepsilon))$. If $E' = \emptyset$, since G is connected, G' consists of one vertex; assigning a uniformly random spin to it and applying the function Lift takes $O(n)$ time. Hence, we may assume that $E' \neq \emptyset$ and $|V'| \ge 2$.

We next sample from the random-cluster model on G' with $p := 1 - 1/\beta$ using the Glauber dynamics. We start from the empty configuration and initialize the fully dynamic connectivity data structure with vertex set V' and no open edge. Suppose that the current configuration is X. The Glauber dynamics draws a uniformly random edge $e = (u,v) \in E'$, updates $X \leftarrow X \setminus \{e\}$ if $e \in X$, and queries whether u and v are connected in (V', X). It then updates $X \leftarrow X \cup \{e\}$ with probability p if they are connected and with probability $p/(2-p)$ otherwise. Consequently, it suffices to support the following operations:

- update $X \leftarrow X \cup \{e\}$ for any given $e \in E'$;
- update $X \leftarrow X \setminus \{e\}$ for any given $e \in E'$;

- query whether u and v are connected in (V', X) for any given $u, v \in V'$.

Using the dynamic connectivity data structure in [WN13, CZ23], each update takes $O(\log^2 n)$ amortized time, after $\widetilde{O}(n + |E'|)$ initialization time for the data structure and the edge states.

Combining this implementation with Theorem 4, we obtain an $\varepsilon/2$-approximate random-cluster sample in time $\widetilde{O}_\beta\left(n^3 \log^3 \frac{1}{\varepsilon}\right)$. Applying the Edwards–Sokal coupling as in Proposition 3, followed by the function Lift, takes $O(n + |E'|)$ time and does not increase the $\varepsilon/2$ sampling error. Together with the $\varepsilon/2$ sparsification error, this gives the required error bound. The total running time is therefore $\widetilde{O}(m+n) + \widetilde{O}_\beta\left(n^3 \log^3 \frac{1}{\varepsilon}\right)$. In particular, if G is simple, the running time is $\widetilde{O}_\beta\left(n^3 \log^3 \frac{1}{\varepsilon}\right)$. □

6. Concluding remarks

It is natural to wonder what is the true mixing time for the random-cluster dyanmics. Although we believe the upper bound in Theorem 4 can be improved, we conjecture that it cannot be down to near-linear time. The reason is as follows. First, the closely related Swendsen-Wang dynamics mixes in time $\Theta\left(n^{1/4}\right)$ at the critical temperature for the mean-field case [LNNP14]. While this does not directly translate into a lower bound for the random-cluster dynamics, it is a strong sign that the RC dynamics does not mix in near-linear time.

Another piece of evidence is the conjectured or predicted behaviour of the Ising model on the 3D torus at the critical temperature. Basically, the mixing time can be lower bounded by the inverse of the spectral gap, which, in turn, can be bounded using the variational characterisation, similar to (4). Taking the size of the open edge set as the test function, the mixing time is lower bounded by the variance of the size of this set. Using predictions of the critical Ising model on the 3D torus, this variance is super-linear. See [Dum22, Section 4.2.1] for some rigorous background and [HP98, KPSDV16] for some numerical results.

References

[AJK+22] Nima Anari, Vishesh Jain, Frederic Koehler, Huy Tuan Pham, and Thuy-Duong Vuong, *Entropic independence: optimal mixing of down-up random walks*, In STOC, 2022, pp. 1418–1430. 1

[BK15] András A. Benczúr and David R. Karger, *Randomized approximation schemes for cuts and flows in capacitated graphs*, SIAM J. Comput. **44** (2015), no. 2, 290–319. 2, 4

[CFJ+25] Xiaoyu Chen, Weiming Feng, Zhe Ju, Tianshun Miao, Yitong Yin, and Xinyuan Zhang, *Faster mixing of the Jerrum–Sinclair chain*, In FOCS, 2025, pp. 1007–1028. 3, 6, 7

[CFYZ22] Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang, *Optimal mixing for two-state anti-ferromagnetic spin systems*, In FOCS, 2022, pp. 588–599. 1

[CGZZ24] Xiaoyu Chen, Heng Guo, Xinyuan Zhang, and Zongrui Zou, *Near-linear time samplers for matroid independent sets with applications*, In APPROX/RANDOM, 2024, pp. 32:1–32:12. 2, 19, 20

[CJM+26] Xiaoyu Chen, Zhe Ju, Tianshun Miao, Yitong Yin, and Xinyuan Zhang, *Edge-tilting field dynamics: Rapid mixing at the uniqueness threshold and optimal mixing for Swendsen-Wang dynamics*, arXiv **abs/2604.10525** (2026). 1

[CLV23] Zongchen Chen, Kuikui Liu, and Eric Vigoda, *Rapid mixing of Glauber dynamics up to uniqueness via contraction*, SIAM J. Comput. **52** (2023), no. 1, 196–237. 1

[CZ23] Xiaoyu Chen and Xinyuan Zhang, *A near-linear time sampler for the Ising model with external field*, In SODA, 2023, pp. 4478–4503. 1, 3, 18

[DGU14] Martin Dyer, Catherine Greenhill, and Mario Ullrich, *Structure and eigenvalues of heat-bath Markov chains*, Linear Algebra Appl. **454** (2014), 57–71. 6

[DS96] Persi Diaconis and Laurent Saloff-Coste, *Logarithmic Sobolev inequalities for finite Markov chains*, Ann. Appl. Probab. **6** (1996), no. 3, 695–750. 6

[Dum22] Hugo Duminil-Copin, *100 years of the (critical) Ising model on the hypercubic lattice*, Proceedings of the International Congress of Mathematicians 2022, EMS Press, 2022, pp. 164–210. 18

[ES88] Robert G. Edwards and Alan D. Sokal, *Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm*, Phys. Rev. D **38** (1988), no. 6, 2009–2012. 2, 3

[FK72] Cees M. Fortuin and Piet W. Kasteleyn, *On the random-cluster model. I. Introduction and relation to other models*, Physica **57** (1972), 536–564. 2

[GJ09] Geoffrey R. Grimmett and Svante Janson, *Random even graphs*, Electron. J. Combin. **16** (2009), no. 1, R46. 9, 10

[GJ18] Heng Guo and Mark Jerrum, *Random cluster dynamics for the Ising model is rapidly mixing*, Ann. Appl. Probab. **28** (2018), no. 2, 1292–1313. 3, 14

[GJ19] ______, *A polynomial-time approximation algorithm for all-terminal network reliability*, SIAM J. Comput. **48** (2019), no. 3, 964–978. 2

[GŠV16] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda, *Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models*, Comb. Probab. Comput. **25** (2016), no. 4, 500–559. 1

[Har60] Theodore E. Harris, *A lower bound for the critical probability in a certain percolation process*, Proc. Cambridge Philos. Soc. **56** (1960), no. 1, 13–20. 20

[HP98] Martin Hasenbusch and Klaus Pinn, *A_+/A_-, α, ν, and $f_s\xi^3$ from 3D Ising energy and specific heat*, J. Phys. A **31** (1998), no. 29, 6157–6166. 18

[JS93] Mark Jerrum and Alistair Sinclair, *Polynomial-time approximation algorithms for the Ising model*, SIAM J. Comput. **22** (1993), no. 5, 1087–1116. 1, 3, 9

[Kar93] David R. Karger, *Global min-cuts in RNC, and other ramifications of a simple min-cut algorithm*, In SODA, 1993, pp. 21–30. 2, 4

[Kar99] ______, *Random sampling in cut, flow, and network design problems*, Math. Oper. Res. **24** (1999), no. 2, 383–413. 2, 4

[KPSDV16] Filip Kos, David Poland, David Simmons-Duffin, and Alessandro Vichi, *Precision islands in the Ising and $O(N)$ models*, J. High Energy Phys. **2016** (2016), no. 8, No. 36. 18

[LLY13] Liang Li, Pinyan Lu, and Yitong Yin, *Correlation decay up to uniqueness in spin systems*, In SODA, 2013, pp. 67–84. 1

[LNNP14] Yun Long, Asaf Nachmias, Weiyang Ning, and Yuval Peres, *A power law of order 1/4 for critical mean-field Swendsen–Wang dynamics*, Mem. Amer. Math. Soc. **232** (2014), no. 1092, vi+84. 18

[RW99] Dana Randall and David Wilson, *Sampling spin configurations of an Ising system*, In SODA, 1999, pp. 959–960. 1

[Sin92] Alistair Sinclair, *Improved bounds for mixing rates of Markov chains and multicommodity flow*, Combin. Probab. Comput. **1** (1992), no. 4, 351–370. 3

[SS14] Allan Sly and Nike Sun, *Counting in two-spin models on d-regular graphs*, Ann. Probab. **42** (2014), no. 6, 2383–2416. 1

[SST14] Alistair Sinclair, Piyush Srivastava, and Marc Thurley, *Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs*, J. Stat. Phys. **155** (2014), no. 4, 666–686. 1

[WN13] Christian Wulff-Nilsen, *Faster deterministic fully-dynamic graph connectivity*, In SODA, 2013, pp. 1757–1769. 3, 18

Appendix A. Sparsification for all-terminal network reliability

In the last section we show that the same contraction sparsification in Section 2 also works for all-terminal network reliability. This improves the previous running time from $\widetilde{O}(mn/\varepsilon^2)$ [CGZZ24] to $\widetilde{O}(m+n^2/\varepsilon^2)$.

Let G be a multigraph. We may assume that G is connected and loopless. For $p \in (0,1)$, let $\mathrm{Rel}_p(G)$ denote the probability that G_p is connected, where G_p is the spanning subgraph obtained by deleting every edge independently with probability p.

Theorem 20 (All-terminal reliability sparsification)**.** *For every constant $p \in (0,1)$, there is an algorithm that, given a connected loopless multigraph $G=(V,E)$ and an error bound $\eta \in (0,1)$, constructs a multigraph H via contraction in time $O(m\log^2 n + n)$ such that*

$$|V(H)| \le n, \qquad \text{and} \qquad |E(H)| = O_p\big(n\log(n/\eta)\big),$$

where $n := |V|$ and $m := |E|$, and

$$(1-\eta)\,\mathrm{Rel}_p(H) \le \mathrm{Rel}_p(G) \le \mathrm{Rel}_p(H).$$

Consequently, for every $\varepsilon \in (0,1)$, there is an FPRAS for $\mathrm{Rel}_p(G)$ with running time $\widetilde{O}_p(m + n^2/\varepsilon^2)$, where $\widetilde{O}_p$ hides factors that are polylogarithmic in n/ε and constants that depend on p.

Proof. Apply Proposition 7 to G to obtain tight strengths $(\widetilde{s}_e)_{e\in E}$ satisfying (1), and let

$$K := \left\lceil \frac{\log(2n^5/\eta)}{\log(1/p)} \right\rceil.$$

Let $E_{\ge K} := \{e \in E : \widetilde{s}_e \ge K\}$, and let $C_1,\dots,C_N$ be the connected components of $(V, E_{\ge K})$. Construct H by contracting each C_i to a single vertex, deleting the resulting self-loops, and retaining every edge between distinct components with its multiplicity. Every retained edge e satisfies $\widetilde{s}_e < K$, and hence

$$|V(H)| = N \leq n, \text{ and } |E(H)| < K \sum_{e \in E(H)} \frac{1}{\widetilde{s}_e} \leq K \cdot c_0 n = O_p\big(n \log(n/\eta)\big).$$

As in the proof of Theorem 2, let $B_1, \ldots, B_L$ be the inclusion-maximal K-strong vertex sets of G. These sets partition V, and every C_i is contained in a unique B_ℓ. Let $X \subseteq E$ be the random set of retained edges, and $G_\ell := (B_\ell, X \cap E(G[B_\ell]))$ be the retained subgraph of B_ℓ. Let $\mathcal{A}$ be the event that G_ℓ is connected for every $\ell \in [L]$. For a fixed B_ℓ with at least two vertices, the argument in Theorem 2, via Proposition 8, gives

$$\Pr\big(G_\ell \text{ is disconnected}\big) \leq \sum_{j \geq 1} n^{2(j+1)} \cdot p^{jK} = \frac{n^4 p^K}{1 - n^2 p^K}.$$

The choice of K implies $p^K \leq \eta/(2n^5)$, so the last expression is at most η/n. Since there are at most n maximal K-strong components, a union bound yields $\Pr(\mathcal{A}) \geq 1 - \eta$.

Let $\mathcal{C}$ be the event that (V, X) is connected, and let $\mathcal{D}$ be the event that the retained subgraph of H is connected. In other words, $\mathrm{Rel}_p(G) = \Pr(\mathcal{C})$ and $\mathrm{Rel}_p(H) = \Pr(\mathcal{D})$. Contracting the sets C_i shows that $\mathcal{C} \subseteq \mathcal{D}$. Moreover, $\mathcal{A} \cap \mathcal{D} \subseteq \mathcal{C}$: identifying all vertices of H whose corresponding blocks C_i lie in the same B_ℓ gives a connected quotient, and given $\mathcal{A}$ every B_ℓ is internally connected in (V, X). Both $\mathcal{A}$ and $\mathcal{D}$ are increasing events under the product measure on X. Therefore, the Harris's inequality [Har60] gives

$$\mathrm{Rel}_p(G) \geq \Pr(\mathcal{A} \cap \mathcal{D}) \geq \Pr(\mathcal{A}) \Pr(\mathcal{D}) \geq (1 - \eta) \mathrm{Rel}_p(H).$$

The inclusion $\mathcal{C} \subseteq \mathcal{D}$ gives the reverse inequality $\mathrm{Rel}_p(G) \leq \mathrm{Rel}_p(H)$.

The tight strengths can be computed in $O(m \log^2 n)$ time, and the components and contraction can be constructed in $O(m + n)$ additional time. This proves the sparsification claim.

For the FPRAS, apply the sparsification with $\eta := \varepsilon/4$ and run the reliability FPRAS obtained from the near-linear sampler of Chen, Guo, Zhang, and Zou [CGZZ24] on H with relative error $\varepsilon/4$. If its output is $\widehat{R}$, then, with its stated success probability,

$$(1 - \varepsilon/4) \mathrm{Rel}_p(H) \leq \widehat{R} \leq (1 + \varepsilon/4) \mathrm{Rel}_p(H).$$

Together with the reliability comparison above, this implies that $\widehat{R}$ is a relative ε-approximation to $\mathrm{Rel}_p(G)$. The FPRAS of [CGZZ24] runs in $\widetilde{O}_p(|V(H)||E(H)|/\varepsilon^2)$ time. Since $|V(H)| \leq n$ and $|E(H)| = O_p(n \log(n/\varepsilon))$, the total running time, including sparsification, is $\widetilde{O}_p(m + n^2/\varepsilon^2)$. $\square$