Mixing of Markov Chains for Independent Sets on Chordal Graphs with Bounded Separators

Ivona Bezáková and Wenbo Sun

Rochester Institute of Technology, Rochester NY 14623, USA

Abstract. We prove rapid mixing of well-known Markov chains for the hardcore model on a new graph class, the class of chordal graphs with a bound on minimal separator size. In the hardcore model, for a given graph G and a fugacity parameter $\lambda \in \mathbb{R}^+$, the goal is to produce an independent set S of G with probability proportional to $\lambda^{|S|}$. In general graphs and arbitrary λ, producing a sample from this distribution in polynomial time is provably difficult. However, natural Markov chains converge to the correct distribution for any graph, leading to the study of their mixing times for different graph classes. Rapid mixing for graphs of bounded degrees and a range of λs dependent on the maximum degree has attracted attention since the 1990's. Recent results showed rapid mixing for arbitrary λ and two other classes of graphs: graphs of bounded treewidth and graphs of bounded bipartite pathwidth. In this work, we extend these results by showing rapid mixing in a new graph class, class of chordal graphs with bounded minimal separators. Graphs in this class have no bound on the vertex degrees, the treewidth, or the bipartite pathwidth. Similarly to the results dealing with bounded treewidth and with bounded bipartite pathwidth, we prove rapid mixing using the canonical paths technique. However, unlike in the previous works, we need to process the data using a non-linear, tree-like, approach.

1 Introduction

Independent sets, that is, sets of vertices in a graph without any edges between them, are heavily studied in computer science and other fields. Among their many applications is the hardcore model of a gas in statistical physics, where the goal is to sample independent sets of a given graph according to a specific probability distribution. In particular, for a given parameter (also known as fugacity) $\lambda \in \mathbb{R}^+$, the goal is to generate an independent set S with probability proportional to $\lambda^{|S|}$.
Markov chains have attracted attention as a sampling technique for the hardcore distribution since the late 1990's. The mixing time of a Markov chain is the time it takes to converge to its stationary distribution, and a Markov chain is said to be rapidly mixing if its mixing time is polynomial in the size of the input. Luby and Vigoda [15] showed rapid mixing of a natural insert/delete chain (a single-site Glauber dynamics) for independent sets of triangle-free graphs with degree bound Δ and $\lambda < 2/(\Delta - 2)$, which was soon extended by Vigoda [23] to general graphs with degree bound Δ. Independently, Dyer and Greenhill [4]

analyzed an insert/delete chain with an added drag transition, showing rapid mixing for the same graph class and range of λs. All these works used the coupling technique to obtain their mixing results.

Dyer, Frieze, and Jerrum [6] established a hardness result, showing that even for $\lambda = 1$, no Markov chain for sampling independent sets that changes only a "small" number (that is, a linear fraction) of vertices per step mixes rapidly for general graphs, even if the maximum degree is six. This was followed by an influential non-Markov-chain-based approach of Weitz [24] which leads to rapid mixing of the Glauber dynamics for subexponentially growing graphs with maximum degree Δ and $\lambda < \lambda_c := (\Delta - 1)^{\Delta-1}/(\Delta - 2)^{\Delta}$. Efthymiou et al. [7] used belief propagation to obtain rapid mixing for graphs with sufficiently large maximum degree and girth (that is, the length of the smallest cycle) ≥ 7, and $\lambda < \lambda_c$. Very recently, Anari, Liu, and Oveis Gharan [1] established rapid mixing for bounded degree graphs and $\lambda < \lambda_c$ using a new notion of spectral independence. We refer the reader to [1] for an overview of rapid mixing results for restricted graph classes with bounded degrees. On the complementary hardness side, assuming $NP \neq RP$, a celebrated result of Sly [20], together with [17, 21, 8, 9], imply hardness of polynomial-time approximate sampling for bounded degree graphs and $\lambda > \lambda_c$.

Beyond graphs of bounded degrees, Bordewich and Kang [2] studied an insert/delete Markov chain (a multi-site Glauber dynamics) to sample vertex subsets, a generalization of the hardcore model, and proved that its mixing time is $n^{O(\mathrm{tw})}$ for an arbitrary λ and n-vertex graphs of treewidth tw. Recently, generalizing work of Matthews [16] on claw-free graphs, Dyer, Greenhill, and Müller [5] introduced a new graph parameter, the bipartite pathwidth, obtaining a mixing time of $n^{O(p)}$ for the insert/delete chain for an arbitrary λ and graphs with bipartite pathwidth bounded by p. These works used the canonical paths technique [13] to prove rapid mixing.

We extend this line of work to another graph class, the class of chordal graphs with bounded minimum separator size. In particular, we obtain a mixing time of $O(n^{O(\log b)})$ for arbitrary λ and chordal graphs with minimal separators of size at most b (or, equivalently, bound b on the intersection size of any pair of maximal cliques). Graphs in this class have no bound on their degrees, treewidth, or bipartite pathwidth[1]. Chordal graphs, where each cycle of length at least four has a chord, are a widely studied graph class, playing an important role in many real-world applications such as inference in probabilistic graphical models [14].

We also use the canonical paths technique but we need to overcome the "non-linearity" of our data. The technique relies on finding a Markov chain path (a canonical path) between every pair of states in such a way that no transition gets overloaded (congested). This is typically done by considering the symmetric difference of the two states (a pair of independent sets), gradually removing one vertex from the initial independent set while adding a vertex from the final independent set. If the symmet-

[1] This can be seen by taking a complete binary tree of $\sqrt{n}$ vertices, where each vertex is replaced by a clique of size $\sqrt{n}$, and each pair of adjacent cliques is connected by an edge.

ric difference induces a collection of paths in the original graph, we can "switch" each path from initial to final starting at one end-point of the path and gradually going to the other end-point, never violating the independent set property. Bounded bipartite pathwidth guarantees that the symmetric difference can be viewed as "wider" paths, as does bounded treewidth due to its relation to the (non-bipartite) pathwidth. However, for our graphs, the symmetric difference is tree-like, which leads to the need to recursively "switch" entire subtrees from initial to final before being able to process the root vertex from the final independent set. Due to this "tree-like" process we also need to overcome corresponding complications in the analysis of the congestion.

We note that if one is interested purely in sampling independent sets from the hardcore distribution on chordal graphs, just like for graphs of bounded treewidth, polynomial-time sampling algorithms exist: Okamoto, Uno, and Uehara [18] designed polynomial-time dynamic programming algorithms to count independent sets, maximum independent sets, and independent sets of fixed size on chordal graphs without any restrictions, which can then be used to sample independent sets from the hardcore distribution. In contrast, our work contributes to the understanding of the conditions under which the well-studied and easy-to-implement insert/delete(/drag) Glauber dynamics Markov chain mixes rapidly.

2 Preliminaries

For an undirected graph G, an *independent set* is a set of vertices $S \subseteq V(G)$ such that there is no edge $(u, v) \in E(G)$ with $u, v \in S$. Let Ω_G be the set of all independent sets of G. We study the problem of sampling independent sets from Ω_G, where the probability distribution is parameterized by a given constant $\lambda > 0$ as follows: A set $S \in \Omega_G$ is to be generated with probability $\pi(S) := \lambda^{|S|}/Z_G(\lambda)$, where the normalization factor $Z_G(\lambda) := \sum_{S\in\Omega_G} \lambda^{|S|}$ is known as the partition function.

We will use Markov chains to obtain a *fully polynomial almost uniform sampler (FPAUS)* from the target distribution π: For a given $\epsilon \in (0, 1)$, we will produce a random element (a sample) from Ω_G chosen from a distribution μ that is ϵ-close to π. In particular, $d_{TV}(\mu, \pi) \leq \epsilon$, where $d_{TV}(\mu, \pi) := \sum_{S\in\Omega_G} |\mu(S) - \pi(S)|/2$ is the *total variation distance*. The sample will be provided in time polynomial in $|V(G)|$ and $\log(1/\epsilon)$.

Next we briefly review chordal graphs and Markov chains.

Chordal graphs

An undirected graph is *chordal* if every cycle of four or more vertices has a chord, that is, an edge that connects two vertices of the cycle but is not part of the cycle. For a nice treatment of chordal graphs, we refer the reader to [22]; in this section we briefly describe the concepts relevant to our work. Every chordal graph G has a *clique tree* T_G which satisfies the following conditions:

(i) T_G is a tree whose vertices are maximal cliques in G and
(ii) T_G has the *induced subtree property*: For every vertex $v \in V(G)$, the maximal cliques containing v form a subtree of T_G. We refer to this tree as the v-induced subtree $T_G(v)$.

In fact, this is a complete characterization of chordal graphs: such a clique tree exists if and only if G is chordal [3, 10]. We note that in some literature the term "clique tree" refers only to the first condition above — in this text by a "clique tree" we mean a clique tree with the induced subtree property, that is, satisfying both conditions above. It follows that since an edge is a clique of size 2, it is a part of a maximal clique, and, therefore, for each edge $(u, v) \in E(G)$ there is a clique in $V(T_G)$ that contains both u and v. The existence of clique trees implies that the *treewidth* of a chordal graph is one less than the size of its largest clique, and, as such, chordal graphs can have unbounded treewidth.
In any graph, a *vertex separator* is a set of vertices whose removal leaves the remaining graph disconnected, and a separator is *minimal* if it has no subset that is also a separator. Suppose we root a clique tree T_G at an arbitrary clique $R \in V(T_G)$, denoting the rooted tree by T_G^R. Then, each clique $C \in T_G^R$ can be partitioned into a separator set $\mathrm{Sep}(C) = C \cap p(C)$ and a residual set $\mathrm{Res}(C) = C \setminus \mathrm{Sep}(C)$, where $p(C)$ denotes the parent clique of $C \neq R$ in T_G^R and $p(R) := \emptyset$. The induced subtree property implies the following theorem, see, for example [22]:

Theorem 1. *Let T_G^R be an R-rooted clique tree of a graph G.*
- *The separator sets $\mathrm{Sep}(C)$ where C ranges over all non-root cliques of T_G^R, are the minimal vertex separators of G.*
- *For each $v \in V(G)$, there is exactly one clique $C_v \in T_G^R$ that contains v in its residual set, that is, $v \in \mathrm{Res}(C_v)$. In particular, C_v is the root of the v-induced subtree $T_G(v)$ in T_G^R. (Therefore, the other cliques in $T_G(v)$ contain v in their separator sets.)*

In this work we assume, without loss of generality, that the given chordal graph G is connected. We also assume that there exists a constant $b \in \mathbb{N}^+$, which we refer to as the *separator bound*, that upper-bounds the size of every minimal separator of G.

Markov chains
In this section we briefly review Markov chains and the canonical paths technique for bounding their mixing times. For more details we refer the reader to, for example, [12]. A (finite discrete) *Markov chain* is a pair (P, Ω) where Ω denotes the state space and P is its transition matrix: a stochastic matrix of dimensions $|\Omega| \times |\Omega|$, indexed by elements from Ω, where $P(u, v)$ is the probability of transitioning from state $u \in \Omega$ to state $v \in \Omega$. The transition from one state to the next is also referred to as a *step* of the Markov chain. A distribution π on Ω (viewed as a vector) is said to be *stationary* if $\pi P = \pi$. If a Markov chain is so-called ergodic, its stationary distribution is unique and it is the limiting distribution the Markov chain converges to as the number of its steps goes to infinity. In this work we deal with ergodic Markov chains and reserve the symbol π for the stationary distribution. For a start distribution μ on Ω, after t steps the chain is in distribution μP^t. For an ergodic chain, $\lim_{t\to\infty} \mu P^t = \pi$. For Markov chains with exponentially large state space the transition matrix is typically very sparse and not given explicitly but instead described implicitly by an algorithm that, for a current state, describes the random process of getting to the next state.

The *mixing time* of the Markov chain is the number of steps needed for the chain to get ϵ-close to its stationary distribution. For a start state $x \in \Omega$, let μ_x denote the distribution where $\mu_x(x) = 1$ and $\mu_x(y) = 0$ for every $y \in \Omega \setminus \{x\}$. Then, for a given $\epsilon \in (0, 1)$, the mixing time $\tau_x(\epsilon)$ from the state x is the smallest t such that $d_{TV}(\mu_x P^t, \pi) \leq \epsilon$. Therefore, a polynomial mixing time for a polynomially-computable start state provides an FPAUS.

Canonical paths [13, 19] is a technique for bounding the mixing time. The idea is to define, for every pair of states $x, y \in \Omega$, a path $\gamma_{x,y} = (x = z_0, \ldots, z_\ell = y)$ such that (z_i, z_{i+1}) are adjacent states in the Markov chain, that is, $P(z_i, z_{i+1}) > 0$. Let $\Gamma := \{\gamma_{xy} \mid x, y \in \Omega\}$ be the set of all canonical paths. The *congestion* through a transition $e = (u, v)$, where $P(u, v) > 0$, is

$$\varrho(\Gamma, e) := \frac{1}{\pi(u)P(u, v)} \sum_{x,y:\gamma_{xy} \text{ uses } e} \pi(x)\pi(y)\,|\gamma_{xy}| \qquad (1)$$

where $|\gamma_{xy}|$ is the length of the path γ_{xy}. The overall congestion of the paths Γ is defined as $\varrho(\Gamma) := \max_{e=(u,v):P(u,v)>0} \varrho(\Gamma, e)$. The mixing time of the chain is bounded by $\tau_x(\epsilon) \leq \varrho(\Gamma)\left(\ln(\frac{1}{\pi(x)}) + \ln(\frac{1}{\epsilon})\right)$ [19].

3 Rapid mixing for Chordal Graphs with Bounded Separators

Recall that we are given a graph G and a parameter $\lambda \in \mathbb{R}^+$. The most commonly used Markov chain for sampling independent sets is the Glauber dynamics (also known as the Luby-Vigoda chain or the insert/delete chain): Let S be the current independent set. Pick a random vertex $u \in V(G)$. If $u \in S$, remove it from S with probability dependent on λ to maintain the desired target distribution (this probability turns out to be $\frac{1}{1+\lambda}$). If $u \notin S$ and if none of its neighbors are in S, add it to S with probability $\frac{\lambda}{1+\lambda}$. Our polynomial mixing time results hold for the Glauber dynamics but in this work we prove mixing time bounds for a closely related Markov chain by Dyer and Greenhill [4].

The Dyer-Greenhill chain: Let $S \in \Omega_G$ be the current independent set. Pick a vertex u uniformly at random from $V(G)$. Then:

- [Delete $\downarrow$:] If $u \in S$, remove it with probability $\frac{1}{1+\lambda}$.
- [Insert $\uparrow$:] If $u \notin S$ and none of the neighbors of u are in S, add u with probability $\frac{\lambda}{1+\lambda}$.
- [Drag $\leftrightarrow$:] if $u \notin S$ and it has a unique neighbor $v \in S$, add u and remove v.

Let S' be the resulting independent set (if none of the above holds for u, let $S' = S$), which is the next state of the Markov chain. The chain is ergodic with the desired stationary distribution $\pi(S) = \frac{\lambda^{|S|}}{Z_G(\lambda)}$ [4].

3.1 Canonical paths for chordal graphs

From now on we assume that G is a connected chordal graph with n vertices. We will define a canonical path between every pair of independent

sets I ("initial") and F ("final") in G. As is often done in canonical paths construction, we will work only with vertices of $I \oplus F$, the symmetric difference of I and F: we will gradually remove vertices from $I \setminus F$ while adding vertices in $F \setminus I$. (Notice that vertices in $I \cap F$ do not neighbor $I \oplus F$, and hence we do not need to touch them.)
We first observe that the symmetric difference of two independent sets in a chordal graph forms an induced forest.

Lemma 1. *Let G be a chordal graph and let I and F be its two independent sets. Then, the subgraph of G induced by $I \oplus F$ is a forest.*

Proof. Let $H = G[I \oplus F]$ be the subgraph induced by $I \oplus F$. By contradiction, assume that H contains a cycle c. Since G is chordal and H is induced, c must have a chord in H, obtaining a shorter cycle. Applying this argument inductively, H contains a triangle which has a pair of adjacent vertices in I, or in F, a contradiction with $I, F \in \Omega_G$. $\square$

We assume that the vertices of G are labeled $1, \ldots, n$. Before defining our canonical paths, we fix a clique tree T_G corresponding to G and we root it at a vertex R (for example, let R be the clique that has vertex 1 in its residual set), obtaining T_G^R. For a vertex u in $V(G)$, let C_u be the clique of T_G^R that contains u in its residual set. We define the *depth* of u in T_G^R as $d(u) := d(C_u)$, where $d(C_u)$ is the depth of C_u in T_G^R (that is, $d(C_u)$ is the distance of C_u from the root R).
For a pair $I, F \in \Omega_G$, we define the canonical path from I to F as follows. By Lemma 1, each connected component of $G[I \oplus F]$, the subgraph of G induced by $I \oplus F$, is a tree. Since the connected components of $G[I \oplus F]$ form a partition of $I \oplus F$, we refer to the vertex sets of the connected components as components of $I \oplus F$. We process components in $I \oplus F$ in the ascending order of their smallest vertex. We first define a start vertex for each component: For a current component D, its *start vertex* $u_D \in I \oplus F$ is the vertex with the smallest depth. If there are multiple such vertices, we pick the smallest one.
We define the canonical way to convert the current component D from I to F as follows: We process D by doing depth-first search of $G[D]$ from its start vertex u_D, processing the children vertices of the current vertex in increasing order of the sizes of the subtrees associated with the children vertices. We break ties by processing smaller children first. Let u be the current vertex in the depth-first search. Then:

- If $u \in I$: If its parent has no other neighbors in the current independent set, we apply the drag transition $\leftrightarrow$ on u and its parent. Otherwise, we apply the delete transition $\downarrow$ on u.
- If $u \in F$: If u has no children, we apply the insert transition $\uparrow$ on u. Otherwise, we proceed to process the children of u.

In other words, we always remove an I-vertex before visiting its children, and we add an F-vertex (either by the insertion $\uparrow$ or by dragging $\leftrightarrow$) to the independent set after we process all its descendants. Clearly, the transitions for $u \in I$ maintain the current state as an independent set. Notice that an F-vertex is added after its I-children have been removed, and we have removed its I-parent prior to visiting this vertex; therefore, these transitions are also legal and maintain the current state as an independent set throughout the process.

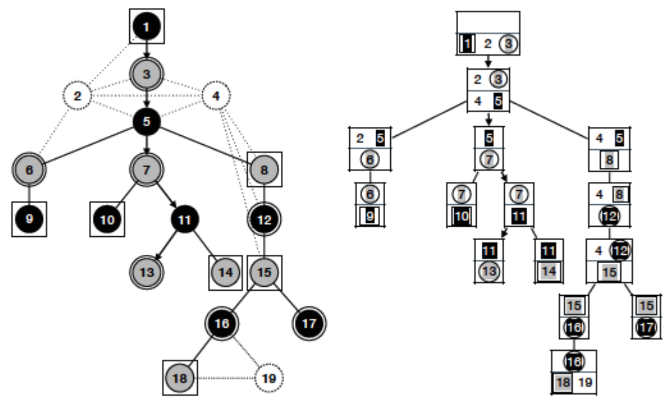

Fig. 1. On the canonical path from I to F: On the left is a chordal graph with 19 vertices. An initial and a final independent set I and F are shown in black and grey, respectively. Solid lines indicate the edges of the induced subgraph $G[I \oplus F]$, the other edges are dotted. $G[I \oplus F]$ is processed from vertex 1, the current transition t is adding $u = 13$, and the current independent set is shown in double circles. Path p_u is shown using arrows, $Q_{p_u} = \{5, 11\}$. Vertices in $\hat{\eta}_t(I, F)$ are squared. A corresponding clique tree is on the right, each clique represented by a rectangle with the separator set and the residual set at the top and bottom, respectively.

3.2 Bounding the congestion

Let $t = (S, S')$ be a transition for which we want to bound the congestion $\rho(\Gamma, t)$, see (1), the definition of which involves a sum through all canonical paths that use t. To bound this sum, one typically defines an "encoding" for each canonical path $\gamma_{I,F}$ through t. The goal for the encoding is to comprise of a state of Ω_G, and possibly some additional information chosen from a set of polynomial size.

Suppose $I, F \in \Omega_G$ are such that $\gamma_{I,F}$ uses t. Our encoding $\eta_{I,F}$ of $\gamma_{I,F}$ will consist of multiple parts. We start by defining its first part $\hat{\eta}_t(I, F)$, see Figure 1:

- Let D be the component of $I \oplus F$ on which t is applied (that is, t inserts, deletes, or drags a vertex $u \in D$).
- Let $p_u = (u_D, \ldots, u)$ be the path from the start vertex u_D to u in $G[D]$.
- Let Q_{p_u} be the set of vertices in $p_u \setminus \{u\}$ that (a) have more than one child in the tree $G[D]$ rooted at u_D, and (b) their successor vertex on p_u is not their last child.
- Then, let

$$\hat{\eta}_t(I, F) = (I \oplus F \oplus (S \cup S')) \setminus Q_{p_u}. \tag{2}$$

Denote by $\operatorname{cp}(t) := \{(I, F) \mid t \in \gamma_{I,F}\}$ the set of pairs $(I, F) \in \Omega_G^2$ whose canonical path $\gamma_{I,F}$ uses transition t. Then we have the following lemma.

Lemma 2. *For a transition t and an independent set pair (I, F) such that $(I, F) \in \operatorname{cp}(t)$, $\hat{\eta}_t(I, F)$ is an independent set.*

Proof. Let S, S', u, and D be defined as above. Let $A = I \oplus F \oplus (S \cup S')$. Then by the definition of the canonical paths, components of $I \oplus F$ prior to D have been already processed, that is, for every such component D', the current state S (and S') contains vertices in $D' \cap F$. Likewise, every component D' after D is untouched, that is S (and S') contains vertices in $D' \cap I$. Thus, A, and therefore also $\hat{\eta}_t(I,F)$, contains the I-vertices in the processed components and the F-vertices in the untouched components. These I-vertices in A (and $\hat{\eta}_t(I,F)$) form an independent set since $I \in \Omega_G$, the same is true for the F-vertices. Moreover, if D' and D'' are two different components of $I \oplus F$ (that correspond to two different connected components of $G[I \oplus F]$), there is no edge in $G[I \oplus F]$ connecting D' and D''. Thus, so far, $A \setminus D$ forms an independent set, whose vertices do not neighbor D.

It remains to analyze D itself. The path p_u splits the tree $G[D]$ into processed parts and untouched parts. Then A agrees with I on the processed parts and it agrees with F on the untouched parts. If $v, v' \in D$ are adjacent and neither is on p_u, then v and v' cannot be both in A because these two vertices are either both in the processed or both in the untouched part of the tree. Therefore, for any two adjacent vertices in D that are both in A, at least one of them is on p_u. We will prove that it is sufficient to remove Q_{p_u} from A to make it an independent set. Let v be a vertex in A and $v \in p_u \setminus Q_{p_u}$ (that is, v's last child in D is on p_u). Then, v must be in I because if v were in F, it would have been added to S (and therefore not be in A) by the drag transition $\leftrightarrow$ when processing v's last child. We will show that there is no neighbor v' of v in $A \setminus Q_{p_u}$, which will conclude our proof of $\hat{\eta}_t(I,F) = A \setminus Q_{p_u}$ being an independent set. Since $v \in I$, every neighbor of v in D is in F. We first consider v's neighbors on p_u: Let v_{parent} be the parent of v (if available) and v_{child} be the last child of v. Notice that $u \neq v$ because $u \in S \cup S'$ and, therefore, it is not in A. Thus, v_{child} exists.

We claim that a vertex v' is in $A \cap F \cap p_u$ if and only if it is in $Q_{p_u} \cap F$. This is because $v' \in A \cap F \cap p_u$ if and only if $v' \in F \cap p_u$ has not yet been added to S, which means that when we were processing the child v'' of v' on p_u, the transition over v'' was remove $\downarrow$ and thus v'' was not the last child of v', which is equivalent to $v' \in Q_{p_u} \cap F$. Therefore, for neighbor $v' \in \{v_{\text{parent}}, v_{\text{child}}\}$ of v we have $v' \in F$ and $v' \notin A \setminus Q_{p_u}$.

Finally, consider a neighbor v' of v in $A \setminus p_u$. Since v_{child} is the last child of v, we have that v' has been already processed. Since $v' \in F$, it has been already added to S. Therefore, $v' \notin A$, concluding the proof. $\square$

The following lemma bounds $|Q_{p_u}|$, the proof is omitted due to space constraints. A similar argument was made by Ge and Štefankovič [11].

Lemma 3. *Let $t = (S, S')$ be a transition, $(I,F) \in \text{cp}(t)$, and u, D, u_D, and Q_{p_u} be defined as above. Then, $|Q_{p_u}| \leq \log_2 n$.*

In our congestion bounds, we will view $I \oplus F$ through the lens of the clique tree T_G. The following observation follows directly from I and F being independent sets.

Observation. *Every clique in T_G can contain at most two vertices of $I \cup F$, and at most one vertex of I and at most one vertex of F.*

Next we relate components of $I \oplus F$ to subtrees of the rooted clique tree T_G^R. Recall that C_u refers to the clique in $V(T_G^R)$ that contains $u \in V(G)$ in its residual set, that is $u \in \mathrm{Res}(C_u)$.

Lemma 4. *Let D be a component of $I \oplus F$ and let X be the subtree of T_G spanned by clique set $\{C_u \mid u \in D\}$. Let X^R be the corresponding rooted tree of X, with edge directions consistent with T_G^R. Then the following holds:*

(i) C_{u_D} is the root of X^R.

(ii) Let $v \in D$ and let $p = (u_D = u_0, \ldots, u_\ell = v)$ be the path from u_D to v in $G[D]$. Then the directed path p_C in X^R from C_{u_D} to C_v passes through $C_{u_0}, C_{u_1}, \ldots, C_{u_\ell}$ in this order. Moreover, C_{u_i}'s are all distinct, with a possible exception of $C_{u_0} = C_{u_1}$.

Proof. We begin by proving (i). By contradiction, assume that a clique vertex $R' \neq C_{u_D}$ is the root of X^R. Then there is a directed path p' from R' to C_{u_D} in X^R. Since $u_D \in D$ is the vertex of the smallest depth (which we defined as the depth of C_{u_D} in T_G^R), none of the cliques on the path p' contain a vertex in D in their residual set. Therefore, the only reason why R' would be included in X^R is that there is another vertex $w \in D$ such that the path from C_{u_D} to C_w in X goes through R'. Let p'' be the path from R' to C_w in X^R. Notice that p'' intersects p' only at R'. If u_D were of depth 0, we would have $C_{u_D} = R = R'$. Therefore, the depth of u_D, and thus also of w, is at least 1. Since both $u_D, w \in D$, there is a path from u_D to w in $G[D]$. Then, this path needs to pass through the separator set $\mathrm{Sep}(C_{u_D})$, which means that $\mathrm{Sep}(C_{u_D})$ contains a vertex $u' \in D$. But then u' is in the parent clique of C_{u_D}, which would mean that u' has a smaller depth than u_D. This is a contradiction and, therefore, R' must be equal to C_{u_D}.

To prove (ii), we will use induction on the depth of v. For the base case, when v is of the same depth as u_D, we have two possibilities. If $v = u_D$, then $p = (u_D)$ and $p_C = (C_{u_D})$ and the statement holds. If $v \neq u_D$, since v and u_D are of the same depth and, by (i), C_{u_D} is the root of X^R, it follows that $C_v = C_{u_D}$. Then $p = (u_D, v)$ and $p_C = (C_{u_D})$ and the statement holds.

For the inductive claim, let v be of depth larger than u_D. If $\mathrm{Sep}(C_v)$ contains u_D, then $p = (u_D, v)$ and p_C starts at C_{u_D} and ends at C_v, so the statement holds. Otherwise, $\mathrm{Sep}(C_v)$ separates u_D from v. Therefore, there must be u_k, where $k \in \{1, \ldots, \ell-1\}$, such that $u_k \in \mathrm{Sep}(C_v)$. We will show that $k = \ell - 1$, that is, $u_{\ell-1} \in \mathrm{Sep}(C_v)$. By contradiction, suppose that $k < \ell - 1$. By the observation on page 9, there are at most two vertices of D in C_v. Therefore, C_v contains u_k and v, and not $v_{\ell-1}$. But since u_k and v are in the same clique C_v, there is an edge between them. Therefore, $u_k, u_{k+1}, \ldots, u_{\ell-1}, u_\ell = v$ is a cycle, contradicting Lemma 1 which states that $G[D]$ is a tree. Thus, $k = \ell - 1$.

The subtree of X of cliques containing $u_{\ell-1}$ has its root $C_{u_{\ell-1}}$ at smaller depth than C_v, since this subtree contains C_v. Therefore, the path p_C needs to pass through $C_{u_{\ell-1}}$. Since the depth of $u_{\ell-1}$ is smaller than the depth of v, we may use the inductive hypothesis for $v' := u_{\ell-1}$. We get that the directed path p'_C in X^R from C_{u_D} to $C_{v'}$ passes through

$C_{u_0}, C_{u_1}, \ldots, C_{u_{\ell-1}}$ in this order. Since p_C passes through $C_{v'}$, it is formed by extending p'_C to C_v. Therefore, p_C passes through $C_{u_0}, \ldots, C_{u_\ell}$ in this order. Moreover, $C_{u_{\ell-1}} \neq C_{u_\ell}$, finishing the proof. □

The following corollary characterizes the appearance of paths from $I \oplus F$ in T_G^R. The proof is omitted due to space constraints.

Corollary 1. *Let D be a component of $I \oplus F$, let $v \in D$, and let $p = (u_D = u_0, \ldots, u_\ell = v)$ be the path from u_D to v in $G[D]$. Then the following holds for the directed path $p_C = (C_{u_D} = C_0, \ldots, C_{\ell'} = C_v)$ in T_G^R from C_{u_D} to C_v:*

(i) For every $i \in \{0, \ldots, \ell\}$, there exist j_i, k_i, $0 \leq j_i \leq k_i \leq \ell'$ such that the cliques on the path p_C that contain u_i are exactly cliques $C_{j_i}, C_{j_i+1}, \ldots, C_{k_i}$. Moreover, $u_i \in \mathrm{Res}(C_{j_i})$.

(ii) For every $i \in \{2, \ldots, \ell\}$, $u_{i-1} \in \mathrm{Sep}(C_{j_i})$.

We are ready to define the encoding of the canonical path from I to F, passing through a transition $t = (S, S')$. Let $\mathbb{N}_b := \{1, \ldots, b\}$, where b is the separator bound of G. The encoding $\eta_t(I, F)$ consists of an independent set, a vertex, and a vector from $\mathbb{N}_b^{\lfloor \log n \rfloor}$:

$$\eta_t(I, F) := (\hat{\eta}_t(I, F), u_D, s_1, s_2, \ldots, s_{\lfloor \log n \rfloor}),$$

where the role of the vector s is to indicate the vertices of Q_{p_u} that were removed from $I \oplus F \oplus (S \cup S')$ during the construction of $\hat{\eta}_t(I, F)$, see (2). We define each s_x, $x \in \{1, \ldots, \lfloor \log n \rfloor\}$, as follows. We apply Corollary 1 to the path $p = p_u = (u_D = u_0, \ldots, u_\ell = u)$. For each vertex $u_i \in Q_{p_u}$, we have that $u_i \in \mathrm{Sep}(C_{j_{i+1}})$ (notice that $u \notin Q_{p_u}$, thus j_{i+1} is always well-defined). Suppose we ordered Q_{p_u} in increasing order of distance from u. Let x be the position of u_i in this ordering, thus s_x will encode u_i. Since $|\mathrm{Sep}(C_{j_{i+1}})| \leq b$, we can specify u_i by its position in $\mathrm{Sep}(C_{j_{i+1}})$. Thus, s_x is such that u_i is the s_x-th smallest vertex in $\mathrm{Sep}(C_{j_{i+1}})$. Notice that we will need as many s_x's as is the size of Q_{p_u}, which is bounded by $\lfloor \log_2 n \rfloor$ by Lemma 3. For $x > |Q_{p_u}|$, we let $s_x = 1$.

Lemma 5. *Let t be a transition of the Markov chain. The above-described function $\eta_t : \mathrm{cp}(t) \to \Omega_G \times V \times \mathbb{N}_b^{\lfloor \log n \rfloor}$ is injective.*

Proof (Sketch due to space constraints.). To prove the injectivity, we need to show that given a state $\hat{\eta}_t(I, F)$, a vertex u_D, a vector $(s_1, s_2, \ldots, s_{\lfloor \log n \rfloor})$, and the current transition $t = (S, S')$, we can uniquely recover the initial and final independent sets I and F.

Suppose we know $I \oplus F$. Then, due to the canonical order of processing the components, and also since we know S, we can reconstruct I and F. It remains to recover $I \oplus F$. Notice that $I \oplus F = (\hat{\eta}_t(I, F) \oplus (S \cup S')) \cup Q_{p_u}$. Therefore, the only missing part in order to determine $I \oplus F$ is Q_{p_u}. Let $B = \hat{\eta}_t(I, F) \oplus (S \cup S')$. Since u_D is given and u is known from t, we can construct the path p_C from Corollary 1 applied to $p := p_u$. Notice that we do not yet have the path p_u constructed—if we did, we would get $I \oplus F$ as $B \cup p_u$ and we would not need to reconstruct Q_{p_u}—but, despite not having p_u, we can construct p_C uniquely just from u and u_D. Our next step will be to construct p_u.

Let $p_C = (C_{u_D} = C_0, \ldots, C_{\ell'} = C_u)$. We want to reconstruct $p_u = (u_D = u_0, u_1, \ldots, u_\ell = u)$. We know all the vertices in B and we have that $u \in C_u$. We will work our way backwards, reconstructing u_i for $i = \ell - 1, \ell - 2, \ldots, 1$. Suppose we know u_{i+1} and so far $x - 1$ vertices of Q_{p_u} have been reconstructed. We consider $C_{u_{i+1}}$. By Corollary 1(ii), we have that $u_i \in \mathrm{Sep}(C_{u_{i+1}})$. By the observation on page 9 we know that u_i and u_{i+1} are the only two vertices of $I \oplus F$ in $C_{u_{i+1}}$. Therefore, we start by checking if clique $C_{u_{i+1}}$ contains a vertex from B in its separator set. If yes, it must be u_i. If not, we will use s_x to recover u_i as the s_x-th smallest vertex in $\mathrm{Sep}(C_{u_{i+1}})$. This process uniquely determines p_u, and hence also $I \oplus F$, from which we obtain I and F. □

Combining Lemmas 3 and 5 allows us to bound the congestion, which then leads to our mixing time bound. The proofs of the corresponding theorems are omitted for space reasons.

Theorem 2. *The congestion of the canonical paths defined above is bounded by $n^{(3+\log_2 b\bar{\lambda})}\bar{\lambda}$, where b is the separator bound of G and $\bar{\lambda} := \max\{1, \lambda\}$.*

Theorem 3. *Let G be a connected chordal graph with separator bound $b \in \mathbb{N}^+$, and let $\lambda \in \mathbb{R}^+$. If $\lambda < 1$, let $x = \emptyset$, otherwise, let x be a maximum independent set of G. The mixing time of the Dyer-Greenhill Markov chain from the start state x is $O(n^{(4+\log_2 b\bar{\lambda})})$.*

We remark that obtaining the start state x, a maximum independent set, is computable in polynomial time for chordal graphs. We conclude with a natural open problem, in addition to extending rapid mixing results to other graph classes: extending our results to arbitrary chordal graphs.

Acknowledgements. We thank the anonymous referees for their valuable feedback. Research supported in part by NSF grant DUE-1819546.

References

1. Anari, N., Liu, K., Gharan, S.O.: Spectral independence in high-dimensional expanders and applications to the hardcore model. In: Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC). ACM (2020)
2. Bordewich, M., Kang, R.J.: Subset Glauber dynamics on graphs, hypergraphs and matroids of bounded tree-width. Electr. J. Comb. **21**(4), P4.19 (2014)
3. Buneman, P., et al.: A characterisation of rigid circuit graphs. Discrete mathematics **9**(3), 205–212 (1974)
4. Dyer, M., Greenhill, C.: On Markov chains for independent sets. Journal of Algorithms **35**(1), 17–49 (2000)
5. Dyer, M., Greenhill, C., Müller, H.: Counting independent sets in graphs with bounded bipartite pathwidth. In: International Workshop on Graph-Theoretic Concepts in Computer Science. pp. 298–310. Springer (2019)

6. Dyer, M.E., Frieze, A.M., Jerrum, M.: On counting independent sets in sparse graphs. SIAM J. Comput. **31**(5), 1527–1541 (2002)
7. Efthymiou, C., Hayes, T.P., Štefankovič, D., Vigoda, E., Yin, Y.: Convergence of MCMC and loopy BP in the tree uniqueness region for the hard-core model. SIAM J. Comput. **48**(2), 581–643 (2019)
8. Galanis, A., Ge, Q., Štefankovič, D., Vigoda, E., Yang, L.: Improved inapproximability results for counting independent sets in the hard-core model. Random Struct. Algorithms **45**(1), 78–110 (2014)
9. Galanis, A., Štefankovič, D., Vigoda, E.: Inapproximability of the partition function for the antiferromagnetic ising and hard-core models. Comb. Probab. Comput. **25**(4), 500–559 (2016)
10. Gavril, F.: The intersection graphs of subtrees in trees are exactly the chordal graphs. Journal of Combinatorial Theory, Series B **16**(1), 47–56 (1974)
11. Ge, Q., Štefankovič, D.: A graph polynomial for independent sets of bipartite graphs. Comb. Probab. Comput. **21**(5), 695–714 (2012)
12. Jerrum, M.: Counting, Sampling and Integrating: Algorithms and Complexity. Birhäuser (2003)
13. Jerrum, M., Sinclair, A.: Approximating the permanent. SIAM J. Comput. **18**(6), 1149–1178 (1989)
14. Koller, D., Friedman, N.: Probabilistic graphical models: principles and techniques. MIT press (2009)
15. Luby, M., Vigoda, E.: Fast convergence of the Glauber dynamics for sampling independent sets. Random Structures & Algorithms **15**(3-4), 229–241 (1999)
16. Matthews, J.: Markov chains for sampling matchings. PhD Thesis, University of Edinburgh (2008)
17. Mossel, E., Weitz, D., Wormald, N.: On the hardness of sampling independent sets beyond the tree threshold. Probab. Theory Rel. Fields **143**, 401–439 (2009)
18. Okamoto, Y., Uno, T., Uehara, R.: Linear-time counting algorithms for independent sets in chordal graphs. In: International Workshop on Graph-Theoretic Concepts in Computer Science. pp. 433–444. Springer (2005)
19. Sinclair, A.: Improved bounds for mixing rates of Markov chains and multicommodity flow. Combinatorics, probability and Computing **1**(4), 351–370 (1992)
20. Sly, A.: Computational transition at the uniqueness threshold. In: 2010 IEEE 51st Annual Symposium on Foundations of Computer Science. pp. 287–296. IEEE (2010)
21. Sly, A., Sun, N.: Counting in two-spin models on d-regular graphs. Annals of Probability **42**(6), 2383–2416 (2014)
22. Vandenberghe, L., Andersen, M.S., et al.: Chordal graphs and semidefinite optimization. Foundations and Trends® in Optimization **1**(4), 241–433 (2015)
23. Vigoda, E.: A note on the Glauber dynamics for sampling independent sets. Electr. J. Comb. **8**(1) (2001)
24. Weitz, D.: Counting independent sets up to the tree threshold. In: Proceedings of the 38th annual ACM symposium on Theory of computing (STOC). pp. 140–149 (2006)