

Fast Sampling via Spectral Independence Beyond Bounded-degree Graphs

IVONA BEZÁKOVÁ, Rochester Institute of Technology, USA

ANDREAS GALANIS and LESLIE ANN GOLDBERG, University of Oxford, UK

DANIEL ŠTEFANKOVIČ, University of Rochester, USA

Spectral independence is a recently developed framework for obtaining sharp bounds on the convergence time of the classical Glauber dynamics. This new framework has yielded optimal $O(n \log n)$ sampling algorithms on bounded-degree graphs for a large class of problems throughout the so-called uniqueness regime, including, for example, the problems of sampling independent sets, matchings, and Ising-model configurations. Our main contribution is to relax the bounded-degree assumption that has so far been important in establishing and applying spectral independence. Previous methods for avoiding degree bounds rely on using L^p -norms to analyse contraction on graphs with bounded connective constant (Sinclair, Srivastava, and Yin, FOCS'13). The non-linearity of L^p -norms is an obstacle to applying these results to bound spectral independence. Our solution is to capture the L^p -analysis recursively by amortising over the subtrees of the recurrence used to analyse contraction. Our method generalises previous analyses that applied only to bounded-degree graphs. As a main application of our techniques, we consider the random graph $G(n, d/n)$, where the previously known algorithms run in time $n^{O(\log d)}$ or applied only to large d . We refine these algorithmic bounds significantly, and develop fast nearly linear algorithms based on Glauber dynamics that apply to all constant d , throughout the uniqueness regime.

CCS Concepts: • **Theory of computation** → **Random walks and Markov chains**; *Generating random combinatorial structures; Design and analysis of algorithms*;

Additional Key Words and Phrases: Glauber dynamics, spectral independence, sparse random graphs

ACM Reference format:

Ivona Bezákova, Andreas Galanis, Leslie Ann Goldberg, and Daniel Štefankovič. 2024. Fast Sampling via Spectral Independence Beyond Bounded-degree Graphs. *ACM Trans. Algor.* 20, 1, Article 7 (January 2024), 26 pages.

<https://doi.org/10.1145/3631354>

For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. All data is provided in full in the results section of this article.

Ivona Bezákova was supported in part by NSF Grant No. DUE-1819546. Daniel Štefankovič was supported by NSF Grant No. CCF-2007287.

Authors' addresses: I. Bezákova, Rochester Institute of Technology, Rochester, NY, USA, e-mail: ib@cs.rit.edu; A. Galanis and L. A. Goldberg, University of Oxford, Oxford, UK, e-mails: {andreas.galanis, leslie.goldberg}@cs.ox.ac.uk; D. Štefankovič, University of Rochester, Rochester, NY, USA, e-mail: stefanko@cs.rochester.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

1549-6325/2024/01-ART7 \$15.00

<https://doi.org/10.1145/3631354>

1 INTRODUCTION

Spectral independence method was introduced by Anari, Liu, and Oveis Gharan [3] as a framework to obtain polynomial bounds on the mixing time of Glauber dynamics. Originally based on a series of works on high-dimensional expansion [1, 15, 27, 28, 31], it has since then been developed further using entropy decay by Chen, Liu, and Vigoda [13] who obtained optimal $O(n \log n)$ mixing results on graphs of bounded maximum degree Δ whenever the framework applies. This article focuses on relaxing the bounded-degree assumption of Reference [13], in sparse graphs where the maximum degree is not the right parameter to capture the density of the graph.

As a running example, we will use the problem of sampling (weighted) independent sets, also known as the sampling problem from the hard-core model. For a graph $G = (V, E)$, the hard-core model with parameter $\lambda > 0$ specifies a distribution $\mu_{G, \lambda}$ on the collection of independent sets of G , where for an independent set I it holds that $\mu_{G, \lambda}(I) = \lambda^{|I|} / Z_{G, \lambda}$ where $Z_{G, \lambda}$ is the partition function of the model (the normalising factor that makes the probabilities add up to 1). For bounded-degree graphs of maximum degree $d + 1$ (where $d \geq 2$ is an integer), it is known that the problems of sampling and approximately counting from this model undergo a computational transition at $\lambda_c(d) = \frac{d^d}{(d-1)^{d+1}}$, the so-called uniqueness threshold [20, 38, 40]: they are poly-time solvable when $\lambda < \lambda_c(d)$, and computationally intractable for $\lambda > \lambda_c(d)$. Despite this clear complexity picture, prior to the introduction of spectral independence, the algorithms for $\lambda < \lambda_c(d)$ were based on elaborate enumeration techniques whose running times scale as $n^{O(\log d)}$ [29, 32, 33, 40]. The analysis of Glauber dynamics¹ using spectral independence in the regime $\lambda < \lambda_c(d)$ yielded initially $n^{O(1)}$ algorithms for any d [3] (see also Reference [14]), and then $O(n \log n)$ for bounded-degree graphs [13]. More recently, Chen, Feng, Yin, and Zhang [10] obtained $O(n^2 \log n)$ results for arbitrary graphs $G = (V, E)$ that apply when $\lambda < \lambda_c(\Delta_G - 1)$, where Δ_G is the maximum degree of G (see also Reference [23] for related results when Δ_G grows like $\log n$); this has been further refined to $O(n \log n)$ in References [2, 9, 11].

The maximum degree is frequently a bad measure of the density of the graph, especially for graphs with unbounded-degree. One of the most canonical examples is the random graph $G(n, d/n)$ where the maximum degree grows with n but the average degree is d , and therefore one would hope to be able to sample from $\mu_{G, \lambda}$ for λ up to some constant, instead of $\lambda = o(1)$ that the previous results yield. In this direction, [35, 37] obtained an $n^{O(\log d)}$ algorithm based on correlation decay that applies to all $\lambda < \lambda_c(d)$ for all graphs with “connective constant” bounded by d (meaning, roughly, that for all $\ell = \Omega(\log n)$ the number of length- ℓ paths starting from any vertex is bounded by d^ℓ). The result of Reference [35] applies to $G(n, d/n)$ for all $d > 0$. In terms of Glauber dynamics on $G(n, d/n)$, [30] showed an $n^{1+\Omega(1/\log \log n)}$ lower bound on the mixing time in the case of the Ising model; this lower bound actually applies to most well-known models, and in particular rules out $O(n \log n)$ mixing time results for the hard-core model when $\lambda = \Omega(1)$. The mixing-time lower bound on $G(n, d/n)$ has only been matched by complementary fast mixing results in models with strong monotonicity properties, see Reference [30] for the ferromagnetic Ising model and Reference [7] for the random-cluster model. Such monotonicity properties unfortunately do not hold for the hard-core model, and the best known results [16, 18] for Glauber dynamics on $G(n, d/n)$ give an n^C algorithm for $\lambda < 1/d$ and sufficiently large d (where C is a constant depending on d). Our goal in this article is to go all the way up to $\lambda_c(d)$ (which converges to e/d so is larger than $1/d$).

¹Recall, for a graph G , the Glauber dynamics for the hard-core model iteratively maintains a random independent set $(I_t)_{t \geq 0}$, where at each step t a vertex v is chosen **uniformly at random (u.a.r.)** and, if $I_t \cup \{v\}$ is independent, then it sets $I_{t+1} = I_t \cup \{v\}$ with probability $\frac{\lambda}{\lambda+1}$; otherwise, $I_{t+1} = I_t \setminus \{v\}$. The mixing time is the maximum number (over the starting I_0) of steps t needed to get within total variation distance $1/4$ of $\mu_{G, \lambda}$; see Section 4.1 for the precise definitions.

Our main contribution is to obtain nearly linear-time algorithms on $G(n, d/n)$, for all of the models considered in Reference [35], i.e., the hard-core model, the monomer-dimer model (weighted matchings), and the antiferromagnetic Ising model. Key to our results are new spectral independence bounds for any $d > 0$ in the regime $\lambda < \lambda_c(d)$ for arbitrary graphs $G = (V, E)$ in terms of their “ d -branching value” (which resembles the connective-constant notion of Reference [35]). To state our main theorem for the hard-core model on $G(n, d/n)$, we first extend the definition of $\lambda_c(d)$ to all reals $d > 0$ by setting $\lambda_c(d) = \frac{d^d}{(d-1)^{d+1}}$ for $d > 1$, and $\lambda_c(d) = \infty$ for $d \in (0, 1)$. We use the term “**with probability (whp)** over the choice of $G \sim G(n, d/n)$ ” as a shorthand for “as n grows large, with probability $1 - o(1)$ over the choice of $G(n, d/n)$.” An ε -sample from a distribution μ supported on a finite set Ω is a random $\sigma \in \Omega$ whose distribution ν satisfies $\|\nu - \mu\|_{\text{TV}} \leq \varepsilon$, where $\|\nu - \sigma\|_{\text{TV}} = \frac{1}{2} \sum_{\sigma \in \Omega} |\nu(\sigma) - \mu(\sigma)|$.

THEOREM 1.1. *Let $d, \lambda > 0$ be such that $\lambda < \lambda_c(d)$. For any arbitrarily small constant $\theta > 0$, there is an algorithm such that, whp over the choice of $G \sim G(n, d/n)$, when the algorithm is given as input the graph G and an arbitrary rational $\varepsilon > 0$, it outputs an ε -sample from $\mu_{G, \lambda}$ in time $n^{1+\theta} \log \frac{1}{\varepsilon}$.*

The reader might wonder why is there no constant in front of the running time (in Theorems 1.1, 1.2, and 1.3) or why is there no requirement that n is sufficiently large? The assumption that n is sufficiently large is taken care of in the whp condition: there is a function $f_{d, \lambda, \theta} : \mathbb{Z} \rightarrow \mathbb{R}$ such that $\lim_{n \rightarrow \infty} f_{d, \lambda, \theta}(n) = 0$ and the “whp” means with probability $\geq 1 - f_{d, \lambda, \theta}(n)$; the function $f_{d, \lambda, \theta}$ equals 1 for small n (making the conclusion trivial for such n). Moreover, the family of $O(n^{1+\theta})$ algorithms from Theorem 1.1 can be turned into an $n^{1+o(1)}$ algorithm, see Remark 6.7 for a discussion.

We remark also here that the algorithm of Theorem 1.1 (as well as Theorems 1.2 and 1.3 below) can also recognise in time $n^{1+o(1)}$ whether the graph $G \sim G(n, d/n)$ is a “good” graph, i.e., we can formulate graph properties that guarantee the success of the algorithm, are satisfied whp, and are also efficiently verifiable, see Section 6.4 for details.

The key to obtaining Theorem 1.1 is to bound the spectral independence of the Gibbs distribution on $G(n, d/n)$. The main strategy that has been applied so far to bound spectral independence is to adapt suitably correlation decay arguments and, therefore, it is tempting to use the correlation decay analysis of Reference [35]. This poses new challenges in our setting, since Reference [35] uses an L^p -norm analysis of correlation decay on trees, and the non-linearity of L^p -norms is an obstacle to converting their analysis into spectral independence bounds (in contrast, for bounded-degree graphs, the L^∞ -norm is used, which can be converted to spectral independence bounds using a purely analytic approach, see Reference [14]). Our solution to work around that is to “linearise” the L^p -analysis by taking into account the structural properties of subtrees. This allows us to amortise over the tree-recurrence using appropriate combinatorial information (the d -branching values) and to bound subsequently spectral independence; details are given in Section 3, see Lemmas 3.6 and 3.8 (and Equation (2) that is at the heart of the argument). Once the spectral independence bound is in place, further care is needed to obtain the fast nearly linear running time, paying special attention to the distribution of high-degree vertices inside $G(n, d/n)$ and to blend this with the entropy-decay tools developed in Reference [13], see Section 4.2 for this part of the argument.

In addition to our result for the hard-core model, we also obtain similar results for the Ising model and the Momomer-Dimer model. The configurations of the Ising model on a graph $G = (V, E)$ are assignments $\sigma \in \{0, 1\}^V$, which assign the spins 0 and 1 to the vertices of G . The Ising model with parameter $\beta > 0$ corresponds to a distribution $\mu_{G, \beta}$ on $\{0, 1\}^V$, where for an assignment $\sigma \in \{0, 1\}^V$, it holds that $\mu_{G, \beta}(\sigma) = \beta^{m(\sigma)} / Z_{G, \beta}$ where $m(\sigma)$ is the number of edges whose endpoints have the same spin assignment under σ , and $Z_{G, \beta}$ is the partition function of the model. The model is antiferromagnetic when $\beta \in (0, 1)$, and ferromagnetic otherwise. For $d \geq 1$, let

$\beta_c(d) = \frac{d-1}{d+1}$; for $d \in (0, 1)$, let $\beta_c(d) = 0$. It is known that on bounded-degree graphs of maximum degree $d+1$ the sampling/counting problem for the antiferromagnetic Ising model undergoes a phase transition at $\beta = \beta_c(d)$, analogous to that for the hard-core model [22, 29, 36, 39].

THEOREM 1.2. *Let $d, \beta > 0$ be such that $\beta \in (\beta_c(d), 1)$. For any constant $\theta > 0$, there is an algorithm such that, whp over the choice of $G \sim G(n, d/n)$, when the algorithm is given as input the graph G and an arbitrary rational $\varepsilon > 0$, it outputs an ε -sample from $\mu_{G, \beta}$ in time $n^{1+\theta} \log \frac{1}{\varepsilon}$.*

For a graph $G = (V, E)$, the monomer-dimer model with parameter $\gamma > 0$ corresponds to a distribution $\mu_{G, \gamma}$ on the set of matchings of G , where for a matching M , it holds that $\mu_{G, \gamma}(M) = \gamma^{|M|} / Z_{G, \gamma}$ where $Z_{G, \gamma}$ is the partition function. For general graphs $G = (V, E)$ and $\gamma = O(1)$, [25, 26] gave an $O(n^2 m \log n)$ algorithm (where $n = |V|, m = |E|$), which was improved for bounded-degree graphs to $O(n \log n)$ in Reference [13] using spectral independence. For $G(n, d/n)$, [35] gave an $O(n^{\log d})$ deterministic algorithm using correlation decay, and Reference [23] showed that Glauber dynamics mixes in $n^{2+o(1)}$ steps when $\gamma = 1$.

THEOREM 1.3. *Let $d, \gamma > 0$. For any constant $\theta > 0$, there is an algorithm such that, whp over the choice of $G \sim G(n, d/n)$, when the algorithm is given as input the graph G and an arbitrary rational $\varepsilon > 0$, it outputs an ε -sample from $\mu_{G, \gamma}$ in time $n^{1+\theta} \log \frac{1}{\varepsilon}$.*

In the next section, we give the main ingredients of our algorithm for the hard-core model, and we give the proof of Theorem 1.1. The proofs of Theorems 1.2 and 1.3 build on similar ideas, though there are some modifications needed to obtain the required spectral independence bounds. We give their proofs in Section 5.3.

Before proceeding let us finally mention that, to go beyond the two-spin models studied here, the main obstacle is to establish the spectral independence bounds for graphs with potentially unbounded degrees. As it is pointed out in Reference [35, Section 7], their correlation-decay analysis does not extend to other models in a straightforward manner, and hence it is natural to expect that the same is true for spectral independence as well.

1.1 Further Developments

Our algorithms are based on running Glauber dynamics on (relatively) low-degree vertices. Subsequent to our work, Eftymiou and Feng [17] obtained for the hard-core model an $n^{1+O(1/\log \log n)}$ mixing-time bound for Glauber dynamics on $G(n, d/n)$ when $\lambda < \lambda_c(d)$ (and similarly for the monomer-dimer model), i.e., without the need to restrict to low-degree vertices. Their spectral independence arguments build upon the “linearisation” of the L^p analysis we introduce here, which are then combined with the framework of Reference [10] to obtain the improved mixing-time bounds.

2 PROOF OUTLINE FOR THEOREM 1.1

Our algorithm for sampling from the hard-core model on a graph $G = (V, E)$ is an adaptation of Glauber dynamics on an appropriate set of “small-degree” vertices U , the details of the algorithm are given in Figure 1. Henceforth, analogously to the Ising model, it will be convenient to view the hard-core model as a two-spin model supported on $\Omega \subseteq \{0, 1\}^V$, where Ω corresponds to the set of independent sets of G (for an independent set I , we obtain $\sigma \in \{0, 1\}^V$ by setting $\sigma_v = 1$ iff $v \in I$). Note that for general graphs G , implementing Steps 2 and Steps 3 of the algorithm might be difficult. The following lemma exploits the sparse structure of $G(n, d/n)$ and in particular the fact that high-degree vertices are sparsely scattered. We will use this in the proof of our main theorems to show that the algorithm SAMPLE(G, T) can be implemented very efficiently for appropriate D , paying only $O(\log n)$ per loop operation in Step 2 and only $O(n \log n)$ in Step 3. The *tree-excess* of a graph $G = (V, E)$ is defined as $|E| - |V| + 1$.

Algorithm SAMPLE(G, T)

Parameters: $D > 0$ (threshold for small/high degree vertices).

Input: Graph $G = (V, E)$, integer $T \geq 1$ (number of iterations).

1. Initialisation: Let U be the set of all vertices with degree $\leq D$.

Let X_0 be the empty independent set on U .

2. Main loop: For $t = 1, \dots, T$,

- Pick a vertex u uniformly at random from U .

- For every vertex $v \in U \setminus \{u\}$, set $X_t(v) = X_{t-1}(v)$.

- Sample the spin $X_t(u)$ according to $\mu_{G,\lambda}(\sigma_u \mid \sigma_{U \setminus \{u\}} = X_t(U \setminus \{u\}))$, i.e., update u according to the hard-core distribution on the *whole graph* G , conditioned on the spins of $U \setminus \{u\}$.

3. Finalisation: Sample $\sigma \sim \mu_{G,\lambda}(\cdot \mid \sigma_U = X_T)$, i.e., extend X_T to the whole vertex set of G by sampling from $\mu_{G,\lambda}$ conditioned on the configuration on U .

Fig. 1. The SAMPLE(G, T) subroutine for sampling from the hard-core distribution $\mu_{G,\lambda}$. We use the analogue of this algorithm for the Ising model with parameter β (replacing $\mu_{G,\lambda}$ by $\mu_{G,\beta}$). For the monomer-dimer model, the only difference is that the algorithm needs to update (single) edges in F , where F is the set of edges whose both endpoints lie in U (i.e., degree $\leq D$).

LEMMA 2.1. *Let $d > 0$ be an arbitrary real. There exist constants $D, \ell > 0$ such that the following holds whp over the choice of $G = (V, E) \sim G(n, d/n)$. Each of the connected components of $G[V \setminus U]$, where U is the set of vertices of degree $\leq D$, has size $O(\log n)$ and tree-excess at most ℓ .*

Lemma 2.1 follows using relatively standard techniques from random graphs and is proved in Section 6. Later, we will establish a more refined version of this property that will allow us to bound the mixing time of the single-site dynamics that we consider (the main loop of SAMPLE(G, T)).

The key ingredient needed to prove our main result is to show that the main loop of our sampling algorithm returns a good sample on the induced hard-core distribution on the set U . More precisely, for a graph $G = (V, E)$ and $U \subseteq V$, we let $\mu_{G,\lambda,U}(\cdot)$ denote the induced distribution on the spins of U , i.e., the marginal distribution $\mu_{G,\lambda}(\sigma_U = \cdot)$.

LEMMA 2.2. *Let $d, \lambda > 0$ be constants such that $\lambda < \lambda_c(d)$. For any arbitrarily small constant $\theta > 0$, there is $D > 0$ such that the following holds whp over the choice of $G \sim G(n, d/n)$.*

Let U be the set of vertices in G of degree $\leq D$. Then, for any $\varepsilon > 0$, for $T = \lceil n^{1+\theta/2} \log \frac{1}{\varepsilon} \rceil$, the main loop of SAMPLE(G, T) returns a sample X_T from a distribution that is ε -close to $\mu_{G,\lambda,U}$.

We will prove Lemma 2.2 in Section 4.2. We are ready to prove Theorem 1.1.

PROOF OF THEOREM 1.1. We give first the details for the more interesting case $d \geq 1$. Consider arbitrarily small $\theta > 0$ and D, ℓ as in Lemmas 2.1 and 2.2, so that whp G satisfies the properties therein. (It is clear from the proofs of Lemmas 2.1 and 2.2 that D and ℓ can be computed given d, λ , and θ .) Let $\varepsilon > 0$ be the desired accuracy for sampling from $\mu_{G,\lambda}$; it is sufficient to consider $\varepsilon < 1/e$. Let U be the set of vertices with degree $\leq D$, and set $T = \lceil n^{1+\theta/2} \log \frac{1}{\varepsilon} \rceil$.

By Lemma 2.2, whp over the choice of G , the main loop of SAMPLE(G, T) returns a configuration $X_T : U \rightarrow \{0, 1\}$ that is ε -close to $\mu_{G,\lambda,U}$. Note that each iteration of the main loop of SAMPLE(G, T) can be implemented in $O(\log n)$ time, since $G[V \setminus U]$ has components of size $O(\log n)$ and tree excess at most ℓ . In particular, any vertex $u \in U$ can be adjacent to at most D of these components, and therefore the component of u in $G[(V \setminus U) \cup \{u\}]$ has size $O(\log n)$ and tree excess at most

$k = D[\ell] = O(1)$. We can therefore sample the spin of u under $\mu_{G,\lambda}$ conditioned on the spins of $U \setminus \{u\}$ in time $O(4^k \log n) = O(\log n)$.² Therefore, the main loop of $\text{SAMPLE}(G, T)$ runs in time $O(T \log n)$. Analogously, the finalisation step of $\text{SAMPLE}(G, T)$, i.e., extending the configuration X_T on U to a configuration σ on the whole vertex set V , can be implemented in time $O(n \log n)$ by iterating over the vertices in $V \setminus U$ and using the fact that the components of $G[V \setminus U]$ have excess at most ℓ . Therefore, the overall running time of the algorithm is bounded by $O(T \log n) + O(n \log n)$, which is less than $\lceil n^{1+\theta} \log \frac{1}{\epsilon} \rceil$ for all sufficiently large n . It remains to note that, since X_T is ϵ -close to the marginal distribution of $\mu_{G,\lambda}$ on U , and the finalisation step is done perfectly conditioned on the configuration on U , the final configuration σ is ϵ -close to the distribution $\mu_{G,\lambda}$.

For $d < 1$, whp G consists of treelike components of size $O(\log n)$, and therefore we can obtain a perfect sample from $\mu_{G,\lambda}$ in time $O(n \log n)$ by going through the vertices one by one and, for each vertex, taking $O(\log n)$ time to compute its marginal, conditioned on the spins already sampled. \square

3 SPECTRAL INDEPENDENCE VIA BRANCHING VALUES

We first introduce the notions of spectral independence and pairwise vertex influences, which we will later use to bound the mixing time of the main loop of $\text{SAMPLE}(G, T)$, i.e., to prove Lemma 2.2. We will define the terminology in a general way that will be useful both for our analysis of the hard-core model, and for our later analysis of other models.

Let $q \geq 2$ be an integer indicating the number of spins and let V be a set of size n . We consider distributions μ supported on a set $\Omega \subseteq [q]^V$.³ For $S \subseteq V$, let $\Omega_S = \{\tau \in [q]^S \mid \mu(\sigma_S = \tau) > 0\}$ be the set of all partial configurations on $[q]^S$ that have non-zero marginal under μ . For $\tau \in \Omega_S$, let μ_τ be the conditional distribution on Ω induced by τ , i.e., $\mu_\tau(\cdot) = \mu(\cdot \mid \sigma_S = \tau)$. Let $\mu_{\min} = \min_{\sigma \in \Omega} \mu(\sigma)$.

For $S \subseteq V$ and $\tau \in \Omega_S$, the *influence matrix* conditioned on τ is the matrix Ψ_τ whose rows and columns are indexed by $\tilde{V}_\tau = \{(v, i) \mid v \in V \setminus S, \mu_\tau(\sigma_v = i) > 0\}$, where the entry indexed by $(v, i), (w, k)$ equals $\mu_\tau(\sigma_w = k \mid \sigma_v = i) - \mu_\tau(\sigma_w = k)$ if $v \neq w$, and 0 otherwise. It is a standard fact that the eigenvalues of the matrix Ψ are all real [3], and we denote by $\lambda_1(\Psi)$ its largest eigenvalue.

Definition 3.1. Let $q \geq 2$ be an integer and V be a set of size $n \geq 2$. Let μ be a distribution supported over $\Omega \subseteq [q]^V$. Let $\eta, b > 0$. We say that μ is η -spectrally independent if for all $S \subset V$ and $\tau \in \Omega_S$, it holds that $\lambda_1(\Psi_\tau) \leq \eta$. We say that μ is b -marginally bounded if for all $S \subset V$, $v \in V \setminus S$, $\tau \in \Omega_S$, and $i \in [q]$, it either holds that $\mu_\tau(\sigma_v = i) = 0$ or else $\mu_\tau(\sigma_v = i) \geq b$.

Following References [3, 14], for distributions μ induced by two-spin systems, we work with the following notion of pairwise vertex-influence, which can be used to bound the spectral independence. For a graph $G = (V, E)$ and $\tau \in \{0, 1\}^S$ for some $S \subset V$, for vertices u, v with $u \in V \setminus S$ and $0 < \mu_\tau(\sigma_u = 1) < 1$, we define the *influence* of u on v (under μ_τ) as

$$\mathcal{I}_G^\tau(u \rightarrow v) = \mu_\tau(\sigma_v = 1 \mid \sigma_u = 1) - \mu_\tau(\sigma_v = 1 \mid \sigma_u = 0).$$

For matchings, we will work with an analogous notion from the perspective of edges (see Section 5.2). For all these models, spectral independence will be bounded by summing the absolute value of the influences of an arbitrary vertex u to the rest of the graph. In turn, it has been shown in Reference [14] that summing the influences of a vertex u in a graph G reduces to summing the sum of influences on the self-avoiding walk tree emanating from u (this is the tree whose vertices correspond to self-avoiding walks in G that start from u , where two vertices are adjacent when one

²One “naive” way to do this is by considering a spanning tree and then brute-forcing over all $\leq 4^k$ possibilities for the endpoints of the excess edges (the spins on each edge can be set in at most four ways). For each of these, the marginal probability at u and the corresponding partition function can be computed using dynamic programming on the left-over tree.

³For an integer $k \geq 1$, we denote by $[k]$ the set $\{0, 1, \dots, k - 1\}$.

walk is a one-step extension of the other; it was first introduced in related contexts by References [34, 40]).

We will use the following lemma from Reference [14].

LEMMA 3.2 ([14]). *Consider an arbitrary two-spin system on a graph $G = (V, E)$, with distribution μ supported on $\Omega \subseteq \{0, 1\}^V$. Let $\rho \in V$ be an arbitrary vertex, $T = (V_T, E_T)$ be the tree of self-avoiding walks in G starting from ρ , and v be the distribution of the two-spin system on T .*

Then, for any $S \subseteq V \setminus \{\rho\}$ and $\tau \in \Omega_S$ with $0 < \mu_\tau(\sigma_\rho = 1) < 1$, there is a subset $W \subseteq V_T \setminus \{\rho\}$ and a configuration $\phi \in \{0, 1\}^W$ such that $\mu_\tau(\sigma_\rho = 1) = v_\phi(\sigma_\rho = 1)$ and

$$\sum_{v \in V} |\mathcal{I}_G^\tau(\rho \rightarrow v)| \leq \sum_{u \in V_T} |\mathcal{I}_T^\phi(\rho \rightarrow u)|,$$

where $\mathcal{I}_T^\phi(\rho \rightarrow \cdot)$ denotes the influence of ρ on the vertices of T under v_ϕ .

3.1 The branching value

We will need the following notion to capture the growth of the self-avoiding walk tree from a vertex.

Definition 3.3. Let $d > 0$ be a real number and $G = (V, E)$ be a graph. For a vertex v in G , the d -branching value S_v equals $\sum_{\ell \geq 0} N_{v, \ell} / d^\ell$, where $N_{v, \ell}$ is the number of (simple) paths with a total of $\ell + 1$ vertices starting from v (for convenience, we set $N_{v, 0} = 1$).

We will show the following lemma in Section 6.1, which bounds the d' -branching value of $G(n, d/n)$ for any $d' > d$.

LEMMA 3.4. *Let $d \geq 1$. Then, for every $d' > d$ and $\varepsilon > 0$, whp over the choice of $G \sim G(n, d/n)$, the d' -branching value of every vertex in G is at most $\varepsilon \log n$.*

3.2 Spectral Independence for the Hard-Core Model

In this section, we bound the spectral independence of $G(n, d/n)$ in the hard-core model when $\lambda < \lambda_c(d)$. We will need the following technical lemma that can be derived from Reference [35]. The derivation details are similar to an analogous lemma for matchings (cf. Lemma 5.4 below), which can be found in Reference [5, Lemma 15].

LEMMA 3.5 ([35]). *Let $d > 1$ and $\lambda > 0$ be constants such that $\lambda < \lambda_c(d)$. Let $\chi \in (1, 2)$ be given from $\frac{1}{\chi} = 1 - \frac{d-1}{2} \log(1 + \frac{1}{d-1})$ and set $a = \frac{\chi}{\chi-1}$. Consider also the function $\Phi(y) = \frac{1}{\sqrt{y(1+y)}}$ for $y > 0$. Then, there is a constant $0 < \kappa < 1/d$ such that the following holds for any integer $k \geq 1$.*

Let $x_1, \dots, x_k > 0$ be reals and $x = \lambda \prod_{i=1}^k \frac{1}{1+x_i}$. Then $(\Phi(x))^a \sum_{i=1}^k \left(\frac{x}{(1+x_i)\Phi(x_i)}\right)^a \leq \kappa^{a/\chi}$.

We will show the following.

LEMMA 3.6. *Let $d > 1$ and $\lambda > 0$ be constants such that $\lambda < \lambda_c(d)$. Then, there is a constant $\chi > 1$ such that the following holds.*

Let $T = (V, E)$ be a tree rooted at ρ , whose d -branching value is $\leq \alpha$ and whose root has k children. Then, for the hard-core distribution on T with parameter λ , any $S \subseteq V \setminus \{\rho\}$ and $\tau \in \Omega_S$ with $0 < \mu_\tau(\sigma_\rho = 1) < 1$, it holds that

$$\sum_{v \in V} |\mathcal{I}_T^\tau(\rho \rightarrow v)| \leq W_k \alpha^{1/\chi},$$

where $W_k > 0$ is a real depending only on the degree k of the root (and the constants d, λ).

PROOF. Let $\kappa \in (0, 1/d)$ and $\chi \in (1, 2)$ be the constants from Lemma 3.5, and $\Phi(x) = \frac{1}{\sqrt{x(1+x)}}$ be also as in Lemma 3.5. We may assume without loss of generality that S is empty (and τ is trivial) by truncating the tree T using the following procedure: just remove vertices $u \in S$ with $\tau_u = 0$, and for $u \in S$ with $\tau_u = 1$ remove u and all of its neighbours. Note that for all the removed vertices v it holds that $\mathcal{I}_T^\tau(\rho \rightarrow v) = 0$, so the removal procedure does not decrease the sum of the absolute influences, while at the same time decreasing the d -branching value of the tree T . Henceforth, we will drop τ and S from notation.

To prove the lemma, we will work inductively on the depth of the tree. To this end, we first define for each vertex u in T the following values α_u and R_u ; the α 's capture a rooted analogue of the branching value of internal vertices within T , while the R 's the marginals of the vertices in the corresponding subtrees. More precisely, if v is a leaf, then set $\alpha_v = 1$ and $R_v = \lambda$; otherwise, set $\alpha_v = 1 + \frac{1}{d} \sum_{i=1}^t \alpha_{v_i}$ and $R_v = \lambda \prod_{i=1}^t \frac{1}{1+R_{v_i}}$, where v_1, \dots, v_t are the children of v . Note that for the root ρ , we have that $\alpha_\rho = S_\rho \leq \alpha$, where S_ρ is the d -branching value of ρ in the tree T . Moreover, if we denote by T_v the subtree of T rooted at v and by u the parent of v in T , then it holds that

$$R_v = \frac{\mu_{T_v, \lambda}(\sigma_v = 1)}{\mu_{T_v, \lambda}(\sigma_v = 0)} \quad \text{and} \quad \mathcal{I}_T(u \rightarrow v) = -\frac{R_v}{R_v + 1}. \quad (1)$$

The first equality is fairly standard and can be proved using induction on the height of the tree, while the second one is Reference [14, Lemma 15] (it also follows directly from the definition of influence and the first equality).

For an integer $h \geq 0$, let $L(h)$ be the nodes at distance h from the root ρ . Let $M_k = \sqrt{1 + (1 + \lambda)^k}/\lambda$, where recall that k is the degree of the root ρ . We will show that

$$\sum_{v \in L(h)} \left(\frac{\alpha_v}{\alpha_\rho} \right)^{1/\chi} \frac{|\mathcal{I}_T(\rho \rightarrow v)|}{R_v \Phi(R_v)} \leq M_k (d\kappa)^{h/\chi}. \quad (2)$$

Since $\alpha_v \geq 1$ for $v \in V$, $\alpha_\rho \leq \alpha$ and $R_v \Phi(R_v) \leq 1$, Equation (2) yields $\sum_{v \in L(h)} |\mathcal{I}_T(\rho \rightarrow v)| \leq M_k \alpha^{1/\chi} (d\kappa)^{h/\chi}$ for all integer $h \geq 0$, and therefore summing over h , we obtain that

$$\sum_{v \in V} |\mathcal{I}_T(\rho \rightarrow v)| \leq M_k \alpha^{1/\chi} \sum_{h \geq 0} (d\kappa)^{h/\chi} \leq \frac{M_k \alpha^{1/\chi}}{1 - (d\kappa)^{1/\chi}},$$

which proves the result with $W_k = \frac{M_k}{1 - (d\kappa)^{1/\chi}}$. So it only remains to prove (2).

We will work inductively. The base case $h = 0$ is equivalent to $M_k \geq 1/(R_\rho \Phi(R_\rho)) = \sqrt{1 + 1/R_\rho}$, which is true, since from the recursion Equation (1) for R_ρ , we have that $R_\rho \geq \lambda/(1 + \lambda)^k$ (using the trivial bound $R_v \leq \lambda$ for each v). For the induction step, consider $v \in L(h-1)$ and suppose it has $k_v \geq 0$ children, denoted by v_i for $i \in [k_v]$. Then, for each $i \in [k_v]$, since v is on the unique path joining ρ to v_i , it holds that (see Reference [3, Lemma B.2])

$$\mathcal{I}_T(\rho \rightarrow v_i) = \mathcal{I}_T(\rho \rightarrow v) \mathcal{I}_T(v \rightarrow v_i),$$

so we can write

$$\sum_{v \in L(h)} \left(\frac{\alpha_v}{\alpha_\rho} \right)^{1/\chi} \frac{|\mathcal{I}_T(\rho \rightarrow v)|}{R_v \Phi(R_v)} = \sum_{v \in L(h-1)} \left(\frac{\alpha_v}{\alpha_\rho} \right)^{1/\chi} \frac{|\mathcal{I}_T(\rho \rightarrow v)|}{R_v \Phi(R_v)} \sum_{i \in [k_v]} \left(\frac{\alpha_{v_i}}{\alpha_v} \right)^{1/\chi} R_v \Phi(R_v) \frac{|\mathcal{I}_T(v \rightarrow v_i)|}{R_{v_i} \Phi(R_{v_i})}. \quad (3)$$

Consider an arbitrary $v \in L(h-1)$. Then, since $\frac{1}{\chi} + \frac{1}{a} = 1$, by Hölder's inequality we have that

$$\sum_{i \in [k_v]} \left(\frac{\alpha_{v_i}}{\alpha_v} \right)^{1/\chi} R_v \Phi(R_v) \frac{|\mathcal{I}_T(v \rightarrow v_i)|}{R_{v_i} \Phi(R_{v_i})} \leq \left(\sum_{i \in [k_v]} \frac{\alpha_{v_i}}{\alpha_v} \right)^{1/\chi} \left((R_v \Phi(R_v))^a \sum_{i \in [k_v]} \left(\frac{|\mathcal{I}_T(v \rightarrow v_i)|}{R_{v_i} \Phi(R_{v_i})} \right)^a \right)^{1/a}. \quad (4)$$

Note that for $x = R_v$ and $x_i = R_{v_i}$, $i \in [k_v]$, we have from Equation (1) that $\frac{|\mathcal{I}_T(v \rightarrow v_i)|}{R_{v_i}} = \frac{1}{1+x_i}$ and $x = \lambda \prod_{i \in [k_v]} \frac{1}{1+x_i}$, so by Lemma 3.5, we have that

$$\left((R_v \Phi(R_v))^a \sum_{i \in [k_v]} \left(\frac{|\mathcal{I}_T(v \rightarrow v_i)|}{R_{v_i} \Phi(R_{v_i})} \right)^a \right)^{1/a} \leq \kappa^{1/\chi}.$$

By definition of the d -branching value, we also have $\alpha_v = 1 + \frac{1}{d} \sum_{i \in [k_v]} \alpha_{v_i} \geq \frac{1}{d} \sum_{i \in [k_v]} \alpha_{v_i}$, so plugging these back into Equation (4) yields

$$\sum_{i \in [k_v]} \left(\frac{\alpha_{v_i}}{\alpha_v} \right)^{1/\chi} R_v \Phi(R_v) \frac{|\mathcal{I}_T(v \rightarrow v_i)|}{R_{v_i} \Phi(R_{v_i})} \leq (d\kappa)^{1/\chi}.$$

In turn, plugging this into Equation (3) and using the induction hypothesis yields Equation (2), finishing the proof. \square

Remark 3.7. For simplicity, and since it is not important for our arguments, the constant W_k in the proof depends exponentially on the degree k of the root. With a more careful inductive proof (cf. Reference [14, Proof of Lemma 14]), the dependence on k can be made linear. In either case, because of the high-degree vertices in $G(n, d/n)$, both bounds do not yield sufficiently strong bounds on the spectral independence of the whole distribution $\mu_{G, \lambda}$, and this is one of the reasons that we have to consider the spectral independence on the induced distribution on low-degree vertices.

Recall that for a graph $G = (V, E)$ and $U \subseteq V$, we let $\mu_{G, \lambda, U}(\cdot)$ denote the marginal distribution on the spins of U , i.e., the distribution $\mu_{G, \lambda}(\sigma_U = \cdot)$.

LEMMA 3.8. *Let $d \geq 1$ and $\lambda > 0$ be constants such that $\lambda < \lambda_c(d)$. Then, for any constants $D, \varepsilon > 0$, whp over the choice of $G \sim G(n, d/n)$, the marginal hard-core distribution $\mu_{G, \lambda, U}$, where U is the set of vertices in G with degree $\leq D$, is $(\varepsilon \log n)$ -spectrally independent.*

PROOF. Let $D, \varepsilon > 0$ be arbitrary constants, and let $d' > d$ be such that $\lambda < \lambda_c(d')$; such d' exists, because the function $\lambda_c(\cdot)$ is continuous in the interval $(1, \infty)$ and $\lambda_c(d) \rightarrow \infty$ for $d \downarrow 1$. Let $\chi \in (1, 2)$ and $W = \max\{W_1, \dots, W_D\}$ where χ and the W_k 's are as in Lemma 3.6 (corresponding to the constants d', λ). By Lemma 3.4, whp all of the vertices of the graph $G = (V, E) \sim G(n, d/n)$ have d' -branching value less than $\varepsilon \log n$. We will show that the result holds for all such graphs G (for sufficiently large n).

Let U be the set of vertices in G with degree $\leq D$, and let for convenience $\mu = \mu_{G, \lambda, U}$. Consider arbitrary $S \subset U$ and $\tau \in \Omega_S$. It suffices to bound the largest eigenvalue of the influence matrix Ψ_τ by $\varepsilon \log n$. Analogously to References [3, 14], we do this by bounding the absolute-value row sums of Ψ_τ . Recall that the rows and columns of Ψ_τ are indexed by $\tilde{V}_\tau = \{(v, i) \mid v \in U \setminus S, \mu_\tau(\sigma_v = i) > 0\}$, where the entry indexed by $(v, i), (w, k)$ equals $\mu_\tau(\sigma_w = k \mid \sigma_v = i) - \mu_\tau(\sigma_w = k)$ if $v \neq w$, and 0 otherwise. Consider arbitrary $(v, i) \in \tilde{V}_\tau$; our goal is to show

$$\sum_{(w, k) \in \tilde{V}_\tau} \left| \mu_\tau(\sigma_w = k \mid \sigma_v = i) - \mu_\tau(\sigma_w = k) \right| \leq \varepsilon \log n. \quad (5)$$

Henceforth, we will also assume that $\mu_\tau(\sigma_v = i) < 1$ (in addition to $\mu_\tau(\sigma_v = i) > 0$); otherwise, the sum on the left-hand side is equal to 0. Then, by the law of total probability, for any $(w, k) \in \tilde{V}_\tau$, we have

$$|\mu_\tau(\sigma_w = k \mid \sigma_v = i) - \mu_\tau(\sigma_w = k)| \leq |\mu_\tau(\sigma_w = k \mid \sigma_v = 1) - \mu_\tau(\sigma_w = k \mid \sigma_v = 0)| = |\mathcal{I}_G^\tau(v \rightarrow w)|,$$

where the last equality follows from the fact that μ is the marginal distribution of $\mu_{G, \lambda}$ on U . Therefore, we can bound

$$\sum_{(w, k) \in \tilde{V}_\tau} |\mu_\tau(\sigma_w = k \mid \sigma_v = i) - \mu_\tau(\sigma_w = k)| \leq 2 \sum_{w \in U} |\mathcal{I}_G^\tau(v \rightarrow w)| \leq 2 \sum_{w \in V} |\mathcal{I}_G^\tau(v \rightarrow w)|.$$

By Lemma 3.2, for the self-avoiding walk tree $T = (V_T, E_T)$ from v , there is a subset $Z \subseteq V_T \setminus \{\rho\}$ and a configuration $\phi \in \{0, 1\}^Z$ such that

$$2 \sum_{w \in V} |\mathcal{I}_G^\tau(v \rightarrow w)| \leq 2 \sum_{w \in V_T} |\mathcal{I}_T^\phi(v \rightarrow w)|,$$

where $\mathcal{I}_T^\phi(v \rightarrow \cdot)$ denotes the influence of v on the vertices of T (in the hard-core distribution $\mu_{T, \lambda}$ conditioned on ϕ). Since the d' -branching value of v (and any other vertex of G) is bounded by $\varepsilon \log n$ and the degree of v is $\leq D$, by Lemma 3.6 applied to T , we have that

$$2 \sum_{w \in V_T} |\mathcal{I}_T^\phi(v \rightarrow w)| \leq 2W(\varepsilon \log n)^{1/\chi}.$$

Since $\chi > 1$, for all sufficiently large n , we have that $2W(\varepsilon \log n)^{1/\chi} \leq \varepsilon \log n$, which proves Equation (5). \square

We also record the following corollary of the arguments in Lemma 3.6.

COROLLARY 3.9. *Let $\lambda > 0$ and $D > 0$ be real numbers. For a graph $G = (V, E)$, let U be the set of vertices in G with degree $\leq D$ and suppose that $|U| \geq 2$. Then, the distribution $\mu := \mu_{G, \lambda, U}$ is b -marginally bounded for $b = \min\{\frac{1}{1+\lambda}, \frac{\lambda}{\lambda+(1+\lambda)^D}\}$.*

PROOF. By Lemma 3.2, for any vertex $v \in U$ and any boundary condition τ on (a subset of) $U \setminus \{v\}$, there is a corresponding tree T and a boundary condition ϕ on T such that $\mu_\tau(\sigma_v = \cdot) = \nu_\phi(\sigma_v = \cdot)$. Since v has degree $\leq D$, from the proof of Lemma 3.6, see in particular Equation (1), we have that $\nu_\phi(\sigma_v = \cdot) \geq b$, where b is as in the lemma statement. \square

4 ENTROPY FACTORISATION FOR BOUNDED-DEGREE VERTICES

In this section, we show how to convert the spectral independence results of the previous section into fast mixing results for Glauber dynamics on the set of small-degree vertices on $G(n, d/n)$. Our strategy here follows the technique of Reference [13], though to obtain nearly linear results, we have to pay attention to the connected components induced by high-degree vertices and how these can connect up small-degree vertices.

4.1 Preliminaries

Entropy factorisation for probability distributions. For a real function f on $\Omega \subseteq [q]^V$, we use $E_\mu(f)$ for the expectation of f with respect to μ and, for $f : \Omega \rightarrow \mathbb{R}_{\geq 0}$, $\text{Ent}_\mu(f) = E_\mu[f \log f] - E_\mu(f) \log E_\mu(f)$, with the convention that $0 \log 0 = 0$. Finally, for $S \subset V$, let $\text{Ent}_\mu^S(f) = E_{\tau \sim \mu_{V \setminus S}}[\text{Ent}_{\mu_\tau}(f)]$ i.e., $\text{Ent}_\mu^S(f)$ is the expected value of the conditional entropy of f when the assignment outside of S is chosen according to the marginal distribution $\mu_{V \setminus S}$ (the induced distribution of μ on $V \setminus S$). For convenience, when $S = V$, we define $\text{Ent}_\mu^S(f) = \text{Ent}_\mu(f)$. The following inequality of entropy under tensor product is a special case of Shearer's inequalities.

FACT 4.1. Let $q, k \geq 2$ be integers and suppose that, for $i \in [k]$, μ_i is a distribution supported over $\Omega_i \subseteq [q]^{V_i}$, where V_1, \dots, V_k are pairwise disjoint sets. Let $\mu = \mu_1 \otimes \dots \otimes \mu_k$ be the product distribution on $\Omega = \Omega_1 \times \dots \times \Omega_k$. Then, for any $f : \Omega \rightarrow \mathbb{R}_{\geq 0}$, it holds that $\text{Ent}_\mu(f) \leq \sum_{i=1}^k \text{Ent}_\mu^{V_i}(f)$.

To bound the mixing time of Markov chains such as the Glauber dynamics, we will be interested in establishing inequalities for factorisation of entropy, defined as follows (see Reference [8] for more details).

Definition 4.2. Let $q \geq 2, r \geq 1$ be integers and V be a set of size $n \geq r + 1$. Let μ be a distribution supported over $\Omega \subseteq [q]^V$. We say that μ satisfies the r -uniform-block factorisation of entropy with multiplier⁴ C_r if for all $f : \Omega \rightarrow \mathbb{R}_{\geq 0}$ it holds that $\frac{r}{n} \text{Ent}_\mu(f) \leq C_r \frac{1}{\binom{n}{r}} \sum_{S \in \binom{V}{r}} \text{Ent}_\mu^S(f)$.

The following lemma will be useful to bound the (r -uniform-block) factorisation multiplier for conditional distributions on sets with small cardinality.

LEMMA 4.3 ([13, LEMMA 4.2]). Let $q \geq 2$ be an integer and V be a set of size $n \geq 2$. Let μ be a distribution supported over $\Omega \subseteq [q]^V$, which is b -marginally bounded for some $b > 0$. Then, for any $S \subseteq V$ and $\tau \in \Omega_{V \setminus S}$, for $f : \Omega \rightarrow \mathbb{R}_{\geq 0}$, it holds that $\text{Ent}_{\mu_\tau}(f) \leq \frac{2|S|^2 \log(1/b)}{b^{2|S|+2}} \sum_{v \in S} \text{Ent}_{\mu_\tau}^v(f)$.

The r -uniform-block dynamics and its mixing time. For an integer $r = 1, \dots, n$, the r -uniform-block dynamics for μ is a Markov chain $(X_t)_{t \geq 0}$ where $X_0 \in \Omega$ is an arbitrary configuration and, for $t \geq 1$, X_t is obtained from X_{t-1} by first picking a subset $S \in V$ of size r uniformly at random and updating the configuration on S according to $\mu(\sigma_S = \cdot \mid \sigma_{V \setminus S} = X_{t-1}(V \setminus S))$. Note, the case $r = 1$ corresponds to the single-site dynamics, where at every step the spin of a single vertex, chosen u.a.r., is updated conditioned on the spins of the remaining vertices. For $\varepsilon > 0$, the mixing time of the r -uniform-block dynamics is defined as $T_{\text{mix}}(\varepsilon) = \max_{\sigma \in \Omega} \min\{t \mid X_0 = \sigma, \|v_t - \mu\|_{\text{TV}} \leq \varepsilon\}$, where v_t denotes the distribution of X_t . The following lemma builds upon a well-known connection between factorisation of entropy and modified log-Sobolev inequalities (see, e.g., Reference [8] for more discussion), we will use the following version that can be extracted from recent works.

LEMMA 4.4 (SEE, E.G., REFERENCE [13, LEMMA 2.6 & FACT 3.5(4)] OR REFERENCE [12, LEMMA 3.2.6 & FACT 3.4.2]). Let $q \geq 2, r \geq 1$ be integers and V be a set of size $n \geq r + 1$. Let μ be a distribution supported over $\Omega \subseteq [q]^V$ that satisfies the r -uniform-block factorisation of entropy with multiplier C_r . Then, for any $\varepsilon > 0$, the mixing time of the r -uniform-block dynamics on μ satisfies

$$T_{\text{mix}}(\varepsilon) \leq \left\lceil C_r \frac{n}{r} \left(\log \log \frac{1}{\mu_{\min}} + \log \frac{1}{2\varepsilon^2} \right) \right\rceil, \text{ where } \mu_{\min} = \min_{\sigma \in \Omega} \mu(\sigma).$$

We remark that to deduce the lemma from Reference [13] or Reference [12], which refer to the so-called “entropy decay constant κ ,” one needs to use the equality $C_r \kappa = r/n$ from Reference [13, Lemma 2.6] or Reference [12, Lemma 3.2.6].

From spectral independence to r -uniform-block factorisation multipliers. The following theorem is shown in Reference [13]; while the version that we state here cannot be found verbatim in Reference [13], we explain in Appendix A how to combine the results therein to obtain it.

THEOREM 4.5 ([13]). Let $q \geq 2$ be an integer and V be a set of size $n \geq 2$. Let μ be a distribution supported over $\Omega \subseteq [q]^V$ that is η -spectrally independent and b -marginally bounded for $\eta, b > 0$.

⁴We note that in related works C_r is usually referred to as the “factorisation constant”; we deviate from this terminology, since for us C_r will depend on n (cf. Corollary 4.6 and Lemma 4.8), and referring to it as a constant could cause confusion.

Then, for all integers $r = 1, \dots, n$, the distribution μ satisfies the r -uniform-block factorisation of entropy with multiplier $C_r = \frac{r}{n} \frac{\sum_{k=0}^{n-1} \Gamma_k}{\sum_{k=n-r}^{n-1} \Gamma_k}$, where $\Gamma_k = \prod_{j=0}^{k-1} \alpha_j$ for $k \in [n]$,⁵ and $\alpha_k = \max\{0, 1 - \frac{4\eta}{b^2(n-1-k)}\}$ for $k \in [n-1]$.

4.2 Entropy Factorisation for Bounded-degree Vertices in the Hard-core Model

We begin by noting that our arguments in this section, while developed primarily in the context of the hard-core model, apply more generally and will be used in particular for our Ising and monomer-dimer results on $G(n, d/n)$ (cf. Section 5.3).

The first step of the analysis of Glauber dynamics for the hard-core model on the set of small-degree vertices will be to employ spectral independence results of Section 3.2 to conclude fast mixing for the r -uniform-block dynamics for $r = \theta|U|$ for any arbitrarily small constant θ . This step will follow by applying the recent technology of entropy factorisation described above.

The second step is the more challenging for us. Here, we need to conclude fast mixing for $r = 1$, and in particular prove that $C_1/C_r = n^{o(1)}$. This is done roughly by studying the connected components of G that arise when resampling an r -subset of the low-degree vertices; the factorisation multiplier of these components controls the ratio C_1/C_r . While this resembles the approach of Reference [13], there is a key difference here, in that high-degree vertices are not resampled. This can not only cause potentially large components but also imposes a deterministic lower bound on components sizes (since a component consisting of high-degree vertices will be deterministically present in the percolated graph consisting of the r -subset of low-degree vertices and all of the high-degree vertices). This lower bound on the component sizes is actually more significant than it might initially seem, since the relatively straightforward bound of $\Omega(\log n)$ would unfortunately give a relatively large factorisation multiplier of $n^{\Omega(1)}$ (through Lemma 4.3). Instead, we need to show that components have size $o(\log n)$, which in turn requires more delicate estimates for the distribution of high-degree vertices in connected sets (see Lemma 4.7 below).

We start with the following corollary of Lemma 3.8, which converts a spectral independence bound into a bound on the factorisation multiplier for the r -uniform-block dynamics when r scales linearly with small-degree vertices. This is analogous to Reference [13, Lemma 2.4], where they obtain a $2^{O(\eta/b^2)}$ bound on C_r when $r = \Theta(n)$ via Theorem 4.5 (where η is the spectral independence bound and b is the bound on the marginals). By restricting to small-degree vertices, we obtain that b is a constant, which combined with the bound $\eta = o(\log n)$ from Lemma 3.8 gives the bound $C_r = n^{o(1)}$, as detailed below (for clarity, we show the relevant lemmas using arbitrarily small constants instead of $o(1)$, see Remark 6.7 on how to extract $o(1)$). The proof of the corollary is given for completeness in Appendix A.

COROLLARY 4.6. *Let $d \geq 1$ and $\lambda > 0$ be constants such that $\lambda < \lambda_c(d)$. Then, for any constants $D, \theta > 0$, whp over the choice of $G \sim G(n, d/n)$, the marginal hard-core distribution $\mu_{G, \lambda, U}$, where U is the set of vertices in G with degree $\leq D$, satisfies for any integer $r \in [\theta|U|, |U|]$ the r -uniform-block factorisation of entropy with multiplier $C_r \leq n^\theta$.*

Note that the reason that we are able to use the same θ in the bounds for both r and C_r is that the bound on C_r is loose (we can obtain a sharper result, since we have a bound on the spectral independence of $\varepsilon \log n$ for any $\varepsilon > 0$).

We will now refine the bound of Corollary 4.6 down to $r = 1$ by exploiting the fact that high-degree vertices are sparsely scattered. In particular, we will need the following lemma, which is

⁵We note that for $k = 0$, the product defining Γ_k is empty and therefore $\Gamma_0 = 1$.

a refinement of Lemma 2.1. For a graph $G = (V, E)$, we say that a set $S \subseteq V$ is connected if the induced subgraph $G[S]$ is connected.

LEMMA 4.7. *Let $d > 0$ be an arbitrary real. There exists an $L > 0$ such that for any $\delta \in (0, 1)$, the following holds whp over the choice of $G = (V, E) \sim G(n, d/n)$. For $\Delta = 1/(\delta \log \frac{1}{\delta})$, for all integers $k \geq \delta \log n$ and any $v \in V$, there are $\leq (2e)^{\Delta L k}$ connected sets $S \subseteq V$ containing v with $|S| = k$. Moreover, every such set contains $\geq k/2$ vertices with degree less than Δk .*

The proof of Lemma 4.7 is given in Section 6.3. We are now ready to show the following.

LEMMA 4.8. *Let $d \geq 1$ and $\lambda > 0$ be constants such that $\lambda < \lambda_c(d)$. For any $\theta > 0$, there is a constant $D > 0$ such that whp over the choice of $G \sim G(n, d/n)$, the marginal hard-core distribution $\mu_{G, \lambda, U}$, where U is the set of vertices in G with degree $\leq D$, satisfies the 1-uniform-block factorisation of entropy with multiplier $C_1 \leq n^\theta$.*

PROOF. Let $L > 0$ be as in Lemma 4.7, and consider an arbitrarily small constant $\theta > 0$. Let $\delta \in (0, 1)$ be a sufficiently small constant so that for $D = L/(\delta \log \frac{1}{\delta})$ and $b = \min\{\frac{1}{1+\lambda}, \frac{\lambda}{\lambda+(1+\lambda)^D}\}$ it holds that $\frac{1}{b^{2\delta}} < e^{\theta/4}$; such a constant exists, since $b^{2\delta} \rightarrow 1$ as $\delta \downarrow 0$. Let $\Delta = 1/(\delta \log \frac{1}{\delta})$ and $\zeta > 0$ be a small constant so that $2(2e)^{\Delta L}(2\zeta)^{1/2} \leq b^2/2$.

Let U be the vertices in G with degree $\leq D$, and let $r = \lfloor \zeta |U| \rfloor + 1$. Let $\mu = \mu_{G, \lambda, U}$. By Corollary 4.6, we have that whp over the choice of G , there is $C_r \leq n^{\theta/3}$ such that for every $f : \Omega \rightarrow \mathbb{R}_{\geq 0}$ it holds that

$$\frac{r}{|U|} \text{Ent}_\mu(f) \leq C_r \frac{1}{\binom{|U|}{r}} \sum_{S \in \binom{U}{r}} \text{Ent}_\mu^S(f). \quad (6)$$

For $S \subseteq U$, let $C'(S)$ denote the collection of the connected components of the graph $G[S \cup (V \setminus U)]$, viewed as vertex sets, and let $C(S) = \bigcup_{R \in C'(S)} \{R \cap U\}$ be the restriction of these components to the set U . Note that, for $S \subseteq U$ and $\tau \in \Omega_{U \setminus S}$, μ_τ factorises over the components of $G[S \cup (V \setminus U)]$ and in particular, applying Fact 4.1, we have that

$$\text{Ent}_\mu^S(f) = \mathbf{E}_{\tau \sim \mu_{U \setminus S}} [\text{Ent}_{\mu_\tau}(f)] \leq \mathbf{E}_{\tau \sim \mu_{U \setminus S}} \left[\sum_{R \in C(S)} \text{Ent}_{\mu_\tau}^R(f) \right].$$

Using the bound in Lemma 4.3, we further obtain that

$$\text{Ent}_\mu^S(f) \leq \mathbf{E}_{\tau \sim \mu_{U \setminus S}} \left[\sum_{R \in C(S)} \frac{2|R|^2 \log(1/b)}{b^{2|R|+2}} \sum_{u \in R} \text{Ent}_{\mu_\tau}^u(f) \right] = \sum_{R \in C(S)} \sum_{u \in R} \frac{2|R|^2 \log(1/b)}{b^{2|R|+2}} \text{Ent}_\mu^u(f),$$

where the last equality follows by linearity of expectation and the fact that $\mathbf{E}_{\tau \sim \mu_{U \setminus S}} [\text{Ent}_{\mu_\tau}^u(f)] = \text{Ent}_\mu^u(f)$. Plugging this bound into Equation (6), we obtain that

$$\text{Ent}_\mu(f) \leq \frac{2C_r \log(1/b)}{b^2 \binom{|U|-1}{r-1}} \sum_{S \in \binom{U}{r}} \sum_{R \in C(S)} \sum_{u \in R} \frac{|R|^2}{b^{2|R|}} \text{Ent}_\mu^u(f),$$

which yields that

$$\text{Ent}_\mu(f) \leq \frac{2C_r \log(1/b)}{b^2} \sum_{u \in U} \text{Ent}_\mu^u(f) \sum_{k=1}^n \frac{k^2}{b^{2k}} \Pr[C_u(S) = k], \quad (7)$$

where $\Pr[C_u(S) = k]$ denotes the probability that u belongs to a set of size k in the set $C(S)$, when we pick S uniformly at random from $\{S \in \binom{U}{r} \mid u \in S\}$. Define analogously $\Pr[C'_u(S) = k]$ to be the probability that u belongs to a connected component of size k in the set $C'(S)$. By Lemma 4.7, whp over $G \sim G(n, d/n)$, for all vertices u and any integer $t \geq \delta \log n$, there are at most $(2e)^{L\Delta t}$

connected sets of size t containing a given vertex u , and each of them contains at least $t/2$ vertices from U . In particular, for any integer $k \geq \delta \log n$, it holds that $\Pr[C_u(S) = k] \leq \Pr[k \leq C'_u(S) \leq 2k]$. For all $k \leq 2|U|$, the probability that a specific subset of $k/2$ vertices of U is present in $G[S \cup (V \setminus U)]$ is at most $\frac{\binom{|U|-k/2}{r-k/2}}{\binom{|U|-1}{r-1}} \leq \left(\frac{r}{|U|}\right)^{k/2} \leq (2\zeta)^{k/2}$. Therefore, for all $k \geq \delta \log n$, by a union bound over the connected sets of size k , we have

$$\Pr[C'_u(S) = k] \leq (2e)^{L\Delta k} 2^k (2\zeta)^{k/2} = \left(2(2e)^{L\Delta} (2\zeta)^{1/2}\right)^k \leq (b^2/2)^k,$$

where in the first inequality the first factor is the number of size- k connected sets T of G containing u , the second factor is an upper bound on the number of size $k/2$ subsets W of U that might be included in T and the final factor is the probability that W is included in S . The last inequality is by the choice of ζ . It follows that

$$\Pr[C_u(S) = k] \leq \Pr[k \leq C'_u(S) \leq 2k] \leq 2k(b^2/2)^k.$$

From this bound and the inequality $1/b^{2\delta} < e^{\theta/4}$ by the choice of δ , we can split and bound the rightmost sum in Equation (7) by

$$\sum_{k=1}^n \frac{k^2}{b^{2k}} \Pr[C_u(S) = k] \leq \frac{(\delta \log n)^2}{b^{2\delta \log n}} + \sum_{k \geq \delta \log n} \frac{2k^3}{2^k} \leq n^{\theta/3},$$

where the last inequality holds for all sufficiently large n , since the first term is $O((\log n)^2 n^{\theta/4})$ and the second term is $O(1)$. In turn, plugging this into Equation (7), we obtain that μ satisfies the 1-uniform block factorisation of entropy with multiplier $C_1 = \frac{2C_r \log(1/b)}{b^2} n^{\theta/3} \leq n^\theta$ for all sufficiently large n (since b is a constant and $C_r \leq n^{\theta/3}$), as needed. \square

Lemma 2.2 now follows easily by combining Lemmas 4.4 and 4.8. This was the last ingredient needed in the proof of Theorem 1.1.

5 PROOFS FOR THE ISING AND MONOMER-DIMER MODELS

5.1 Antiferromagnetic Ising Model

In this section, we bound the spectral independence of $G(n, d/n)$ in the antiferromagnetic Ising model. We will need the following well-known lemma; the version we state here is from Reference [6, Lemma 40], but it traces back to Reference [30] based on a lemma from Reference [4] (that was originally stated for the ferromagnetic Ising model).

LEMMA 5.1 ([6, LEMMA 40], SEE ALSO REFERENCES [4, 30]). *Let $\beta \in (0, 1)$. Let $T = (V, E)$ be a tree, $S \subseteq V$ be a subset of the vertices and $v \in V$ be an arbitrary vertex. Let $\tau_1, \tau_2 \in \{0, 1\}^S$ be two configurations on S , which differ only a subset $W \subseteq S$. Then,*

$$\left| \mu_{T, \beta}(\sigma_v = 1 \mid \sigma_S = \tau_1) - \mu_{T, \beta}(\sigma_v = 1 \mid \sigma_S = \tau_2) \right| \leq \sum_{w \in W} \left(\frac{1 - \beta}{1 + \beta} \right)^{\text{dist}(v, w)},$$

where for a vertex $w \in W$, $\text{dist}(v, w)$ denotes the distance between v and w in T .

LEMMA 5.2. *Let $d \geq 1$ and $\beta > 0$ be constants such that $\beta \in (\beta_c(d), 1)$. For any constant $\varepsilon > 0$, whp over the choice of $G \sim G(n, d/n)$, the antiferromagnetic Ising distribution $\mu_{G, \beta}$ is $(\varepsilon \log n)$ -spectrally independent.*

PROOF. Let $d' > d$ be a real number such that $\beta > \beta_c(d')$, i.e., $\frac{1-\beta}{1+\beta} < \frac{1}{d'}$; such d' exists, because of the continuity of the function $\beta_c(\cdot)$ in the interval $(0, 1)$. By Lemma 3.4, whp all of the vertices

the graph $G = (V, E) \sim G(n, d/n)$ have d' -branching value less than $\varepsilon \log n$. We will show that the result holds for all such graphs G .

Let for convenience $\mu = \mu_{G, \beta}$. Consider arbitrary $S \subset V$ and $\tau \in \Omega_S$. We will once again bound the largest eigenvalue of the influence matrix Ψ_τ by bounding the absolute-value row sums of Ψ_τ . In particular, analogously to the proof of Lemma 3.8, it suffices to show that for $v \in V \setminus S$, we have

$$\sum_{w \in V} |\mathcal{I}_G^\tau(v \rightarrow w)| \leq \varepsilon \log n, \quad (8)$$

where $\mathcal{I}_G^\tau(v \rightarrow w)$ denotes the influence of v on the vertices of G conditioned on τ . By Lemma 3.2, for the self-avoiding walk tree $T = (V_T, E_T)$ from v , there is a subset $W \subseteq V_T \setminus \{v\}$ and a configuration $\phi \in \{0, 1\}^W$ such that

$$\sum_{w \in V} |\mathcal{I}_G^\tau(v \rightarrow w)| \leq \sum_{w \in V_T} |\mathcal{I}_T^\phi(v \rightarrow w)|, \quad (9)$$

where $\mathcal{I}_T^\phi(v \rightarrow \cdot)$ denotes the influence of v on the vertices of T (in the Ising distribution $\mu_{T, \lambda}$ conditioned on ϕ). Since the d' -branching value of v (and any other vertex of G) is bounded by $\varepsilon \log n$, and applying Lemma 5.1 to T , we have that

$$\sum_{w \in V_T} |\mathcal{I}_T^\phi(v \rightarrow w)| \leq \sum_{\ell \geq 0} N_{v, \ell} \left(\frac{1 - \beta}{1 + \beta} \right)^\ell \leq \sum_{\ell \geq 0} N_{v, \ell} / (d')^\ell \leq \varepsilon \log n,$$

where $N_{v, \ell}$ is the number of paths in G with a total of $\ell + 1$ vertices starting from v . Combining this with Equation (9) proves Equation (8), as wanted. \square

The following lemma is a crude (and relatively standard) bound on the marginals for the antiferromagnetic Ising distribution, a proof can be found in, e.g., Reference [6, Lemma 26]; there, the result is stated for bounded-degree graphs, but the proof of the marginal bound applies to any vertex whose degree is bounded.

LEMMA 5.3 (SEE, E.G., REFERENCE [6, LEMMA 26]). *Let $\beta \in (0, 1)$ and $D > 0$ be real numbers. For a graph $G = (V, E)$, let U be the set of vertices in G with degree $\leq D$ and suppose that $|U| \geq 2$. Then, the distribution $\mu := \mu_{G, \beta, U}$ is b -marginally bounded for $b = \frac{\beta^D}{1 + \beta^D}$.*

5.2 Monomer-Dimer Model

In this section, we bound the spectral independence of $G(n, d/n)$ in the monomer-dimer model. Instead of vertex-to-vertex influences that we considered for two-spin systems, we need to consider instead edge-to-edge influences. Namely, for a graph $G = (V, E)$ and $\mu = \mu_{G, \gamma}$, fix some $\tau \in \{0, 1\}^S$ for some $S \subset E$. Then, following Reference [13, Section 6], for edges e, f with $e \in E \setminus S$ and $0 < \mu_\tau(\sigma_e = 1) < 1$, we define the influence of e on f (under μ_τ) as

$$\mathcal{I}_G^\tau(e \rightarrow f) = \mu_\tau(\sigma_f = 0 \mid \sigma_e = 0) - \mu_\tau(\sigma_f = 0 \mid \sigma_e = 1).$$

We will need the following technical lemma from Reference [35]. The version below can be more easily derived from Reference [5, Lemma 15] (where it is explained how to combine the results of Reference [35]; in the notation below, the values of a, χ correspond to those of p, q there, respectively).

LEMMA 5.4 ([35]). *Let $d > 1$ and $\gamma > 0$. Let $\chi = (1 + 4\gamma\hat{d})^{1/2}$ where $\hat{d} = \max\{d, 3/(4\gamma)\}$, and set $a = \frac{\chi}{\chi - 1}$. Consider also the function $\Phi(x) = \frac{1}{x(2-x)}$ for $x \in (0, 1]$. Then, there is a constant $0 < \kappa < 1/d$ given by $\kappa = \frac{1}{d}(1 - \frac{2}{1 + \chi})^\chi$ such that the following holds for any integer $k \geq 1$.*

Let $R_1, \dots, R_k \in (0, 1]$ be reals and $R = \frac{1}{1 + \gamma \sum_{j=1}^k R_j}$. Then $(\Phi(R))^a \sum_{j=1}^k (\frac{\gamma R_j^2}{\Phi(R_j)})^a \leq \kappa^{a/\chi}$.

With this in hand, and using the notion of branching values (cf. Definition 3.3), we are now able to bound the total sum of influences.

LEMMA 5.5. *Let $d > 1$ and $\gamma > 0$ be constants. Then, there is a constant $\chi > 1$ such that the following holds.*

Let $T = (V, E)$ be a tree and e be an edge of T , whose endpoints v_1 and v_2 have d -branching values $\leq \alpha$, and let $k = \max\{\deg(v_1), \deg(v_2)\}$. Then, for the monomer-dimer distribution on T with parameter γ , it holds that

$$\sum_{f \in E} |\mathcal{I}_T(e \rightarrow f)| \leq W_k \alpha^{1/\chi},$$

where $W_k > 0$ is a real depending only on k (and the constants d, γ).

PROOF. Let $\kappa \in (0, 1/d)$ and $\chi > 1$ be the constants from Lemma 5.4, and $\Phi(x) = \frac{1}{x(2-x)}$ be also as in Lemma 5.4. Let $M_k = 2k/\kappa^{1/\chi}$ and $W_k = \frac{M_k}{1-(d\kappa)^{1/\chi}}$.

Let $\mu = \mu_{T, \gamma}$. For $i \in \{1, 2\}$, let \hat{T}_i be the subtree containing v_i obtained from T when we remove the edge e , and let T_i be the tree obtained by adding the edge e to \hat{T}_i . We will root T_i at v_i , and set $\mu_i = \mu_{T_i, \gamma}$. We will show that

$$\sum_{f \in E_{T_i}} |\mathcal{I}_{T_i}(e \rightarrow f)| \leq W_k \alpha^{1/\chi}, \quad (10)$$

where $\mathcal{I}_{T_i}(e \rightarrow \cdot)$ denotes the influence of e on the edges of T_i under μ_i . Note that for any $f \in E_{T_i}$, we have that $\mathcal{I}_T(e \rightarrow f) = \mathcal{I}_{T_i}(e \rightarrow f)$. Therefore, by adding Equation (10) for $i = 1, 2$, we obtain the statement of the lemma.

We therefore focus on proving Equation (10) for $i = 1, 2$. The argument is analogous to that used in Lemma 3.6 with suitable adaptations to account for the monomer-dimer model. We will work inductively on the depth of T_i . Analogously to that lemma, we first define for each vertex u in T_i the following values α_u and R_u . If u is a leaf, then set $\alpha_u = 1$ and $R_u = 1$; otherwise, set $\alpha_u = 1 + \frac{1}{d} \sum_{j=1}^t \alpha_{u_j}$ and $R_u = \frac{1}{1 + \gamma \sum_{j=1}^t R_{u_j}}$, where u_1, \dots, u_t are the children of u in the rooted tree T_i . Note that we have that $\alpha_{v_i} \leq \alpha$, where α is the upper bound on the d -branching value of v_i in the tree T_i . Moreover, if we denote by $T_i(u)$ the subtree of T_i rooted at u , then a standard induction argument shows that

$$R_u = \mu_{T_i(u), \gamma} (u \text{ is unmatched}). \quad (11)$$

For an integer $h \geq 1$, let $L_i(h)$ be the set of edges at distance h from the edge e in the tree T_i , so that $L_i(1)$ consists of the edges incident to v_i other than e , and so on. Moreover, for an edge f , let $v(f)$ be the endpoint of f , which is the farthest from v_{3-i} . We will show that

$$\sum_{f \in L_i(h)} \left(\frac{\alpha_{v(f)}}{\alpha_{v_i}} \right)^{1/\chi} \frac{|\mathcal{I}_{T_i}(e \rightarrow f)|}{R_{v(f)} \Phi(R_{v(f)})} \leq M_k (d\kappa)^{h/\chi}. \quad (12)$$

Note that $\alpha_v \geq 1$ and $R_v \Phi(R_v) \leq 1$ for all vertices v in T_i , and for the root $\alpha_{v_i} \leq \alpha$. So, Equation (12) yields $\sum_{f \in L_i(h)} |\mathcal{I}_{T_i}(e \rightarrow f)| \leq M_k \alpha^{1/\chi} (d\kappa)^{h/\chi}$ for all integer $h \geq 0$, and therefore summing over h , yields Equation (10), since $\kappa < 1/d$. So, it only remains to prove Equation (12).

We will work inductively. For the base case $h = 1$, we have for every edge $f \in L_i(h)$ the trivial bounds $|\mathcal{I}_{T_i}(e \rightarrow f)| \leq 1$ and $1/(R_{v(f)} \Phi(R_{v(f)})) = 2 - R_{v(f)} \leq 2$. Moreover, by the definition of α_{v_i} , we have $\alpha_{v_i} \geq \frac{1}{d} \alpha_{v(f)}$, so $\frac{\alpha_{v(f)}}{\alpha_{v_i}} \leq d$. Using these and $M_k = 2k/\kappa^{1/\chi}$, we obtain that Equation (12) holds for $h = 1$. For the induction step, consider $f \in L_i(h)$, and let $v = v(f)$. Suppose v has $k_v \geq 0$ children in the tree T_i , joined by the edges f_j for $j \in [k_v]$. Then, for each $j \in [k_v]$, since f is on

the unique path joining e to f_j and the edges f, f_j cannot simultaneously belong to a matching, it holds that (see Reference [13, Lemmas 6.11 and 6.12])

$$\mathcal{I}_{T_i}(e \rightarrow f_j) = \mathcal{I}_{T_i}(e \rightarrow f) \mathcal{I}_{T_i}(f \rightarrow f_j), \text{ and } |\mathcal{I}_{T_i}(f \rightarrow f_j)| = \mu_{T_i(v(f))}(\sigma_{f_j} = 1) = \gamma R_{v(f_j)} R_{v(f)},$$

where the last equality follows from Equation (11). Therefore, we can write

$$\begin{aligned} & \sum_{f \in L(h+1)} \left(\frac{\alpha_{v(f)}}{\alpha_{v_i}} \right)^{1/\chi} \frac{|\mathcal{I}_{T_i}(e \rightarrow f)|}{R_{v(f)} \Phi(R_{v(f)})} \\ &= \sum_{f \in L(h)} \left(\frac{\alpha_{v(f)}}{\alpha_{v_i}} \right)^{1/\chi} \frac{|\mathcal{I}_{T_i}(e \rightarrow f)|}{R_{v(f)} \Phi(R_{v(f)})} \sum_{j \in [k_{v(f)}]} \left(\frac{\alpha_{v(f_j)}}{\alpha_{v(f)}} \right)^{1/\chi} R_{v(f)} \Phi(R_{v(f)}) \frac{|\mathcal{I}_{T_i}(f \rightarrow f_j)|}{R_{v(f_j)} \Phi(R_{v(f_j)})}, \\ & \leq M_k(d\kappa)^{h/\chi} \max_{f \in L(h)} \left\{ \sum_{j \in [k_{v(f)}]} \left(\frac{\alpha_{v(f_j)}}{\alpha_{v(f)}} \right)^{1/\chi} \Phi(R_{v(f)}) \frac{\gamma R_{v(f)}^2}{\Phi(R_{v(f_j)})} \right\}. \end{aligned} \quad (13)$$

Consider an arbitrary $f \in L(h)$. Since $\frac{1}{\chi} + \frac{1}{a} = 1$, by Hölder's inequality, we have that

$$\sum_{j \in [k_{v(f)}]} \left(\frac{\alpha_{v(f_j)}}{\alpha_{v(f)}} \right)^{1/\chi} \Phi(R_{v(f)}) \frac{\gamma R_{v(f)}^2}{\Phi(R_{v(f_j)})} \leq \left(\sum_{j \in [k_{v(f)}]} \frac{\alpha_{v(f_j)}}{\alpha_{v(f)}} \right)^{1/\chi} \left((\Phi(R_{v(f)}))^a \sum_{i \in [k_v]} \left(\frac{\gamma R_{v(f)}^2}{\Phi(R_{v(f_j)})} \right)^a \right)^{1/a}. \quad (14)$$

By Lemma 5.4, we have that $((\Phi(R_{v(f)}))^a \sum_{i \in [k_v]} (\frac{\gamma R_{v(f)}^2}{\Phi(R_{v(f_j)})})^a)^{1/a} \leq \kappa^{1/\chi}$. Moreover, by the definition of the α values, we have $\alpha_{v(f)} = 1 + \frac{1}{d} \sum_{j \in [k_{v(f)}]} \alpha_{v(f_j)} \geq \frac{1}{d} \sum_{j \in [k_{v(f)}]} \alpha_{v(f_j)}$. Plugging these bounds into Equation (14), and then into Equation (13) completes the inductive proof of Equation (12). \square

We can now conclude the following spectral independence property for the monomer-dimer distribution on $G(n, d/n)$.

LEMMA 5.6. *Let $d \geq 1$ and $\gamma > 0$. Then, for any constants $D, \varepsilon > 0$, whp over the choice of $G \sim G(n, d/n)$, the marginal monomer-dimer distribution $\mu_{G, \gamma, F}$, where F is the set of edges in G whose both endpoints have degree $\leq D$, is $(\varepsilon \log n)$ -spectrally independent.*

PROOF. Consider arbitrary $D, \varepsilon > 0$ and let $d' > d$. Let $\chi > 1$ and $W = \max\{W_1, \dots, W_D\}$ where χ and the W_k 's are as in Lemma 3.6 (corresponding to the constants d', γ). By Lemma 3.4, whp all of the vertices the graph $G = (V, E) \sim G(n, d/n)$ have d' -branching value less than $\varepsilon \log n$. We will show that the result holds for all such graphs G .

Let for convenience $\mu = \mu_{G, \gamma, F}$, where F is the set of edges whose both endpoints have degree $\leq D$. Consider arbitrary $S \subset F$ and $\tau \in \Omega_S$. We will once again bound the largest eigenvalue of the influence matrix Ψ_τ by bounding the absolute-value row sums of Ψ_τ . It suffices to consider the case that S is empty (and τ is trivial), since conditioning on an arbitrary τ is equivalent to the monomer-dimer model on a subgraph of G . Analogously to the hard-core model (cf. proof of Lemma 3.8), to bound the largest eigenvalue, it suffices to show for arbitrary $e \in F$ that

$$\sum_{f \in F} |\mathcal{I}_G(e \rightarrow f)| \leq \varepsilon \log n, \quad (15)$$

where $\mathcal{I}_G(e \rightarrow \cdot)$ denotes the influence of e on the edges of G under the (full) distribution $\mu_{G, \gamma}$. In Reference [14, Theorem 6.2], they showed the analogue of Lemma 3.2 in the case of matchings.

In particular, let v be one of the endpoints of e and consider the self-avoiding walk tree $T = (V_T, E_T)$ emanating from v , then Reference [13, Theorem 6.2] asserts that

$$\sum_{f \in E} |\mathcal{I}_G(e \rightarrow f)| \leq \sum_{f \in E_T} |\mathcal{I}_T(e \rightarrow f)|,$$

where $\mathcal{I}_T(e \rightarrow \cdot)$ denotes the influence of e to the edges of T , under $\mu_{T,y}$. By Lemma 5.5 applied to the tree T , the right-hand side is bounded by $W(\varepsilon \log n)^{1/\chi}$, which is less than $\varepsilon \log n$ for all sufficiently large n . This gives Equation (15), as wanted. \square

5.3 Proof of Theorems 1.2 and 1.3

PROOF OF THEOREM 1.2. This is completely analogous to the proof of Theorem 1.1 presented in Section 2. Lemma 2.1 is about structural properties of the random graph $G \sim G(n, d/n)$, so we can use it verbatim. Therefore, we only need to establish the analogue of Lemma 2.2, i.e., for any $\theta > 0$ and all sufficiently large $D > 0$, whp over $G \sim G(n, d/n)$, running Glauber dynamics on U , where U is the set of vertices with degree $\leq D$, gives an ε -sample from $\mu_{G,\beta,U}$ in time $n^{1+\theta} \log \frac{1}{\varepsilon}$. Completely analogously to the hard-core model, using now the spectral independence bound of Lemma 5.2, we conclude that for any integer $r \in [\theta|U|, |U|]$ the distribution $\mu_{G,\beta,U}$ satisfies the r -uniform-block factorisation of entropy with multiplier $C_r \leq n^{\theta/3}$ (i.e., the analogue of Corollary 4.6 for the antiferromagnetic Ising model). From there, the same argument as in Lemma 4.8 yields that $\mu_{G,\beta,U}$ satisfies the 1-uniform-block factorisation of entropy with multiplier $C_1 \leq n^\theta$. \square

PROOF OF THEOREM 1.3. Again, this is very similar to the proof of Theorems 1.1 and 1.2, we only need to adapt the argument slightly to account for updating edges (instead of vertices).

Once again, we only need to establish the analogue of Lemma 2.2, i.e., for any $\theta > 0$ and all sufficiently large $D > 0$, whp over $G \sim G(n, d/n)$, running Glauber dynamics on F , where F is the set of vertices whose both endpoints have degree $\leq D$, gives an ε -sample from $\mu_{G,y,F}$ in time $n^{1+\theta} \log \frac{1}{\varepsilon}$. Using the spectral independence bound of Lemma 5.6, we obtain the analogue of Corollary 4.6, i.e., for any integer $r \in [\theta|F|, |F|]$ the distribution $\mu_{G,\beta,F}$ satisfies the r -uniform-block factorisation of entropy with multiplier $C_r \leq n^{\theta/3}$. The same argument as in Lemma 4.8 yields that $\mu_{G,\beta,F}$ satisfies the 1-uniform-block factorisation of entropy with multiplier $C_1 \leq n^\theta$. \square

6 RANDOM GRAPH PROPERTIES USED IN OUR PROOFS

In this section, we prove the random graph properties that we have used in our proofs, i.e., Lemmas 2.1, 3.4, and 4.7.

6.1 Bounding the Branching Value

Let $d \geq 1$. In this section, we prove Lemma 3.4, which bounds the d -branching values of the vertices in $G(n, d/n)$. The key ingredient is to bound the following closely related quantity. For a graph G and a vertex v of G , let $\hat{N}_{v,r}$ be the number of vertices at (graph) distance exactly r from v . We call a vertex v ε -good if it holds that

$$\hat{S}_v := \sum_{r \geq 0} \hat{N}_{v,r} / ((1 + \varepsilon)d)^r \leq \varepsilon \log n.$$

Note that the main difference between \hat{S}_v and the d -branching value S_v (cf. Definition 3.3) is that the latter is defined with respect to the number of paths, and therefore we trivially have $\hat{S}_v \leq S_v$. For random graphs, however, we will also be able to show that $S_v \leq 2\hat{S}_v$.

We start with showing that whp all vertices in $G(n, d/n)$ are ε -good.

LEMMA 6.1. For any constants $d \geq 1$ and $\varepsilon > 0$, whp over the choice of $G \sim G(n, d/n)$, all vertices of G are ε -good.

PROOF. For integers $n \geq 1$, consider a (finite) tree rooted at ρ created by a branching process where a node at depth r has $\text{Bin}(n-r, \frac{d}{n-r})$ children, and let $Y_{n,r}$ be the number of nodes at depth r . Set

$$X_n = \sum_{r=0}^{n-1} \alpha^r Y_{n,r}, \text{ where } \alpha := \frac{1}{d(1+\varepsilon)} < 1/d \leq 1.$$

For any fixed vertex v of G , we can couple S_v and X_n , so that $S_v \leq X_n$. Indeed, the BFS exploration process in G starting from v is stochastically dominated by the branching process above, since at depth r of the BFS process the number of unexplored vertices is at most $n-r$ and the connection probability $\frac{d}{n}$ is less than $\frac{d}{n-r}$. In particular, there is a coupling so that $\hat{N}_{v,r} \leq Y_{n,r}$ for all integers $r \geq 0$, therefore yielding $\hat{S}_v \leq X_n$ as well. It follows that

$$\Pr[\hat{S}_v \geq \varepsilon \log n] \leq \Pr[X_n \geq \varepsilon \log n]. \quad (16)$$

To bound the probability $\Pr[X_n \geq \varepsilon \log n]$, we will use moment generating functions, adapting an argument from Reference [30]. For integer $n \geq 1$ and real $t \geq 0$, let $g_n(t) = \mathbb{E}[\exp(tX_n)]$ and consider the functions $\{f_n(t)\}_{n \geq 1}$ defined for $t \geq 0$ by

$$f_1(t) = \exp(t), \quad f_n(t) = \exp(t + d(f_{n-1}(\alpha t) - 1)) \text{ for } n \geq 2.$$

We will show by induction on n that $g_n(t) \leq f_n(t)$ for all $t \geq 0$. Since $X_1 = 1$, the base case is trivial. For $n \geq 2$, observe that

$$X_n = 1 + \alpha(X_{n-1}^{(1)} + \dots + X_{n-1}^{(K)}) \text{ where } K \sim \text{Bin}(n, \frac{d}{n}),$$

and $X_{n-1}^{(1)}, X_{n-1}^{(2)}, \dots$ are i.i.d. variables with distribution X_{n-1} . Therefore,

$$\begin{aligned} g_n(t) &= \sum_{k=0}^n \binom{n}{k} \left(\frac{d}{n}\right)^k \left(1 - \frac{d}{n}\right)^{n-k} \mathbb{E}[\exp(t(1+\alpha(X_{n-1}^{(1)} + \dots + X_{n-1}^{(K)})))] = \exp\left(t + \frac{d}{n}(g_{n-1}(t\alpha) - 1)\right)^n \\ &\leq \exp\left(t + \frac{d}{n}(f_{n-1}(t\alpha) - 1)\right)^n \leq \exp(t + d(f_{n-1}(\alpha t) - 1)) = f_n(t), \end{aligned}$$

completing the inductive proof that $g_n(t) \leq f_n(t)$.

We will next define a function $F(t)$ so that $f_n(t) \leq F(t)$ for all n and t . To define F , let c be a constant such that $c > 1/(1 - \alpha d)$ and note that $c > 1$, since $\alpha d < 1$. Let $t_0 > 0$ be such that

$$\exp(c\alpha t_0) = 1 + \frac{c-1}{d}t_0, \quad (17)$$

and note that such t_0 exists, since both sides at $t = 0$ have value 1, and their derivatives at $t = 0$ are given by $c\alpha$ and $(c-1)/d$, satisfying $c\alpha < (c-1)/d$ from the choice of c . From convexity of $\exp(c\alpha t)$, we obtain that for all $t \in [0, t_0]$ it holds that

$$\exp(c\alpha t) \leq 1 + \frac{c-1}{d}t, \text{ or equivalently } \exp(t + d(\exp(c\alpha t) - 1)) \leq \exp(ct). \quad (18)$$

Now, consider the function F defined by setting $F(t) = \exp(ct)$ for $t \in [0, t_0]$; for $t > t_0$, we define F inductively on intervals $(t_0(1/\alpha)^k, t_0(1/\alpha)^{k+1}]$ by

$$F(t) = \exp(t + d(F(\alpha t) - 1)).$$

It is not hard to see that the function F is increasing and continuous, satisfying the inequality

$$\exp(t + d(F(\alpha t) - 1)) \leq F(t). \quad (19)$$

Using Equation (18), the validity of Equation (19) and monotonicity are trivial; to check continuity, note that F can only be discontinuous at the points $t_k = t_0(1/\alpha^k)$; let $k \geq 0$ be the first such integer. By construction, F is continuous at t_0 (see Equation (17)), and hence it must be the case that $k \geq 1$. But then

$$\lim_{t \downarrow t_k} F(t) = \lim_{t \downarrow t_k} \exp(t + d(F(\alpha t) - 1)) = \exp(t + d(F(t_{k-1}) - 1)) = F(t_k),$$

contradicting the choice of k . Now, we also prove, by induction, that $f_n(t) \leq F(t)$. We have $f_1(t) = \exp(t) \leq F(t)$, since $c > 1/(1 - \alpha d) > 1$. Now suppose $f_n(t) \leq F(t)$. Then, using Equation (19), we have

$$f_{n+1}(t) = \exp(t - d + d f_n(\alpha t)) \leq \exp(t - d + d F(\alpha t)) \leq F(t),$$

as needed.

Using Markov's inequality and the fact that $g_n(t) \leq f_n(t) \leq F(t)$ for arbitrary $t \geq 0$, we have

$$\Pr(X_n \geq \varepsilon \log n) = \Pr(\exp(t X_n) \geq \exp(t \varepsilon \log n)) \leq \frac{F(t)}{\exp(t \varepsilon \log n)} \leq \exp(-h(\varepsilon \log n)), \quad (20)$$

where the function h is defined for $A \geq 0$ by

$$h(A) := \sup_{t \geq 0} \{tA - \log F(t)\}.$$

Let $M = 1 + 5/\varepsilon$. For any $A \geq F(M)$, we have

$$\frac{h(A)}{A} = \sup_{t \geq 0} \left\{ t - \frac{\log F(t)}{A} \right\} \geq M - 1 = 5/\varepsilon,$$

where the inequality follows by considering $t = M$ and observing that $\log F(M) \leq F(M) \leq A$.

Since $F(M)$ is a constant (depending only on ε, d), we have that for all sufficiently large n it holds that $\varepsilon \log n \geq F(M)$, and therefore $h(\varepsilon \log n) \geq 5 \log n$. From Equations (16) and (20), we therefore obtain that $\Pr[S_v \geq \varepsilon \log n] \leq 1/n^5$, and Lemma 6.1 follows by taking a union bound over the n vertices of G . \square

To obtain the desired bound on the branching values from Lemma 6.1, we will need to relate $N_{v,r}$, the number of paths with $r + 1$ vertices from v , with $\hat{N}_{v,r}$, the number of nodes at distance r from v . This will follow by the following treelike property of $G(n, d/n)$. For a graph G , a vertex v of G and real $r > 0$, we use $B(v, r)$ to be the set of vertices in G at distance $\leq r$ from v .

LEMMA 6.2 ([30, LEMMA 7]). *Let $d \geq 1$. The following holds whp over $G = (V, E) \sim G(n, d/n)$ and $R = \lfloor (\log \log n)^2 \rfloor - 1$. For all $v \in V$, $|B(v, R)| \leq d^R \log n$ and the tree-excess of the induced graph $G[B(v, R)]$ is at most 1.*

We can now prove Lemma 3.4, which we restate here for convenience.

LEMMA 3.4. *Let $d \geq 1$. Then, for every $d' > d$ and $\varepsilon > 0$, whp over the choice of $G \sim G(n, d/n)$, the d' -branching value of every vertex in G is at most $\varepsilon \log n$.*

PROOF. It suffices to prove the result for arbitrarily small constant $\varepsilon > 0$ and $d' = (1 + \varepsilon)d$. Let $R = \lfloor (\log \log n)^2 \rfloor - 1$ and $\varepsilon' = \frac{d'-1}{4d'}\varepsilon$. We have that whp $G = (V, E) \sim G(n, d/n)$ satisfies the conclusions of Lemmas 6.1 and 6.2.

For arbitrary $v \in V$, we can bound the d' -branching value of v from above by $\sum_{i \geq 0} s_i$, where for an integer $i \geq 0$, we set $s_i := \sum_{r=i}^{(i+1)R} N_{v,r}/(d')^r$. From Lemma 6.2, the tree-excess of $G[B(v, R+1)]$

is at most 1, and hence for integers $r \leq R + 1$, we have the bound⁶

$$N_{v,r} \leq 2 \sum_{r'=0}^r \hat{N}_{v,r'}. \quad (21)$$

It follows that $s_0 \leq 2 \sum_{r'=0}^R \frac{\hat{N}_{v,r'}}{(d')^{r'}} \sum_{r=r'}^R \frac{1}{(d')^{r-r'}} \leq \frac{\varepsilon}{2} \log n$, where in the last inequality, we used the fact from Lemma 6.1 that all vertices in G are ε' -good, and in particular that $\sum_{r'=0}^R \frac{\hat{N}_{v,r'}}{(d')^{r'}} \leq \varepsilon' \log n$.

To bound s_i for an integer $i \geq 1$, note that for integer $iR \leq r \leq (i+1)R$, we can decompose a path with r vertices starting from v into two paths with $r-R$ and $R+1$ vertices, so we have

$$N_{v,r} \leq N_{v,r-R} \max_{w \in B(v, r-R)} N_{w,R+1} \leq 2d^{R+1} N_{v,r-R} \log n,$$

where the last inequality follows from applying Equation (21) for $r = R+1$ and noting that $\sum_{r'=0}^{R+1} \hat{N}_{v,r'} = |B(v, R+1)| \leq d^{R+1} \log n$ from Lemma 6.2. Therefore,

$$s_i = \sum_{r=iR}^{(i+1)R} \frac{N_{v,r}}{(d')^r} \leq s_{i-1} \sum_{r=iR}^{(i+1)R} \frac{2d^{R+1} \log n}{(d')^R} \leq s_{i-1} \frac{2(R+1)d^{R+1} \log n}{(d')^R} \leq s_{i-1}/2, \quad (22)$$

where the last inequality is true for all sufficiently large n , since $d' > d$ and $R = \omega(\log \log n)$.

Using Equation (22) and summing over $i \geq 0$, we have that the d' -branching value of v is bounded by $\sum_{i \geq 0} s_i \leq 2s_0 \leq \varepsilon \log n$. Since v was an arbitrary vertex of G , this finishes the proof. \square

6.2 Bounding the Tree-excess of Small Connected Sets

We start with the following lemma, which shows that the tree-excess of logarithmically sized connected sets in $G(n, d/n)$ is bounded by an absolute constant (a similar result in a slightly different setting was shown in Reference [21]).

LEMMA 6.3. *Let $d, M > 0$ be arbitrary reals. There exists a positive integer ℓ such that the following holds whp over the choice of $G = G(n, d/n)$. There is no connected set S of vertices such that $|S| \leq M \log n$ and the tree excess is more than ℓ .*

PROOF. It suffices to prove the lemma for all sufficiently large d , since the property is increasing under edge inclusion; in particular, we will assume that $d \geq e$. Let ℓ be an integer bigger than $2M + 5$.

For a positive integer $k \leq \lfloor M \log n \rfloor$, we calculate the expected number of sets S as in the statement of the lemma with $|S| = k$. There are $\binom{n}{k}$ choices for the set S , k^{k-2} labelled trees on S , and at most $\binom{k^2}{\ell}$ choices for ℓ additional edges. The probability that all these $k + \ell - 1$ edges are in G is $(d/n)^{k-1+\ell}$. Therefore, by a union bound, we can upper bound the expected number of such sets S with $|S| = k$ by

$$\binom{n}{k} k^{k-2} \binom{k^2}{\ell} \left(\frac{d}{n}\right)^{k+\ell-1} \leq \frac{e^{k+\ell} d^{k+\ell-1} k^{2\ell}}{n^{\ell-1} \ell^\ell} \leq \frac{1}{n^2},$$

where the last inequality holds for all sufficiently large n using that $\ell > 2M + 5$. By summing over the $O(\log n)$ values for the integer k , we obtain that the expected number of such sets S with $|S| \leq M \log n$ is $o(1)$, and therefore the result follows by Markov's inequality. \square

⁶This can be proved by considering the BFS tree from v . An “excess” edge can only connect vertices either at the same or consecutive levels of the BFS tree. In either case, every path of length r that uses that edge can be mapped injectively to a vertex at distance $\leq r$ from v .

6.3 Bounding the Number of High-degree Vertices in Connected Components

In this section, we prove Lemma 4.7, that bounds the number of connected sets in $G(n, d/n)$ as well as the fraction of high-degree vertices. We also prove Lemma 2.1. (Note, Lemma 4.7 shows the existence of a quantity L , which depends on a quantity M from Lemma 6.6: We take δ to be a sufficiently small constant, $\Delta = 1/(\delta \log \frac{1}{\delta})$, $M = \max\{10d, 50(1 + \log d)\}$, and $L = 4M$. We state these quantities upfront to help with the readability of the expressions in this section.)

For a graph $G = (V, E)$ and a set $S \subseteq V$, we let $\deg_G(S)$ denote the sum of the degrees of the vertices in S . To control the number of connected sets containing a given vertex, we will use the following bound that holds for general graphs G .

LEMMA 6.4 ([21, LEMMA 6]). *Let $G = (V, E)$ be a graph, $v \in V$, and $\ell \geq 1$ be an integer. The number of connected sets $S \subseteq V$ such that $v \in V$ and $\deg_G(S) = \ell$ is at most $(2e)^{2\ell-1}$.*

We will use the following version of Chernoff bounds to upper bound the total degree of connected sets with $\Omega(\log n)$ size.

LEMMA 6.5. *Suppose X_1, \dots, X_n are i.i.d. Bernoulli random variables. Let $X = X_1 + \dots + X_n$ denote their sum and set $\mu = \mathbb{E}[X]$. Then, for any $t \geq 5$, $\Pr(X > t\mu) \leq e^{-\frac{1}{2}t\mu \log t}$.*

PROOF. It is well-known (see, e.g., Reference [24, Theorem 2.1]) that for any $\delta > 0$, it holds that

$$\Pr(X > (1 + \delta)\mu) \leq \left(\frac{e^\delta}{(1 + \delta)^{(1 + \delta)}} \right)^\mu.$$

For $\delta \geq 4$, we have that $\frac{(1+\delta)^{(1+\delta)}}{e^\delta} \geq (1 + \delta)^{(1+\delta)/2}$, and the result follows. \square

LEMMA 6.6. *Let $d > 0$ be an arbitrary real. There is an $M > 0$ such that for any $\delta \in (0, 1)$, whp over the choice of $G = G(n, d/n)$, every connected set S of vertices with $|S| = \lceil \delta \log n \rceil$ satisfies $\deg_G(S) \leq M\Delta|S|$, where $\Delta = 1/(\delta \log \frac{1}{\delta})$.*

PROOF. The proof is close to an argument of Fountoulakis and Reed [19, Lemma 2.4], the only difference is that we have to account for the smaller size of S . We may assume for convenience that $d \geq 1$, since the graph property is increasing with respect to adding edges in the graph G . Let $M = \max\{10d, 50(1 + \log d)\}$, and $\delta \in (0, 1)$ be an arbitrarily small constant, we will in fact assume henceforth that $\delta < 1/e^2$. Let $\Delta = 1/(\delta \log(1/\delta))$ and note that $\Delta > 2$. For convenience, let $k = \lceil \delta \log n \rceil$.

For a set S , let $e_{\text{in}}(S)$ be the number of edges whose both endpoints belong to S , and $e_{\text{out}}(S)$ be the number of edges with exactly one endpoint in S . By Lemma 6.3, whp we conclude the crude bound $e_{\text{in}}(S) \leq 2|S| = 2k$ for all connected sets S with size k (since for such sets S the tree excess of $G[S]$ is bounded by an absolute constant and $|S| = k = \Omega(\log n)$). We also have that $\mathbb{E}[e_{\text{out}}(S)] \leq dk$, so by Lemma 6.5,

$$\Pr\left(e_{\text{out}}(S) \geq \frac{1}{2}kM\Delta\right) \leq e^{-\frac{1}{4}kM\Delta \log \frac{\Delta M}{4d}}.$$

Since $M \geq 4d$, $\delta < 1/e^2$ and $\Delta = 1/(\delta \log \frac{1}{\delta})$, we have that $\log \frac{\Delta M}{4d} \geq \log \Delta \geq \frac{1}{2} \log(1/\delta)$. Therefore, since $k \geq \delta \log n$ and $M \geq 50(1 + \log d)$, we have that

$$\frac{1}{4}kM\Delta \log \frac{\Delta M}{4d} \geq 3(1 + \log d) \log n.$$

Using that the number of labelled trees on a vertex set of size k is k^{k-2} , we have that the expected number of trees with size k in $G(n, d/n)$ is $\binom{n}{k}k^{k-2}(d/n)^k \leq n(ed)^k$, and hence there are in expectation at most $n(ed)^k$ connected sets S with size k . Note also that the event $e_{\text{out}}(S) \geq \frac{1}{2}kM\Delta$ is independent of the event that S is connected. Therefore, by linearity of expectation,

we obtain that the number of connected sets S with $|S| = k$ and $e_{\text{out}}(S) \geq \frac{1}{2}kM\Delta$ is at most $n(ed)^k e^{-3(1+\log d)\log n} \leq n(ed)^{\log n} e^{-3(1+\log d)\log n} \leq 1/n$, and hence whp there are no such sets by Markov's inequality. This finishes the proof of the lemma, since any other connected set S of size k satisfies $\deg_G(S) \leq 2e_{\text{in}}(S) + e_{\text{out}}(S) \leq 4k + \frac{1}{2}kM\Delta \leq M\Delta k$. \square

We next prove Lemma 4.7, which we restate here for convenience.

LEMMA 4.7. *Let $d > 0$ be an arbitrary real. There exists an $L > 0$ such that for any $\delta \in (0, 1)$, the following holds whp over the choice of $G = (V, E) \sim G(n, d/n)$. For $\Delta = 1/(\delta \log \frac{1}{\delta})$, for all integers $k \geq \delta \log n$ and any $v \in V$, there are $\leq (2e)^{\Delta L k}$ connected sets $S \subseteq V$ containing v with $|S| = k$. Moreover, every such set contains $\geq k/2$ vertices with degree less than $L\Delta$.*

PROOF. Let M be as in Lemma 6.6, and let $L = 4M$. Consider arbitrary $\delta \in (0, 1)$. By Lemma 6.6, we have whp that every connected set S of vertices in G with $|S| = r := \lceil \delta \log n \rceil$ has in total at most $M\Delta|S|$ edges incident to it, where $\Delta = 1/(\delta \log \frac{1}{\delta})$. We claim that for all connected sets S with $|S| \geq r$, it holds that $\deg_G(S) \leq 2M\Delta|S|$. Indeed, we can decompose any connected set S into at most $t \leq 2|S|/r$ connected sets (not necessarily disjoint) S_1, \dots, S_t , each of size k .⁷ From this, it follows that

$$\deg_G(S) \leq \sum_{i=1}^t \deg_G(S_i) \leq tM\Delta r \leq 2M\Delta|S|,$$

as claimed. The bound on the number connected sets S with $|S| = k \geq \delta \log n$ containing v follows by applying Lemma 6.4, by aggregating over the possible values of $\deg_G(S)$, which can be at most $2M\Delta k$. Moreover, the number of vertices in such a set S with degree $\geq L\Delta$ has to be at most $k/2$ (otherwise, $\deg_G(S) > L\Delta(k/2) = 2M\Delta k$). This finishes the proof. \square

We can also give now the proof of Lemma 2.1, which we restate here for convenience.

LEMMA 2.1. *Let $d > 0$ be an arbitrary real. There exist constants $D, \ell > 0$ such that the following holds whp over the choice of $G = (V, E) \sim G(n, d/n)$. Each of the connected components of $G[V \setminus U]$, where U is the set of vertices of degree $\leq D$, has size $O(\log n)$ and tree-excess at most ℓ .*

PROOF. Let $D = eL$, where $L > 0$ is the constant from Lemma 4.7. Let ℓ be the integer of Lemma 6.3 corresponding to $M = 1/e$.

By Lemma 4.7 applied to $\delta = 1/e$, whp over the choice of G , every connected set S with size $k \geq \frac{1}{e} \log n$ has at least $k/2$ vertices with degree $\leq D = eL$. Therefore, with U being the set of vertices with degree $\leq D$, we obtain that the components of $G[V \setminus U]$ have size at most $\frac{1}{e} \log n$. Moreover, by Lemma 6.3, whp over the choice of G , all these components have tree excess $\leq \ell$. \square

6.4 Verifying the Random Graph Properties Efficiently

Here, we briefly discuss how to verify in time $n^{1+o(1)}$ that a random graph $G \sim G(n, d/n)$ satisfies the high-probability properties of Lemmas 2.1, 3.4, and 4.7. The property in Lemma 2.1 is immediate, since we only need to do an exploration of the graph $G[V \setminus U]$. To ensure that the property of Lemma 3.4 is satisfied, and following its proof, we only need to check the property in Lemma 6.2, and that for $R = \lfloor (\log \log n)^2 \rfloor - 1$ and all $v \in V$ it holds that $\sum_{r=0}^R N_{v,r}/d^r = o(\log n)$. Both of these properties can be checked via enumeration in time $n^{1+o(1)}$. Similarly for the property in Lemma 4.7, whose proof used only the property in Lemma 6.6 (which is verifiable in time $n^{1+o(1)}$).

⁷One way to do this is to consider a spanning tree of S , double its edges, and obtain an Eulerian tour of the resulting graph. The desired decomposition of S can then be obtained by traversing the tour and extracting connected sets with k vertices.

Remark 6.7. We claimed in the Introduction that the family of $O(n^{1+\theta})$ algorithms from Theorem 1.1 can be turned into an $n^{1+o(1)}$ algorithm. We now explain how to do this for the interested reader.

As noted after the statement of Theorem 1.1, there is a function $f_{d,\lambda,\theta} : \mathbb{Z} \rightarrow \mathbb{R}$ such that $\lim_{n \rightarrow \infty} f_{d,\lambda,\theta}(n) = 0$ and the “whp” bound in Theorems 1.1, 1.2, and 1.3 means with probability $\geq 1 - f_{d,\lambda,\theta}(n)$; the function equals 1 for small n (making the conclusion trivial for such n).

To understand the function $f_{d,\lambda,\theta}(n)$, we need to look at the whp bounds in the lemmas that we use. The whp bound functions are $\geq 1 - 1/n$ for sufficiently large n where sufficiently large is (ultimately) a simple function of d, λ, θ (for example, Lemma 2.1, 4.7, 6.3, and 6.6). Lemma 6.1 (and as a consequence Lemma 3.4) has a more wild dependence on the parameter ε : the function $F(t)$ inside the proof can be bounded from above by a tower of exponentials of depth $\log_d(t)$. The whp bound is then $\geq 1 - 1/n$ assuming $n \geq \exp(F(1+5/\varepsilon)/\varepsilon)$; this whp bound propagates to Lemma 4.8, then to Lemma 2.2 and finally to Theorems 1.1, 1.2, and 1.3.

For fixed d and λ , the dependence of ε on θ (used in Lemma 6.1) can be extracted from Corollary 4.6 and Lemma 4.8, where ε (controlling the leading constant in the branching value) needs to be small enough to compensate the leading constant from the spectral-independence bound η ; in particular, ε scales roughly as $1/e^{C/\theta}$, where C is a constant depending only on d and λ . It follows that we can set $f_{d,\lambda,\theta}(n) = 1/n$ when $n \geq N_{d,\lambda}(\theta)$ and 1 otherwise, where $N_{d,\lambda}(\theta)$ is a computable function satisfying $n \geq N_{d,\lambda}(1/k)$ for some function $k = k(n) = \omega(1)$ (for example, it suffices to take $k(n) = \lfloor \hat{C} \log^{(3)} \log^* \log^{(2)} n \rfloor$ where \hat{C} is a constant depending only on d and λ).

The $n^{1+o(1)}$ algorithm then proceeds as follows. Given the input size n , it first computes $k = k(n)$ as above in $O(\log n)$ time (say). Then, by the definition of $k(n)$, we have $n \geq N_{d,\lambda}(1/k)$ for all sufficiently large n . The new algorithm runs the algorithm of Theorem 1.1 with $\theta = 1/k$ in time $n^{1+1/k} \log \frac{1}{\varepsilon} = n^{1+o(1)} \log \frac{1}{\varepsilon}$, which succeeds with probability $\geq 1 - 1/n = 1 - o(1)$ over the choice of $G(n, d/n)$.

APPENDIX

A OTHER OMITTED PROOFS

PROOF OF THEOREM 4.5. The theorem follows by combining Reference [13, Claim 1.13, Theorem 1.14, Lemma 2.6], which bound the factorisation multiplier of the more general order- (r, s) down-up walk; the result here is the special case $s = n$. The only difference is that in Reference [13] they state the r -uniform-block multiplier $\hat{C}_r = \frac{r}{n} \frac{\sum_{k=0}^{n-1} \hat{\Gamma}_k}{\sum_{k=n-r}^{n-1} \hat{\Gamma}_k}$, with $\hat{\Gamma}_k = \prod_{j=0}^{k-1} \hat{\alpha}_j$, where

$$\hat{\alpha}_k = \max \left\{ 1 - \frac{4\eta_k}{b^2}, \frac{1 - \eta_k}{4 + 2 \log(\frac{1}{2b_k b_{k+1}})} \right\},$$

with $\eta_k = \frac{\eta}{n-1-k}$ for $k \in [n-1]$, $b_k = \frac{b}{n-k}$ for $k \in [n]$. The only thing we need to note, which is also implicit in Reference [13, Proof of Lemma 2.4], is that $\hat{C}_r \leq C_r$. In particular, for each $k \in [n]$, C_r is decreasing with respect to α_k (both numerator and denominator are multi-linear functions and the value of C_r increases as $\alpha_k \downarrow 0$), so we only need to check that $\hat{\alpha}_k \geq \alpha_k$. For this, we need to further note that η_k is an upper bound to the so-called local spectral expansion, which is defined as the second largest eigenvalue of the transition matrix of an appropriate random walk (on a suitable simplicial complex corresponding to μ ; the details are not important for our purposes). This implies that $\eta_k \leq 1$, and hence $\hat{\alpha}_k \geq 0$, since $\hat{\alpha}_k \geq \frac{1 - \eta_k}{4 + 2 \log(\frac{1}{2b_k b_{k+1}})}$, proving that $\hat{\alpha}_k \geq \alpha_k$. \square

PROOF OF COROLLARY 4.6. The proof is analogous to Reference [13, Proof of Lemma 2.4], we highlight the main differences for completeness (since the context is somewhat different). Let

$D, \theta > 0$ be arbitrary constants; we may assume that $D \geq 1$ (otherwise, U is empty) and $\theta \in (0, 1]$ (otherwise, there are no relevant r). Note that since U is the set of vertices with degree $\leq D$, we have from Corollary 3.9 that the marginal distribution $\mu_{G, \lambda, U}$ is b -marginally bounded for $b = \min\{\frac{1}{1+\lambda}, \frac{\lambda}{\lambda+(1+\lambda)^D}\}$. Let $u = |U|$, by a standard balls-and-bins bound, we have that whp $u \geq n^{1/2}$. By Lemma 3.8, we have that whp $\mu := \mu_{G, \lambda, U}$ is η -spectrally independent for $\eta = \frac{b^2\theta^2}{10} \log n$. Let $R = \lceil 4\eta/b^2 \rceil$ and note that for all sufficiently large n , $\theta u \geq \theta^2 \log n \geq 4R$.

Consider an arbitrary integer $r \in [\theta u, u]$ as in the lemma statement so that $u \geq r \geq 2R$. From Theorem 4.5, we therefore have that μ satisfies the r -uniform-block factorisation of entropy with multiplier $C_r = \frac{r}{u} \frac{\sum_{k=0}^{u-1} \Gamma_k}{\sum_{k=u-r}^{u-1} \Gamma_k} \leq \frac{\sum_{k=0}^{u-1} \Gamma_k}{\sum_{k=u-r}^{u-1} \Gamma_k}$, with $\Gamma_k = \prod_{j=0}^{k-1} \alpha_j$ for $k \in [u]$ and $\alpha_k = \max\{0, 1 - \frac{4\eta}{b^2(u-k-1)}\}$ for $k \in [u-1]$. Then, we can apply the exact same reasoning as in Reference [13] to obtain that

$$C_r \leq \left(\frac{u-R}{r-R}\right)^R \leq \left(\frac{2u}{r}\right)^R \leq \left(\frac{2}{\theta}\right)^R.$$

Note that for $x \in (0, 1)$, we have the inequality $(1/x)^{x^2} \leq e^x$, which for $x = \theta/2$ and using that $R \leq (\frac{\theta}{2})^2 \log n$ gives that $C_r \leq n^\theta$, as claimed. \square

ACKNOWLEDGMENTS

We are thankful to the anonymous referees for their helpful comments.

REFERENCES

- [1] V. L. Alev and L. C. Lau. 2020. Improved analysis of higher order random walks and applications. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC'20)*. 1198–1211.
- [2] N. Anari, V. Jain, F. Koehler, H. T. Pham, and T.-D. Vuong. 2022. Entropic independence: Optimal mixing of down-up random walks. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC'22)*. 1418–1430.
- [3] N. Anari, K. Liu, and S. Oveis Gharan. 2020. Spectral independence in high-dimensional expanders and applications to the hardcore model. In *Proceedings of the 61st Annual Symposium on Foundations of Computer Science (FOCS'20)*. 1319–1330.
- [4] N. Berger, C. Kenyon, E. Mossel, and Y. Peres. 2005. Glauber dynamics on trees and hyperbolic graphs. *Probabil. Theory Relat. Fields* 131, 3 (2005), 311–340.
- [5] I. Bezákova, A. Galanis, L. A. Goldberg, and D. Štefankovič. 2021. The complexity of approximating the matching polynomial in the complex plane. *ACM Trans. Comput. Theory* 13, 2, Article 13 (2021).
- [6] A. Blanca, A. Galanis, L. A. Goldberg, D. Štefankovič, E. Vigoda, and K. Yang. 2020. Sampling in uniqueness from the Potts and random-cluster models on random regular graphs. *SIAM J. Discrete Math.* 34, 1 (2020), 742–793.
- [7] A. Blanca and R. Gheissari. 2022. Sampling from Potts on random graphs of unbounded degree via random-cluster dynamics. In *Proceedings of the Conference on Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques (APPROX/RANDOM'22)*, Vol. 245. 24:1–24:15.
- [8] P. Caputo and D. Parisi. 2021. Block factorization of the relative entropy via spatial mixing. *Commun. Math. Phys.* 388, 2 (2021), 793–818.
- [9] X. Chen, W. Feng, Y. Yin, and X. Zhang. 2022. Optimal mixing for two-state anti-ferromagnetic spin systems. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS'22)*. 588–599.
- [10] X. Chen, W. Feng, Y. Yin, and X. Zhang. 2022. Rapid mixing of glauber dynamics via spectral independence for all degrees. In *Proceedings of the IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS'21)*. 137–148.
- [11] Y. Chen and R. Eldan. 2022. Localization schemes: A framework for proving mixing bounds for Markov chains. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS'22)*. 110–122.
- [12] Z. Chen. 2021. *Optimal Mixing of Markov Chains for Spin Systems via Spectral Independence*. Ph.D. Dissertation. Georgia Institute of Technology.
- [13] Z. Chen, K. Liu, and E. Vigoda. 2021. Optimal mixing of glauber dynamics: Entropy factorization via high-dimensional expansion. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing (STOC'21)*. 1537–1550.
- [14] Z. Chen, K. Liu, and E. Vigoda. 2023. Rapid mixing of glauber dynamics up to uniqueness via contraction. *SIAM J. Comput.* 52, 1 (2023), 196–237.

- [15] I. Dinur and T. Kaufman. 2017. High dimensional expanders imply agreement expanders. In *Proceedings of the 58th Annual IEEE Symposium on Foundations of Computer Science (FOCS'17)*. 974–985.
- [16] C. Efthymiou. 2014. MCMC sampling colourings and independent sets of $G(n, d/n)$ near uniqueness threshold. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'14)*. 305–316.
- [17] C. Efthymiou and W. Feng. 2023. On the mixing time of glauber dynamics for the hard-core and related models on $G(n, d/n)$. In *Proceedings of the 50th International Colloquium on Automata, Languages, and Programming (ICALP'23)*. 54:1–54:17.
- [18] C. Efthymiou, T. P. Hayes, D. Štefankovič, and E. Vigoda. 2018. Sampling random colorings of sparse random graphs. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'18)*. 1759–1771.
- [19] N. Fountoulakis and B. A. Reed. 2008. The evolution of the mixing rate of a simple random walk on the giant component of a random graph. *Random Struct. Algor.* 33, 1 (2008), 68–86.
- [20] A. Galanis, Q. Ge, D. Štefankovič, E. Vigoda, and L. Yang. 2014. Improved inapproximability results for counting independent sets in the hard-core model. *Random Struct. Algor.* 45, 1 (2014), 78–110.
- [21] A. Galanis, L. A. Goldberg, and J. Stewart. 2021. Fast mixing via polymers for random graphs with unbounded degree. In *Proceedings of the Conference on Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques (APPROX/RANDOM'21)*. 36:1–36:13. Retrieved from <https://arxiv.org/abs/2105.00524>
- [22] A. Galanis, D. Štefankovič, and Eric Vigoda. 2016. Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models. *Comb. Probab. Comput.* 25, 4 (2016), 500–559.
- [23] V. Jain, H. T. Pham, and T. D. Vuong. 2021. Spectral independence, coupling with the stationary distribution, and the spectral gap of the glauber dynamics. Retrieved from <https://arxiv.org/abs/2105.01201>
- [24] S. Janson, A. Rucinski, and T. Łuczak. 2011. *Random Graphs*. John Wiley & Sons.
- [25] M. Jerrum. 2003. *Counting, Sampling and Integrating: Algorithms and Complexity*. Springer Science & Business Media.
- [26] M. Jerrum and A. Sinclair. 1989. Approximating the permanent. *SIAM J. Comput.* 18, 6 (1989), 1149–1178.
- [27] T. Kaufman and D. Mass. 2017. High dimensional random walks and colorful expansion. In *Proceedings of the 8th Innovations in Theoretical Computer Science Conference (ITCS'17)*. Vol. 67. Art. No. 4, 27.
- [28] T. Kaufman and I. Oppenheim. 2020. High order random walks: Beyond spectral gap. *Combinatorica* 40, 2 (2020), 245–281.
- [29] L. Li, P. Lu, and Y. Yin. 2013. Correlation decay up to uniqueness in spin systems. In *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'13)*. 67–84.
- [30] E. Mossel and A. Sly. 2013. Exact thresholds for ising-gibbs samplers on general graphs. *Ann. Probabil.* 41, 1 (2013), 294–328. <https://doi.org/10.1214/11-AOP737>
- [31] I. Oppenheim. 2018. Local spectral expansion approach to high dimensional expanders part i: Descent of spectral gaps. *Discrete Comput. Geom.* 59, 2 (2018), 293–330.
- [32] V. Patel and G. Regts. 2017. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM J. Comput.* 46, 6 (2017), 1893–1919.
- [33] H. Peters and G. Regts. 2019. On a conjecture of Sokal concerning roots of the independence polynomial. *Mich. Math. J.* 68, 1 (2019), 33–55.
- [34] A. D. Scott and A. D. Sokal. 2005. The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma. *J. Stat. Phys.* 118 (2005), 1151–1261.
- [35] A. Sinclair, P. Srivastava, D. Štefankovič, and Y. Yin. 2017. Spatial mixing and the connective constant: Optimal bounds. *Probabil. Theory Relat. Fields* 168, 1 (2017), 153–197.
- [36] A. Sinclair, P. Srivastava, and M. Thurley. 2012. Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'12)*. 941–953.
- [37] A. Sinclair, P. Srivastava, and Y. Yin. 2013. Spatial mixing and approximation algorithms for graphs with bounded connective constant. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS'13)*. 300–309.
- [38] A. Sly. 2010. Computational transition at the uniqueness threshold. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS'10)*. 287–296.
- [39] A. Sly and N. Sun. 2012. The computational hardness of counting in two-spin Models on d -regular graphs. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS'12)*. 361–369.
- [40] D. Weitz. 2006. Counting independent sets up to the tree threshold. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC'06)*. 140–149.

Received 24 October 2022; revised 8 October 2023; accepted 10 October 2023