CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Glauber dynamics for the hard-core model on bounded-degree H-free graphs

Mark Jerrum

School of Mathematical Sciences, Queen Mary, University of London, London E1 4NS, UK
Email: m.jerrum@qmul.ac.uk

Abstract

The hard-core model has as its configurations the independent sets of some graph instance G. The probability distribution on independent sets is controlled by a 'fugacity' $\lambda > 0$, with higher λ leading to denser configurations. We investigate the mixing time of Glauber (single-site) dynamics for the hard-core model on restricted classes of bounded-degree graphs in which a particular graph H is excluded as an induced subgraph. If H is a subdivided claw then, for all λ, the mixing time is $O(n \log n)$, where n is the order of G. This extends a result of Chen and Gu for claw-free graphs. When H is a path, the set of possible instances is finite. For all other H, the mixing time is exponential in n for sufficiently large λ, depending on H and the maximum degree of G.

Keywords: Glauber dynamics; Spin systems; Hereditary graph classes; Mixing time

2020 MSC Codes: Primary: 05C75; Secondary: 05C69, 60J10, 68W20, 82B20

1. Introduction

The complete bipartite graph $K_{1,3}$ is known as a claw. A graph is said to be claw-free if it does not contain a claw as an induced subgraph. The hard-core model is specified by a graph G and a fugacity $\lambda > 0$. These inputs define a probability distribution on independent sets of a graph, the so-called hard-core distribution. Higher values of λ favour larger independent sets, leading to denser configurations. (Precise definitions will be given presently.)

Glauber dynamics defines an ergodic Markov chain on independent sets in a graph, whose stationary distribution is the hard-core distribution. We are interested in analysing the mixing time (time to near stationarity) of Glauber dynamics. With increasing fugacity λ the density of configurations increases and we expect the mixing time to increase too. For most graph classes, for example, graphs of bounded degree, which have been extensively studied [11, 25, 29], there is a critical λ above which the mixing time becomes exponential in the size of the graph. However, in his PhD thesis, Matthews [23] showed that, for any fixed λ, however, large, the mixing time of Glauber dynamics on a claw-free graph G is polynomial in $n = |V(G)|$.

Recently, Chen and Gu [9] showed (amongst other things) that, with the additional assumption that G has bounded degree Δ, the mixing time of Glauber dynamics on claw-free graphs is in fact $O(n \log n)$. The constant implicit in the O-notation depends on λ and Δ. We call this mixing 'optimal', since $O(n \log n)$ time is necessary for Glauber dynamics to visit every vertex in G, by the coupon-collector argument. The intuition that $O(n \log n)$ is optimal can be formalised, though doing so is surprisingly challenging [19]. Prior to Chen and Gu's work, optimal mixing for matchings in bounded-degree graphs – a related result – had been shown by Chen, Liu, and Vigoda [12].
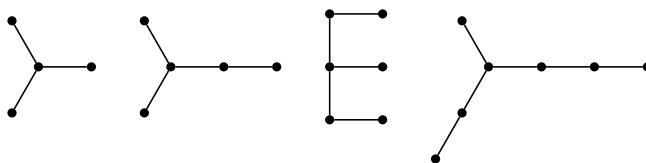
Figure 1. The claw, the fork, the E, and the skew star.

It is natural to wonder whether excluding graphs other than the claw can lead to optimal mixing time for all λ. There is an extensive literature on independent sets in graph families that exclude a fixed connected graph H as an induced subgraph. (From now on we restrict attention to connected H.) In this line of work, subdivided claws play a crucial role [1, 13]. A subdivided claw is obtained from a claw by subdividing the three edges by an arbitrary number of intermediate vertices. Alternatively, a subdivided claw is a tree with a single vertex of degree 3, and all other vertices having degree 1 or 2. For $1 \leq i \leq j \leq k$, we write $S_{i,j,k}$ for the subdivided claw whose leaves are at distance i, j, and k from the degree-3 vertex. The smaller subdivided claws have special names: $S_{1,1,1}$ (as we have seen) is the *claw*, $S_{1,1,2}$ the *fork*, $S_{1,2,2}$ is the *E*, and $S_{1,2,3}$ is the *skew star*. See Fig. 1.

It soon becomes apparent that the interesting cases for us are when the excluded graph H is a subdivided claw. Informally, the reason is this: if H is not a subdivided claw or a path, then it contains either a vertex of degree greater than 3, or a cycle, or two vertices of degree 3 connected by a path. We can exclude the first case by working with instance graphs G with maximum degree 3, and the other two cases by working with graphs whose edges have been 'stretched' or subdivided by sufficiently many vertices. This intuition is formalised in Section 6. Note that excluding a path in a bounded-degree graph leads to a finite graph class. The (fairly routine) conclusion, expressed in Theorem 18, is that Glauber dynamics mixes in exponential time for sufficiently large λ (depending on Δ and H but not on $n = |V(G)|$) when H is not a subdivided claw or a path.

Our main result, Theorem 17 in Section 6, asserts that bounded-degree graphs that exclude any subdivided claw support optimal mixing. Together with the matching negative result just mentioned, we obtain the following dichotomy. If $\mathcal{H}$ is any collection of graphs, we say that a graph G is $\mathcal{H}$-free if G is H-free (does not contain H as an induced subgraph) for all $H \in \mathcal{H}$.

Theorem 1. *Let $\mathcal{H}$ be any finite set of connected graphs, and $\Delta \geq 3$. Consider the hard-core model on $\mathcal{H}$-free graphs of maximum degree Δ. If $\mathcal{H}$ contains a subdivided claw or a path then Glauber dynamics mixes in optimal $O(n \log n)$ time, at any fixed fugacity $\lambda > 0$, where n is the number of vertices of G. Otherwise, for some sufficiently large fixed λ (depending on $\mathcal{H}$), the mixing time is exponential in n in the worst case, even when $\Delta = 3$.*

The above statement is essentially a composite of Theorems 17 and 18, and the (very short) proof can be found in Section 6.

There is a formal similarity between this dichotomy and that of Abrishami, Chudnovsky, Dibek, Cemil, and Rzçżewski [1], who studied the problem of finding the largest independent set in a bounded-degree H-free graph. The current result is technically not as difficult. This is to be expected: the algorithm, for finding a largest independent set in a claw-free graph is quite involved, whereas that for sampling an independent set is quite simple: simulate Glauber dynamics for $O(n \log n)$ steps.

It is interesting to note that claw-free graphs arise in nature as crystal structures, for example, the kagome and pyrochlore lattices. However, I am unaware of any real-life crystals that are free of subdivided claws without in fact being claw-free.

In related work, Chudnovsky and Seymour [14] showed that the partition function of the hard-core model on claw-free graphs has only (negative) real roots. The existence of a zero-free region that includes the positive real axis has implications for the existence of *deterministic* algorithms

- Let the current state (independent set) be I.
- Select a vertex $v \in V(G)$ uniformly at random.
- If $\Gamma_G(v) \cap I = \emptyset$ then:
  - with probability $\lambda/(1 + \lambda)$, set $I' \leftarrow I \cup \{v\}$;
  - with the remaining probability set $I' \leftarrow I \setminus \{v\}$;
- else: set $I' \leftarrow I$.
- The new state is I'.

Figure 2. Glauber dynamics for the hard-core model.

for approximating the partition function, through work of Barvinok [6] and Patel and Regis [24]. For fork-free graphs, a zero-free region, smaller but still including the positive real axis, has been identified by Bencs [8].

Note that the restriction to bounded-degree graphs is crucial. General fork-free graphs include the complete bipartite graphs of arbitrary sizes, which certainly do not support optimal or even polynomial-time mixing. Additional restrictions other than bounded degree can lead to polynomial-time algorithms; for example, Dyer, Jerrum, and Müller treat the case of fork-free, odd-hold free graphs (equivalent to fork-free perfect graphs) [17].

Finally, although we concentrate on the hard-core model here, a result of a similar flavour has been obtained for the antiferromagnetic Ising model. For this model, Glauber dynamics has been shown to have optimal mixing time on the class of line graphs at any non-zero temperature however small (i.e. for any interaction strength however large). This result is due to Chen, Liu, and Vigoda [10], strengthening the mixing time bound of Dyer, Heinrich, Jerrum, and Müller [16] from polynomial to optimal.

2. Preliminaries

Suppose G is a graph and $\lambda > 0$ a 'fugacity'. A set $I \subseteq V(G)$ is an *independent set* in G if no edge in $E(G)$ has both endpoints in I. The *weight* $\text{wt}(I)$ of independent set I is simply $\text{wt}(I) = \lambda^{|I|}$. Denote by $\mathcal{I}_G$ the set of all independent sets in G. The *hard-core distribution* $\mu_{G,\lambda}$ assigns probability $\mu_{G,\lambda}(I) = \text{wt}(I)/Z(G, \lambda)$ to each $I \in \mathcal{I}_G$, where $Z(G, \lambda)$ is the *partition function*.

$$Z(G, \lambda) = \sum_{I \in \mathcal{I}_G} \text{wt}(I).$$

We think of $\mu_{G,\lambda}$ being a distribution on $2^{V(G)}$ with support $\mathcal{I}_G$.

Glauber dynamics defines a Markov chain on the independent sets of G. It is presented in Fig. 2, where $\Gamma_G(v)$ is used to denote the set of neighbours of v in G. This Markov chain is ergodic, and converges to the hard-core distribution on independent sets of G.

Denoting the t-step transition probabilities of this Markov chain by $P^t(\cdot, \cdot)$, convergence to stationarity may be measured by total variation distance (i.e. one half of the ℓ_1 distance) at time t, that is,

$$d(t) = \max_{I \in \mathcal{I}_G} \{d_{\text{TV}}(P^t(I, \cdot), \mu_{G,\lambda})\}.$$

Then the *mixing time* of the Markov chain is defined to be $t_{\text{mix}} = \min\{t : d(t) \leq \frac{1}{4}\}$. The constant $\frac{1}{4}$ is arbitrary, subject to lying in the interval $(0, \frac{1}{2})$. It is standard [21, Section 4.5] that convergence to stationary is exponential, when time is measured in units of t_{mix}, specifically, $d(\ell\, t_{\text{mix}}) \leq 2^{-\ell}$, for all $\ell \in \mathbb{N}$. The central question is then: is t_{mix} bounded by a polynomial in $n = |V(G)|$ and, more particularly, is $t_{\text{mix}} = O(n \log n)$ (with the constant of proportionality being a function of λ and Δ)?

For a short while we change perspective very slightly and view independent sets in G as spin configurations in $\{0, 1\}^{V(G)}$. Naturally, the independent set $I \in \mathcal{I}_G$ corresponds to the spin

configuration σ, where

$$\sigma(v) = \begin{cases} 0, & \text{if } v \notin I; \\ 1, & \text{if } v \in I. \end{cases}$$

Let $\beta_{G,\lambda}$ be the product distribution of independent Bernoulli random variables, with success probability $p = \lambda/(1+\lambda)$ living on the vertices of G. The following simple fact will be important in what follows.

Lemma 2. *The distribution $\beta_{G,\lambda}$ stochastically dominates the distribution $\mu_{G,\lambda}$. That is, there is a coupling of random variables, σ and σ', distributed as $\mu_{G,\lambda}$ and $\beta_{G,\lambda}$, respectively, such that $\sigma \leq \sigma'$ in the product order on $\{0,1\}^{V(G)}$.*

Proof. A simple way to see this is by monotone coupling of Markov chains [18, §3.10.3]. Modify the Markov chain from Fig. 2 by removing the test $\Gamma_G(v) \cap I = \emptyset$ and the 'else' part of the conditional statement. Thus the probabilistic update is done whatever the state of the vertices adjacent to v. This modified process clearly converges to the product distribution $\beta_{G,\lambda}$. Now run the modified process starting from the all-1 configuration in parallel with Glauber dynamics starting from the all-0 configuration. Use the same random choice of vertex v at each step, and optimally couple the updates. It is easy to see that the first process remains above the second in the product order. The upper process converges to $\beta_{G,\lambda}$ and the lower to $\mu_{G,\lambda}$. □

Corollary 3. *Let G be a graph of maximum degree Δ, and v be a vertex of G. If I is sampled from $\mu_{G,\lambda}$ then*

$$\frac{\lambda}{(1+\lambda)^{\Delta+1}} \leq \mathbb{P}(v \in I) \leq \frac{\lambda}{1+\lambda}.$$

Proof. The upper bound is immediate from Lemma 2. For the lower bound, note that $\mathbb{P}(I \cap \Gamma(v) = \emptyset) \geq 1/(1+\lambda)^{\Delta}$, again from Lemma 2, and that $\mathbb{P}(v \in I \mid I \cap \Gamma(v) = \emptyset) \geq \lambda/(1+\lambda)$. □

Suppose μ and μ' are probability distributions on $\{0,1\}^{V(G)}$. The Wasserstein (or Kantorovich-Rubinstein) distance between μ and μ' (specialised to this application) is defined to be

$$W_1(\mu, \mu') = \inf_{\sigma \sim \mu, \sigma' \sim \mu'} d_{\mathrm{H}}(\sigma, \sigma'),$$

where d_{H} denotes Hamming distance, and the infimum is over all couplings of random variables σ, σ' with the appropriate distributions.

For a graph G and vertex $v \in V(G)$ let $\mu_{G,\lambda}^{(v,0)}$ (respectively, $\mu_{G,\lambda}^{(v,1)}$) denote the hard-core distribution on G at fugacity λ conditioned on vertex v being excluded from (respectively, included in) the independent set. We will be using the powerful spectral independence machinery of Anari, Liu, Oveis Gharan, and Vinzant [5] and Anari, Liu, and Oveis Gharan [4], as sharpened by Cryan, Guo, and Mousa [15] and Chen, Liu, and Vigoda [12]. However, we can ease our task by using Chen and Gu's packaging of the machinery for situations such as ours. Indeed, as we shall see presently, to show optimal mixing we just need to construct a coupling that achieves bounded Wasserstein distance $W_1\big(\mu_{G,\lambda}^{(v,0)}, \mu_{G,\lambda}^{(v,1)}\big)$.

A graph class is said to be *hereditary* if it is closed under taking induced subgraphs. Note that graph classes defined by excluding induced subgraphs, with or without a degree bound, are hereditary.

Corollary 4. *Let $\mathcal{G}$ be a hereditary class of graphs with degree bound Δ, and λ be positive. Suppose there exists $\eta > 0$ such that for all $G \in \mathcal{G}$ and all $v \in V(G)$ we have $W_1(\mu_{G,\lambda}^{(v,0)}, \mu_{G,\lambda}^{(v,1)}) \leq \eta$, where $\mu_{G,\lambda}$ is the hard-core distribution with fugacity λ on $V(G)$. Then the mixing time of Glauber dynamics for the hard-core model on $G \in \mathcal{G}$ at fugacity λ is bounded above by $O(n \log n)$, where $n = |V(G)|$. The constant implicit in the O-notation depends on λ, Δ, and η.*

Proof. This follows from [9, Theorem 9 and Lemma 10],[1] which builds on earlier work, such as [12]. In order to make the correspondence we need to discuss briefly the notion of 'pinning'. The distributions $\mu_{G,\lambda}^{(v,0)}$ and $\mu_{G,\lambda}^{(v,1)}$ are examples of distributions obtained by pinning, in this instance by pinning vertex v to 0 and 1, respectively. More generally, we may consider pinning a set of vertices, say $\Lambda \subset V(G)$, to values given by $\tau : \Lambda \to \{0,1\}$, resulting in a marginal distribution $\mu_{G,\lambda}^{(\Lambda,\tau)}$ on $V(G) \setminus \Lambda$. Lemma 10 of [9] quantifies over pairs of pinnings (Λ, τ) and (Λ, τ') where τ and τ' differ at exactly one point. However, the hard-core distribution has a pleasant property: pinning a vertex v to 0 is equivalent to removing v from the graph, while pinning v to 1 is equivalent to removing v and all its neighbours. So, in the case of the hard-core model we need only quantify over single vertex pinnings, since we are restricting attention to graph classes which are closed under taking induced subgraphs.

From our premises, Lemma 10 of [9] tells us that that $\mu_{G,\lambda}$ is 'η-spectrally bounded'. For our purposes we don't need to define this concept as we are merely going to feed this information about $\mu_{G,\lambda}$ into Theorem 9 of [9]. The remaining premiss required for this theorem is that $\mu_{G,\lambda}$ is 'b-marginally bounded'. This just means that $\mathbb{P}(\sigma(v) = 1)$ is bounded away from both 0 and 1 by a margin b. But this is exactly what is guaranteed by Corollary 3, with $b = \min\{\lambda/(1+\lambda)^{\Delta+1}, 1/(1+\lambda)\}$. □

3. $S_{1,1,t}$-free graphs (claw-free and upwards)

Suppose $\Delta \geq 3$ is a degree bound and H a fixed graph. Let G be an H-free (i.e. excluding induced copies of H) connected graph of maximum degree Δ. Chen and Gu [9, Theorem 4] show optimal mixing of Glauber dynamics for the hard-core model on G, for any given fugacity $\lambda > 0$, when the excluded graph H is the claw. In fact, optimal mixing also holds for $H = S_{1,1,t}$ with $t \geq 2$, for a rather dull reason. Define

$$\mathrm{vol}\,(\Delta, t) = 1 + \Delta\,\frac{(\Delta - 1)^t - 1}{\Delta - 2},$$

and note that $\mathrm{vol}\,(\Delta, t)$ bounds the number of vertices in a radius-t ball around any vertex of a graph of maximum degree Δ.

Lemma 5. *Suppose that G is a connected graph of maximum degree Δ, and that $t \geq 2$. If $|V(G)| > \mathrm{vol}\,(\Delta, t+1)$ and G contains an induced claw, then G contains an induced $S_{1,1,t}$.*

Proof. Let G be as in the statement and suppose $\{v, a, b, c\} \subseteq V(G)$ induces a claw centred at v. As $\mathrm{vol}\,(\Delta, t+1)$ is an upper bound on the number of vertices of G within distance $t+1$ of v, there must exist a vertex w at distance $t+2$ from v. Consider a minimum length path from $\{v, a, b, c\}$ to w, which clearly has length at least $t+1$. Let this path be $P = (w_0, w_1, \ldots, w_s)$, where $w_0 \in \{v, a, b, c\}$, $w_s = w$ and $s \geq t+1$. Consider the graph induced by the vertex set $U = \{v, a, b, c, w_1, w_2, \ldots, w_s\}$. By minimality of P, the graph $G[U]$ contains edges $\{\{v, a\}, \{v, b\}, \{v, c\}, \{w_1, w_2\}, \{w_2, w_3\}, \ldots, \{w_{s-1}, w_s\}\}$ together with some non-empty subset A of $\{\{v, w_1\}, \{a, w_1\}, \{b, w_1\}, \{c, w_1\}\}$. The 15 possibilities for A fall into four cases, up to symmetry:

- $A = \{\{v, w_1\}\}$: the set $\{v, a, b, w_1, \ldots, w_t\}$ induces $S_{1,1,t}$;
- $A = \{\{a, w_1\}\}$: the set $\{v, b, c, a, w_1, \ldots, w_{t-1}\}$ induces $S_{1,1,t}$;
- $A = \{\{v, w_1\}, \{a, w_1\}\}$: the set $\{v, b, c, w_1, \ldots, w_t\}$ induces $S_{1,1,t}$;
- $A \supseteq \{\{a, w_1\}, \{b, w_1\}\}$: the set $\{w_1, a, b, w_2, \ldots, w_{t+1}\}$ induces $S_{1,1,t}$.

The cases are exhaustive, completing the proof. □

---

[1]The quoted results are stated for edge-based configurations rather than vertex-based, but this is inessential.

- Let R be a random 'red' independent set from $\mu_{G,\lambda}^{(v,1)}$.
- Let B be a random 'blue' independent set from $\mu_{G,\lambda}^{(v,0)}$, with B independent of R.
- Consider the subgraph $G[R \cup B]$ of G induced by the union of the two independent sets. Let $C \subseteq V(G)$ be the connected component of $G[R \cup B]$ containing v. We refer to C as a cluster.
- Modify the blue independent set by setting $B \leftarrow (B \cap C) \cup (R \setminus C)$. Thus, B is made to agree with B inside C and with R outside. The independent set R remains unchanged.
- (R, B) is the required coupling.

Figure 3. The coupling of van den Berg and Brouwer. Here, G is a graph, $v \in V(G)$, and $\lambda > 0$.

Corollary 6. *Restricted to connected graphs of maximum degree Δ, the class of $S_{1,1,t}$-free graphs is a finite variation of the class of claw-free graphs, for any $t \geq 2$.*

Corollary 7. *For any $\Delta \geq 3$ any $t \geq 1$ and any $\lambda > 0$, Glauber dynamics for the hard-core model at fugacity λ has mixing time $O(n \log n)$ on the class of $S_{1,1,t}$-free graphs of maximum degree Δ.*

Proof. This is immediate from [9, Theorem 4] and Corollary 6. □

4. $S_{1,2,t}$-free graphs (E-free and upwards)

Before heading off to larger graph classes it is instructive to go back to a precursor of [9]. Van den Berg and Brouwer [26] construct an explicit coupling for matchings in a bounded-degree graph that (in modern terms) bounds spectral independence.[2] This almost immediately gives optimal mixing of Glauber dynamics for the monomer-dimer model at all temperatures. (Of course, van den Berg, and Brouwer did not have the spectral independence technology, so they were only able to deduce low-degree polynomial mixing time.) Their method extends easily to independent sets in claw-free graphs (though this was not stated in their paper). In this section, we rederive the result of Chen and Gu, as the proof will provide a base from which we may progress to large subdivided claws.

The coupling of van den Berg and Brouwer, translated to independent sets, is given in Fig. 3.

Lemma 8. *Let G be a connected claw-free graph, $v \in V(G)$ be a vertex, and $\lambda > 0$. Then the coupling of Figure 3 is correct: R is a sample from $\mu_{G,\lambda}^{(v,1)}$ and B is a sample from $\mu_{G,\lambda}^{(v,0)}$.*

Proof. Since the independent set R was never modified, it is clearly a sample from $\mu_{G,\lambda}^{(v,1)}$.

For B, we argue as follows. Think of the pair R, B being selected as follows. First we select the cluster $C \subseteq V(G)$, as if it had been formed from a pair of independent sets R, B as above. Since $G[C]$ is a connected bipartite graph, and $v \in R$, we can reconstruct the portion of R and B lying within C. Denote by

$$\partial C = \{w \in V(G) \setminus C : \{w, u\} \in E(G) \text{ for some } u \in C\}$$

the boundary of C. A subset $R_{\text{ext}} \subseteq V(G) \setminus C$ extends $R_{\text{int}} = R \cap C$ to an independent set $R = R_{\text{int}} \cup R_{\text{ext}}$ on the whole of $V(G)$ if and only if R_{ext} is an independent set of the graph $G[V(G) \setminus (C \cup \partial C)]$. (A vertex in R_{ext} certainly cannot be adjacent to a red vertex in the cluster C, and it cannot be adjacent to a blue vertex by maximality of C.) Exactly the same is true of extensions B_{ext} of $B_{\text{int}} = B \cap C$. In short, the set of possible extensions to B_{int} and R_{int} are identical. Informally, R_{ext} is equally valid as an extension of B_{int} as B_{ext} is. Thus $B = B_{\text{int}} \cup R_{\text{ext}}$ is a true sample from $\mu_{G,\lambda}^{(v,0)}$. □

---

[2]Although this reference is closest to our application, the general concept of 'disagreement percolation' appears earlier in work of van den Berg and Maes [27] and van den Berg and Steif [28].

Algorithm 1 Growing a red-blue cluster by Breadth-First Search (BFS).

Cluster BFS(v)   // R and B are random independent sets, revealed as the algorithm progresses.
$L_0 \leftarrow \{v\}$
$U \leftarrow V(G) \setminus \{v\}$ // the set of unprocessed vertices
$d \leftarrow 0$ // current depth of search
while $L_d \neq \emptyset$ **do**
   $N \leftarrow U \cap \partial L_d$ // $\partial L_d = \{u \in V(G) \setminus L_d : \{u, w\} \in E(G) \text{ for some } w \in L_d\}$
   if d is even **then**
      reveal the blue vertices within N, and set $L_{d+1} \leftarrow N \cap B$
   else
      reveal the red vertices within N, and set $L_{d+1} \leftarrow N \cap R$
   end if
   $U \leftarrow U \setminus N$
   $d \leftarrow d + 1$
end while
$d \leftarrow d - 1$ // The red-blue cluster containing v is $C = L_0 \cup L_1 \cup \cdots \cup L_d$

Using this coupling we can rederive the result of Chen and Gu.

Corollary 9. *Suppose that $\Delta \geq 3$ and $\lambda > 0$. If G is a claw-free graph of maximum degree Δ and $v \in V(G)$, then $W_1(\mu_{G,\lambda}^{(v,0)}, \mu_{G,\lambda}^{(v,1)}) = O(1)$.*

Proof. The claimed bound on W_1 is evidenced by the coupling just analysed. As $R \oplus B = C$, we just have to demonstrate that $|C|$ is small in expectation. Note that $G[C]$ is bipartite. It is straightforward to see that a connected bipartite claw-free graph can have no vertices of degree 3 or greater, so is either a path or a cycle. As paths and cycles are one-dimensional structures, we do not expect them to propagate far, and this is indeed the case. Imagine that we reveal the colours of vertices in a breadth-first fashion, starting with v at time 0. The breadth-first search (BFS) is presented as Algorithm 1. This is a little heavy for current purposes, but its generality will be useful later. As the cluster is either a path or a cycle, we have $|L_j| \leq 2$ for all levels j. Since the hard-core distribution is dominated by the product of independent Bernoulli distributions (Lemma 2), on each iteration, with probability at least $p = (1/(\lambda + 1))^{2\Delta}$, the BFS will halt. So $\mathbb{E}|C| = \mathbb{E}[|R \oplus B|]$ is bounded above by the expectation of a random variable having a geometric distribution with success probability $(\lambda + 1)^{-2\Delta}$, which is $(\lambda + 1)^{2\Delta}$. □

Then, by Corollary 4 we obtain:

Corollary 10. *For any $\Delta \geq 3$ and $\lambda > 0$, Glauber dynamics at fugacity λ has optimal (i.e. $O(n \log n)$) mixing time on the class of claw-free graphs of maximum degree Δ.*

Let's make some modest steps towards excluding larger subdivided claws.[3] Fork-free graphs were characterised by Alekseev [2]. See also Bencs [8]. A connected bipartite fork-free graph is either a path, a cycle, or a complete bipartite graph minus a (possibly empty) matching. Since the last of these can have size at most $2(\Delta + 1)$, the argument just given in the Proof of Lemma 8 for claw-free graphs carries over to fork-free graphs, as the expected size of a cluster is still constant. However, as we already saw in Corollary 6, this generalisation is illusory.

In light of Corollary 6, the first interesting case beyond claw-free is E-free (i.e. $S_{1,2,2}$-free). Fig. 4 illustrates an infinite family of E-free graphs which are not claw-free. Indeed, by iterating the basic building block either horizontally or vertically, or both, we can appreciate that the family includes

---

[3]The remainder of the section is not essential to the technical development of the main result, Theorem 17, relating to general subdivided claws. However, it contains some contextual material on graphs with excluded subdivided claws, and also motivates the approach taken in Section 5.
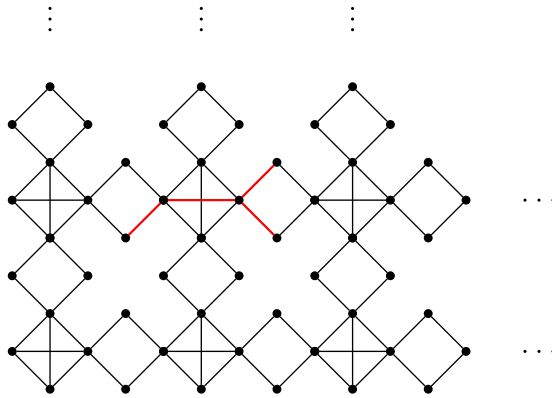
Figure 4. An (infinite family of) E-free but not fork-free graph(s).

graphs that contain induced copies of $S_{1,1,t}$ for arbitrarily large t. (An instance of a fork is picked out in red in Fig. 4.) So, by excluding E, we have escaped from the constraints of Corollary 6 into new, apparently interesting families of graphs.

It is worth recording the fact that there is no analogue of Corollary 6 for the sequence of $S_{1,2,t}$-free graph classes. For example, suppose we construct a graph G as a disjoint union of an E together with a path of arbitrary length, with endpoints a and b, and then add edges from a to all vertices in the E. We claim that G is skew-star-free (i.e. $S_{1,2,3}$-free). Suppose to the contrary that G contains an induced skew star. This induced graph would have to include vertex a together with at least one edge e from the E. But the endpoints of e together with a induce a triangle, which is a contradiction. So there are infinitely many skew-star-free graphs of degree at most 7 that are not E-free.

At this point we skip the class of E-free graphs and proceed directly to the larger class of skew-star-free graphs as these have been analysed in detail by Lozin [22] in his 'Decomposition Theorem'. He showed that a connected bipartite skew-star-free graph is of one of four kinds. Three of these are of bounded size, while the fourth is composed of blow-ups of paths and cycles. (A blow-up of a graph H is obtained by replacing each vertex of H by an independent set, and each edge by a complete bipartite graph connecting the corresponding independent sets.) As the unbounded components are still 'linear', we continue to have spectral independence and optimal mixing of Glauber dynamics. The only change to the proof of Corollary 9 lies in the analysis of the level sets L_i generated by the BFS: we now have $|L_i| \leq 2(\Delta - 1)$ rather than $|L_j| \leq 2$.

Corollary 11. *Suppose that $\Delta \geq 3$ and $\lambda > 0$. If G is a skew-star-free graph of maximum degree Δ and $v \in V(G)$, then $W_1(\mu_{G,\lambda}^{(v,0)}, \mu_{G,\lambda}^{(v,1)}) = O(1)$.*

Corollary 12. *For any $\Delta \geq 3$ and $\lambda > 0$, Glauber dynamics at fugacity λ has optimal (i.e. $O(n \log n)$) mixing time on the class of skew-star-free (and hence E-free) graphs of maximum degree Δ.*

As we saw earlier, there is no direct analogue of Lemma 5 with $S_{1,2,2}$ replacing $S_{1,1,1}$, and $S_{1,2,t}$ replacing $S_{1,1,t}$. But if we add the condition of bipartiteness we do have:

Lemma 13. *Suppose that G is a connected bipartite graph of maximum degree Δ, and that $t \geq 3$. If $|V(G)| > \mathrm{vol}\,(\Delta, t + 2)$ and G contains an induced $S_{1,2,2}$, then G contains an induced $S_{1,2,t}$.*

This may be proved in a similar way to Lemma 5, but we omit the proof as it is not on the direct line to the main result.

Corollary 14. *For any $\Delta \geq 3$, $\lambda > 0$, and $t \geq 2$, Glauber dynamics at fugacity λ has optimal (i.e. $O(n \log n)$) mixing time on the class of $S_{1,2,t}$-free graphs of maximum degree Δ.*

Sketch proof. The cluster containing v is bipartite, and hence, by Lemma 13 and Lozin's decomposition theorem, is either of bounded size or is the blow-up of a path or cycle. The layers L_i created by the BFS of Algorithm 1 are of bounded size, so the depth of the search is again bounded by a geometric random variable. □

We have seen that in the case of $S_{1,1,t}$-free graphs and $S_{1,2,t}$-free graphs the cluster generated by Algorithm 1 is either of bounded size, or is 'linear' (in this case a blow-up of either a path or a cycle). We may suspect that this remains true for graphs excluding larger subdivided stars. While this is morally true, it is difficult to provide a suitable definition of 'linear' that meshes with the coupling argument. We therefore slightly change our approach.

5. $S_{2,2,2}$-free graphs and beyond

Instead of trying to analyse the structure of bounded-degree bipartite graphs excluding a large subdivided claw, we will analyse the BFS procedure of Algorithm 1 directly. Recall the layers L_i that are produced by the search. The key observation is that these layers do not grow large.

Lemma 15. *If G is $S_{t,t,t}$-free then $|L_i| \leq 2 \operatorname{vol}(\Delta, 2t)$, for all $i \in \mathbb{N}$ for which L_i is defined.*

Proof. Suppose to the contrary that $|L_i| > 2 \operatorname{vol}(\Delta, 2t)$. Choose vertices $a, b, c \in L_i$ with the property that $d_G(a, b)$, $d_G(a, c)$ and $d_G(b, c)$ are all greater than $2t$. This is always possible: select a arbitrarily, and eliminate from consideration the at most $\operatorname{vol}(\Delta, 2t)$ vertices in L_i that are within distance $2t$ from a; choose b from the remaining vertices and eliminate the vertices in L_i that are within distance $2t$ from b; at least one choice for c remains.

Now orient all edges away from v, so that each edge goes from level L_j to L_{j+1} for some j. Choose a minimum cardinality directed Steiner tree T connecting v (as root) to all of a, b, c. Consider the graph T' induced by $V(T)$. Suppose that T' has some edge not in $E(T)$, say $e = \{u_j, u_{j+1}\}$ where $u_j \in L_j$ and $u_{j+1} \in L_{j+1}$. (Recall that T' is bipartite, so there are no edges within a layer; also the levels were constructed by BFS, so no edges of T' skip a layer.)

First assume that T has a single vertex w such that either $w \neq v$ and $\deg(w) = 4$ or $w = v$ and $\deg(w) = 3$. Note that all vertices of T other than v, w, a, b, c have degree 2. By adding e to $E(T)$ and removing the edges of T lying between w and u_{j+1} we would obtain a smaller tree, contradicting minimality of T. (Notice that we must remove at least two edges.) So such an edge e does not exist. By choice of a, b, c we know that $d_T(w, a), d_T(w, b), d_T(w, c) \geq t + 1$, so G contains an induced $S_{t,t,t}$. (Note that we are measuring graph distances within the tree T here.)

Next assume that T has two vertices w and w', with w' in a higher numbered level than w, and such that either (i) $w \neq v$ and and $\deg(w) = \deg(w') = 3$ or (ii) $w = v$ and $\deg(w) = 2$ and $\deg(w') = 3$. Note that all vertices of T other than v, w, w', a, b, c have degree 2. Suppose, without loss of generality, that a and b are descendants of w'. With two exceptions, we may add e to $E(T)$ and remove all the edges between u_{j+1} and w, or w' as appropriate, to obtain a smaller tree, contradicting minimality of T. Those exceptions occur when $w' \in L_j$, that is, u_j is in the same level as w'. In that case, one edge is added and one removed and the size of the tree remains unchanged. There are two candidate vertices for u_{j+1}; these are the vertices in $L_{j+1} \cap V(T)$, say x and y, that are adjacent to w'. Let x (respectively, y) be the vertex on the path in T from w' to a (respectively b). To summarise: The graph induced by $V(T)$ is tree T, with the possible addition of one or two of the edges $\{u_j, x\}$ and $\{u_j, y\}$. (Refer to Fig. 5.)

By choice of a, b, c we know that $d_T(w', a), d_T(w', b) \geq t + 1$. In the case that neither of the exceptional edges is present, we have an induced $S_{t,t,t}$ with its centre at w' and with the path from w' to c routing via w. In the case of one exceptional edge, say $\{u_j, x\}$, we have an induced $S_{t,t,t}$ with its centre at x and with the path from x to b routing via w', and the path from x to c routing via u_j. In the case of two exceptional edges, we have an induced $S_{t,t,t}$ with its centre at u_j and with the path from u_j to a routing via x, and the path from u_j to b routing via y.

As the above cases are exhaustive, we obtain a contradiction to G being $S_{t,t,t}$-free. □
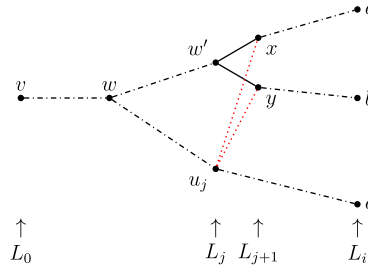
Figure 5. The two exceptional edges (shown red, dotted). The solid lines are edges and the dot-dashed lines are paths.

Corollary 16. *Suppose that $\Delta \geq 3$, $\lambda > 0$, and $t \geq 2$. If G is a $S_{t,t,t}$-free graph of maximum degree Δ and $v \in V(G)$, then $W_1(\mu_{G,\lambda}^{(v,0)}, \mu_{G,\lambda}^{(v,1)}) = O(1)$.*

Proof. The levels of the BFS remain bounded. There is a constant probability of the search terminating at each iteration. □

Theorem 17. *For any $\Delta \geq 3$, $\lambda > 0$, and $t \geq 2$, Glauber dynamics for the hard-core distribution at fugacity λ has optimal (i.e. $O(n \log n)$) mixing time on the class of $S_{t,t,t}$-free graphs of maximum degree Δ.*

Proof. Follows immediately using Corollary 4. □

Naturally, the theorem also covers $S_{i,j,k}$-free graphs by taking $t = \max\{i, j, k\}$.

6. Exponential-time mixing of *H*-free graphs when *H* is neither a path nor a subdivided claw

To complete the dichotomy, it only remains to show:

Theorem 18. *Let H be a connected graph that is neither a subdivided claw nor a path. There exists a fugacity λ and an infinite family $\mathcal{G}$ of H-free graphs of maximum degree 3 such that Glauber dynamics for the hard-core distribution on these graphs has exponential mixing time on $\mathcal{G}$.*

Proof. Let $\alpha > 0$ be sufficiently small. Given n, let G be a cubic bipartite graph on $n + n$ vertices that is a $1 + \alpha$ expander for sets up to size $\frac{2}{3}n$. By this we mean the following. Let $V_L \cup V_R = V(G)$ be the bipartition of the vertex set of G. Then, every subset $S \subset V_L$ of size at most $\frac{2}{3}n$ is adjacent to at least $(1 + \alpha)|S|$ vertices in V_R, and the same is true with V_L and V_R interchanged. Such graphs exist for some fixed $\alpha > 0$ by a lemma due to Bassalygo [7]; see Alon [3, Lemma 4.1]. Let ℓ be a number and let G^* denote the $(2\ell + 1)$-stretch of G; that is, G^* is obtained from G by subdividing each edge by 2ℓ new vertices. Fix ℓ sufficiently large that G^* contains no induced H. Let V_L and V_R continue to denote the vertices in G^* that are inherited from G.

Let $\mathcal{I}_< \cup \mathcal{I}_= \cup \mathcal{I}_> = \mathcal{I}_G$ be the partition of the independent sets $\mathcal{I}_G$ of G given by:

$$\mathcal{I}_\circ = \{I \in \mathcal{I} : |I \cap V_L| \circ |I \cap V_R|\},$$

where $\circ \in \{<, =, >\}$. Note that, to pass from any independent set in $\mathcal{I}_<$ to any independent set in $\mathcal{I}_>$, Glauber dynamics must pass through an independent set in $\mathcal{I}_=$. To show an exponential lower bound on mixing time it is enough to verify that $\mathrm{wt}(\mathcal{I}_=) \leq e^{-cn} \min\{\mathrm{wt}(\mathcal{I}_<), \mathrm{wt}(\mathcal{I}_>)\}$ for some $c > 0$.[4] This inequality implies that the 'conductance' of Glauber dynamics on G^* is exponential small, and hence that the mixing time is exponentially large [20, Claim 7.14].

Note that G^* is bipartite, with a balanced bipartition, and having $|V(G^*)| = (6\ell + 2)n$ vertices. By selecting all vertices on one side of the bipartition we obtain an independent set of size

---

[4] Here, we have extended $\mathrm{wt}(\cdot)$ from configurations to sets of configurations in the obvious way.

$(3\ell + 1)n$ in $\mathcal{I}_<$, and by selecting all vertices in the opposite side an independent set of the same size in $\mathcal{I}_>$. Just considering these two independent sets we see that

$$\mathrm{wt}(\mathcal{I}_<), \mathrm{wt}(\mathcal{I}_>) \geq \lambda^{(3\ell+1)n}.$$

We will show that the independent sets in $\mathcal{I}_=$ all have cardinality significantly less than $(3\ell + 1)n$. So if λ is sufficiently large, the total weight $\mathrm{wt}(\mathcal{I}_=)$ of independent sets in $\mathcal{I}_=$ will be exponentially smaller than $\mathrm{wt}(\mathcal{I}_<)$. So consider an independent set I with $|I \cap V_L| = |I \cap V_R| = k$. In general, each of the paths of length $2\ell + 1$ in G^* joining a vertex u in V_L to a vertex v in V_R is able to support an independent set of cardinality ℓ on its internal vertices (the vertices that are neither in V_L nor in V_R). However, if u and v are both in I then the internal vertices can only support an independent set of size $\ell - 1$.

By expansion, if $n/(2 + \alpha) \leq k \leq \frac{2}{3}n$, there must be at least $(2 + \alpha)k - n$ of these deficient paths. Thus, when $0 \leq k \leq \frac{2}{3}n$, the cardinality of an independent set I in G^* with $|I \cap V_L| = |I \cap V_R| = k$ is at most

$$|I| = 2k + 3\ell n - \max\{(2 + \alpha)k - n, 0\}.$$

The maximum of this expression is achieved at $k = n/(2 + \alpha)$ where it takes the value $(3\ell + 2/(2 + \alpha))n$. When $k > \frac{2}{3}n$ there is a simpler argument. There are $3k$ edges of G leaving the set $I \cap V_L$ and at most $3(n - k)$ of these can be accommodated in $V_R \setminus I$. So there are at least $6k - 3n$ deficient paths and the cardinality of an independent set In G^* in this range is at most $2k + 3\ell n - (6k - 3n) = (3\ell + 3)n - 4k$. This expression is maximised at $k = \frac{2}{3}n$, at which point it evaluates to $(3\ell + \frac{1}{3})n$. Over the whole range of k, the maximum cardinality of an independent set in $\mathcal{I}_=$ is bounded by $(3\ell + 2/(2 + \alpha))n$.

As $|V(G^*)| = (6\ell + 2)n$, the total weight of independent sets in $\mathcal{I}_=$ is at most

$$\mathrm{wt}(\mathcal{I}_=) \leq 2^{(6\ell+2)n}\lambda^{(3\ell+2/(2+\alpha))n},$$

and hence the ratio of $\mathrm{wt}(\mathcal{I}_=)$ to $\mathrm{wt}(\mathcal{I}_<)$ is at most

$$\frac{\mathrm{wt}(\mathcal{I}_=)}{\mathrm{wt}(\mathcal{I}_<)} \leq \left(2^{6\ell+2}\lambda^{2/(2+\alpha)-1}\right)^n.$$

By setting λ sufficiently large in terms of ℓ (which in turn is determined by H) we can ensure that the above ratio is at most 2^{-n}. Thus, Glauber dynamics on G^* has exponentially small conductance, and exponentially large mixing time. □

Proof of Theorem 1. The positive direction is immediate from Theorem 17. Now choose ℓ sufficiently large such that the stretched graph G^* in the proof of Theorem 18 contains none of the graphs $H \in \mathcal{H}$. The construction of Theorem 18 yields a fugacity λ and an infinite family of graphs (both depending on ℓ) on which Glauber dynamics requires exponential mixing time at fugacity λ. □

Acknowledgements

References

[1] Abrishami, T., Chudnovsky, M., Dibek, C. and Rzążewski, P. (2022) Polynomial-time algorithm for maximum independent set in bounded-degree graphs with no long induced claws. *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA).* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, pp. 1448–1470,

[2] Alekseev, V. E. (2004) Polynomial algorithm for finding the largest independent sets in graphs without forks, Russian translations. *Discrete Appl. Math.* **135**, 1–3, 3–16. Russian translations. II.

[3] Alon, N. (1986) Eigenvalues and expanders. *Combinatorica* **6**(2) 83–96. Theory of computing (Singer Island, Fla., 1984).

[4] Anari, N., Liu, K. and Gharan, S. O. (2024) Spectral independence in high-dimensional expanders and applications to the hardcore model. *SIAM J. Comput* **53**(6) FOCS20–1–FOCS20-37.

[5] Anari, N., Liu, K., Gharan, S. O. and Vinzant, C. (2024) Log-concave polynomials II: High-dimensional walks and an FPRAS for counting bases of a matroid. *Ann. Math.* **199**(1) 259–299.

[6] Barvinok, A. (2016) Computing the permanent of (some) complex matrices. *Found. Comput. Math.* **16**(2) 329–342.

[7] Bassalygo, L. A. (1981) Asymptotically optimal switching circuits. *Probl. Peredachi Informat.* **17**(3) 81–88.

[8] Bencs, F. (2019) Graphs, Groups and Measures (*PhD thesis*). Central European University.

[9] Chen, Z. and Gu, Y. (2024) Fast sampling of *b*-matchings and *b*-edge covers. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM, Philadelphia, pp. 4972–4987.

[10] Chen, Z., Liu, K. and Vigoda, E. (2022) Spectral independence via stability and applications to Holant-type problems. *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science – FOCS 2021*, IEEE Computer Society, Los Alamitos, pp. 149–160.

[11] Chen, Z., Liu, K. and Vigoda, E. (2023) Rapid mixing of Glauber dynamics up to uniqueness via contraction. *SIAM J. Comput.* **52**(1) 196–237.

[12] Chen, Z., Liu, K. and Vigoda, E. ([2021] ©2021) Optimal mixing of Glauber dynamics: entropy factorization via high-dimensional expansion. In *STOC '21–Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. ACM, New York, pp. 1537–1550.

[13] Chudnovsky, M., Pilipczuk, M., Pilipczuk, M. and Thomassé, S. (2024) Quasi-polynomial time approximation schemes for the maximum weight independent set problem in *H*-free graphs. *SIAM J. Comput.* **53**(1) 47–86.

[14] Chudnovsky, M. and Seymour, P. (2007) The roots of the independence polynomial of a clawfree graph. *J. Combin. Theory Ser. B* **97**(3) 350–357.

[15] Cryan, M., Guo, H. and Mousa, G. (2021) Modified log-Sobolev inequalities for strongly log-concave distributions. *Ann. Probab.* **49**(1) 506–525.

[16] Dyer, M., Heinrich, M., Jerrum, M. and Müller, H. (2021) Polynomial-time approximation algorithms for the antiferromagnetic Ising model on line graphs. *Combin. Probab. Comput.* **30**(6) 905–921.

[17] Dyer, M., Jerrum, M., Müller, H. and Vusković, K. (2021) Counting weighted independent sets beyond the permanent. *SIAM J. Discrete Math.* **35**(2) 1503–1524.

[18] Friedli, S. and Velenik, Y. (2018) *Statistical mechanics of lattice systems*. Cambridge University Press, Cambridge. A concrete mathematical introduction.

[19] Hayes, T. P. and Sinclair, A. (2007) A general lower bound for mixing of single-site dynamics on graphs. *Ann. Appl. Probab.* **17**(3) 931–952.

[20] Jerrum, M. (2003) Counting, sampling and integrating: algorithms and complexity. *Lectures in Mathematics ETH Zürich*. Basel: Birkhäuser Verlag.

[21] Levin, D. A., Peres, Y. and Wilmer, E. L. (2009) *Markov chains and mixing times, With a chapter by James G.* American Mathematical Society, Propp and David B. Wilson, Providence

[22] Lozin, V. V. (2002) Bipartite graphs without a skew star. *Discrete Math.* **257**(1) 83–100.

[23] Matthews, J. (2008) Markov Chains for Sampling Matchings. *PhD thesis*. School of Informatics, University of Edinburgh. http://hdl.handle.net/1842/3072.

[24] Patel, V. and Regts, G. (2017) Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM J. Comput.* **46**(6) 1893–1919.

[25] Sly, A. (2010) Computational transition at the uniqueness threshold. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science – FOCS 2010*. IEEE Computer Society, Los Alamitos, pp. 287–296.

[26] van den Berg, J. and Brouwer, R. (2000) Random sampling for the monomer-dimer model on a lattice, Probabilistic techniques in equilibrium and nonequilibrium statistical physics. *J. Math. Phys.* **41**(3) 1585–1597.

[27] van den Berg, J. and Maes, C. (1994) Disagreement percolation in the study of Markov fields. *Ann. Probab.* **22**(2) 749–763.

[28] van den Berg, J. and Steif, J. E. (1994) Percolation and the hard-core lattice gas model. *Stochastic Process. Appl.* **49**(2) 179–197.

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