Rapid mixing of subset Glauber dynamics on graphs of bounded tree-width*

Magnus Bordewich and Ross J. Kang

Durham University, Durham, UK
m.j.r.bordewich@durham.ac.uk, ross.kang@gmail.com

Abstract. Motivated by the 'subgraphs world' view of the ferromagnetic Ising model, we develop a general approach to studying mixing times of Glauber dynamics based on subset expansion expressions for a class of graph polynomials. With a canonical paths argument, we demonstrate that the chains defined within this framework mix rapidly upon graphs of bounded tree-width. This extends known results on rapid mixing for the Tutte polynomial, the adjacency-rank (R_2-)polynomial and the interlace polynomial.

Keywords: Markov chain Monte Carlo, graph polynomials, subset expansion, tree-width, canonical paths, randomised approximation schemes.

1 Introduction

We analyse a subset-sampling Markov chain on simple graphs that is derived from certain graph functions — usually, in fact, graph polynomials. We show that this chain mixes rapidly on graphs of constant tree-width.

The graph functions $\mathcal{P}$ we consider are formulated using subset expansion. An *edge subset expansion formula for* $\mathcal{P}$ is written as follows: for any simple graph $G = (V, E)$,

$$\mathcal{P}(G) = \sum_{S \subseteq E} w((V, S)) \tag{1}$$

for some graph function w, where (V, S) denotes the graph with vertex set V and edge set S. If the function w is non-negative, that is, $w(G) \geq 0$ for all graphs G, we refer to (1) as an *edge subset weighting for* $\mathcal{P}$ and to w as its *weight function*. In fact, we shall need the weight function to be *positive* on all subgraphs — from a statistical physics viewpoint, this results in a so-called 'soft-core model'.

Before moving on, let us anchor the general formula (1) with an example that is prominent in statistical physics, theoretical computer science, and discrete

* This research was supported by the Engineering and Physical Sciences Research Council (EPSRC), grant EP/G066604/1.

probability. The *partition function of the random cluster model* can be defined for any $G = (V, E)$ and parameters q, μ as

$$Z_{RC}(G; q, \mu) := \sum_{S \subseteq E} q^{\kappa(S)} \mu^{|S|}, \tag{2}$$

where $\kappa(S)$ is the number of components in (V, S). For more on the random cluster model, see an extensive treatise by Grimmett [21]. Notice that, if $q, \mu \geq 0$, then $w((V, S)) := q^{\kappa(S)} \mu^{|S|}$ provides an edge subset weighting for $Z_{RC}(G; q, \mu)$. Under a suitable transformation, $Z_{RC}(G; q, \mu)$ is equivalent to the Tutte polynomial [43], defined for any $G = (V, E)$ and parameters x, y as

$$T(G; x, y) := \sum_{S \subseteq E} (x-1)^{r(E)-r(S)} (y-1)^{|S|-r(S)}, \tag{3}$$

where $r(S)$ is the $\mathbb{F}_2$-rank of the incidence matrix for (V, S). A wealth of combinatorial and structural information can be obtained from evaluations of this function. The Tutte polynomial specialises to several key univariate graph polynomials, including the chromatic polynomial of Birkhoff [5]. It specialises to the Jones polynomial in knot theory [28]. By its connection with the random cluster model, it also generalises the partition functions of the Ising [24] and Potts [38] models[1]. Consult the monograph of Welsh [44] for more on these crucial connections. In addition to $Z_{RC}(G; q, \mu)$ and $T(G; x, y)$, we shall highlight a few other specific polynomials from the literature, but for a broad account of the development of graph polynomials, consult the recent surveys by Makowsky [31] and Ellis-Monaghan and Merino [13, 14].

It was shown in 1990 by Jaeger, Vertigan and Welsh [25] that, in general, for fixed (rational) values of x and y, the evaluation of $T(G; x, y)$ is #P-hard, except on a few special points and curves in the (x, y)-plane. As a result, there have been substantial efforts since then to pin down the *approximation complexity* of computing $T(G; x, y)$. For large swaths of the (x, y)-plane, it is now known that the computation of $T(G; x, y)$ either does not admit a fully polynomial-time randomised approximation scheme (FPRAS) unless RP = NP, or is at least as hard as #BIS (the problem of counting independent sets in bipartite graphs) under approximation-preserving reductions, cf. Goldberg and Jerrum [19]. The sole positive approximation result applicable to general graphs is the breakthrough FPRAS by Jerrum and Sinclair [27] for the partition function of the ferromagnetic Ising model — this corresponds to computation of $T(G; x, y)$ along the portion of the parabola $(x-1)(y-1) = 2$ with $y > 1$. Various approaches have given efficient approximations in some regions of the Tutte plane for specific classes of graphs — cf. e.g. [1, 9, 29]. To obtain their seminal result, Jerrum and Sinclair used a Markov chain Monte Carlo (MCMC) method, a principal tool in the design of efficient approximation schemes for counting problems. MCMC methods are widespread in computational physics, computational biology, machine learning, and statistics. There have been steady advances in our

[1] If $x, y \geq 1$ or $q, \mu \geq 0$, then, respectively, $T(G; x, y)$ or $Z_{RC}(G; q, \mu)$ generalise the partition functions of the *ferromagnetic* Ising and Potts models.

understanding of such random processes and in showing how quickly they produce good approximations of useful probability distributions in huge, complex data sets. See the lecture notes of Jerrum [26] or a survey by Randall [39] for an overview of the application of these techniques in theoretical computer science.

We postpone the precise statement of our main result, Theorem 1, as it requires a host of definitions, but here we give a cursory description. In this paper, we are interested in the rate of convergence to stationarity of a natural Markov chain closely associated to a subset weighting of $\mathcal{P}$ (of form (1)), when some mild restriction is placed upon the weight function w. That restriction — which we have dubbed λ-*multiplicative* — is described in Subsection 2.1: for now, we remark that some important graph polynomials and partition functions from statistical physics (e.g. $Z_{RC}(G;q,\mu)$ and $T(G;x,y)$) obey it. The state space of our chain is the set of all edge subsets, upon which we have set up a MCMC method using Glauber dynamics [17]. Each possible transition in the chain is either the addition or deletion of exactly one edge to/from the subset and the transition probabilities are defined according to the weights $w((V,S))$, subject to a Metropolis-Hastings filter [22, 34][2]. Our main finding is that on graphs of bounded tree-width this Markov chain converges to the stationary distribution in time that is polynomial in the number of vertices of the graph.

Our approach is inspired in part by the 'subgraphs world' in which Jerrum and Sinclair designed their FPRAS for the partition function of the ferromagnetic Ising model. It is also motivated by recent work of Ge and Štefankovič [16], who introduced the R_2-polynomial in an attempt to devise a FPRAS for #BIS. Their *adjacency-rank polynomial* is defined for any $G=(V,E)$ and parameters q,μ as

$$R_2(G;q,\mu) := \sum_{S\subseteq E} q^{\mathrm{rk}_2(S)}\mu^{|S|}, \tag{4}$$

where $\mathrm{rk}_2(S)$ is the $\mathbb{F}_2$-rank of the *adjacency* matrix for (V,S). Using a combinatorial interpretation of rk_2 applicable only to bipartite graphs, they showed that the edge subset Glauber dynamics (using the weighting in (4)) mixes rapidly on trees. They conjectured that the chain mixes rapidly on all bipartite graphs, cf. Conjecture 1 in [16]. In addition, Ge and Štefankovič showed that the Markov chain for the (soft-core) random cluster model — i.e. weighted according to (2) — mixes rapidly upon graphs of bounded tree-width. We have extended both of these results under a unified framework. In particular, we show that the R_2-polynomial fits in our framework without recourse to the combinatorial interpretation for bipartite graphs, and hence that the Markov chain for the R_2-polynomial mixes rapidly upon all graphs of bounded tree-width. We also remark here that the conjectured rapid mixing of this chain on all bipartite graphs was disproved by Goldberg and Jerrum [18].

The polynomials and Markov chains that we capture in our framework are defined for any graph; however, we obtain rapid mixing results only on graphs of

[2] A Metropolis-Hastings filter is applied in order to ensure that the resulting process is a reversible Markov chain and thus guaranteed to converge to a unique stationary distribution with state probabilities proportional to the weight.

constant tree-width. For brevity, we will not define tree-width here, but merely say that it is an essential concept in structural graph theory and parameterised complexity — see modern surveys on the topic by Bodlaender [8] and Hliněný et al. [23]. The restriction of tree-width is commonly used in graph algorithms to reduce the complexity of a computationally difficult problem, usually by way of dynamic programming. For example, it is already known that many of the polynomials covered here can be evaluated efficiently for graphs of bounded tree-width. Independently, Andrzejak [2] and Noble [35] exhibited polynomial-time algorithms to compute the Tutte polynomial of graphs with bounded tree-width. Works of Makowsky and Mariño [32] and Noble [36] have significantly generalised this, in the former case, to a wide array of polynomials under the framework of monadic second order logic (MSOL), and, in the latter case, to the so-called U-polynomial [37], a polynomial that includes not only the Tutte polynomial but also a powerful type of knot invariant as a special case.

Even though many of the polynomials we refer to can be computed exactly in polynomial time for graphs of bounded tree-width, it remains of interest to show that the associated Glauber dynamics is rapidly mixing. One hope is that for some polynomials the chain mixes rapidly for a wider class of graphs. There have been significant and concerted endeavours by researchers spanning physics, computer science and probability to determine the mixing properties of Glauber dynamics on many related Markov chains. Spin systems have been of particular interest; indeed, the main thrust of the work of Jerrum and Sinclair was to tackle the partition function for the 'spins world' of the ferromagnetic Ising model (using a translation to the rapidly mixing 'subgraphs world'). Many recent projects on spin systems have been restricted to trees or tree-like graphs, cf. e.g. [4, 11, 12, 20, 33, 41].

Our primary focus in this paper is to establish results for polynomials defined according to *edge* subset expansion, but we can also extend our methodology to polynomials defined according to *vertex* subset expansion, which may be viewed as the 'induced subgraphs world'. To our knowledge, this form of Markov chain has not been greatly examined, but it handles one important graph polynomial that was recently introduced by Arratia, Bollobás and Sorkin [3]: the *bivariate interlace polynomial* is defined for any graph $G = (V, E)$ and parameters x, y as

$$q(G; x, y) := \sum_{S \subseteq V} (x-1)^{\mathrm{rk}_2(S)} (y-1)^{|V| - \mathrm{rk}_2(S)}, \tag{5}$$

where $\mathrm{rk}_2(S)$ is the $\mathbb{F}_2$-rank of the adjacency matrix for $G[S]$. This polynomial specialises to the independence polynomial and is intimately related to Martin polynomials. Just as for the Tutte polynomial, computation of the bivariate interlace polynomial is #P-hard in almost the entire plane [7]. The multivariate interlace polynomial, a generalisation of the interlace polynomial, can be evaluated efficiently for graphs of bounded tree-width [10], cf. [6]. Subject to a condition on the weightings, which we call *vertex λ-multiplicativity*, we establish rapid mixing for vertex subset Glauber dynamics on graphs of constant tree-width.

For all of our results, we need that the weight function is strictly positive for all (induced) subgraphs. Many of the classical enumeration polynomials such as the matching, independence, clique and chromatic polynomials are captured by the general polynomials that we mention as examples throughout this work. However, these are 'hard-core models', in which some (induced) subgraphs have a zero weighting, and hence are not included in our approach. Many of these are evaluations that fall at the boundary of the regions that we can handle. For example, the Tutte polynomial evaluated at the point $(2, 1)$ counts the number of forests of the graph. We have shown rapid mixing at all fixed points $(2, 1+\delta)$, for $\delta > 0$, with a mixing time that depends on δ. It would be interesting to consider whether the chains associated with these boundary points mix rapidly for graphs of bounded tree-width.

The structure of this paper is as follows. In the next section, we give the definitions that are necessary for a detailed description of the main theorem. We give the main theorem in Section 3 and then indicate some of its consequences. We present an outline of the proof in Section 4. In Section 5, we state how our results extend to Glauber dynamics on vertex subsets.

2 Definitions

2.1 λ-multiplicative weight functions

In this subsection, we describe the condition we require on our graph functions $\mathcal{P}$. This condition prescribes that the weight function is multiplicative with respect to the operation of disjoint graph union as well as "nearly multiplicative" with respect to the operation of composition via small vertex cuts.

We use the notation $\hat{\lambda} := \max\{\lambda, 1/\lambda\}$. For a graph $G = (V, E)$, a vertex cut K is said to separate sets V_1 and V_2 if (V_1, K, V_2) is a partition of V and there is no edge of E that is incident to both a vertex of V_1 and a vertex of V_2. A partition (E_1, E_2) of E is appropriate (for K) if E_1 has no edge adjacent to a vertex in V_2 and E_2 has no edge adjacent to a vertex in V_1.

For fixed $\lambda > 0$, we say that the weight function w is λ-*multiplicative*, if for any $G = (V, E)$, any vertex cut K that separates sets V_1 and V_2, and any appropriate partition (E_1, E_2), we have

$$\hat{\lambda}^{-|K|} \leq \frac{w((V_1 \cup K, E_1))w((V_2 \cup K, E_2))}{w(G)} \leq \hat{\lambda}^{|K|}. \tag{6}$$

As mentioned above, if w is λ-multiplicative, then w is multiplicative with respect to disjoint union (by taking $K = \emptyset$); furthermore, taking $V_2 = \emptyset$ implies that the addition or deletion of a few edges in the graph does not change w wildly.

2.2 Examples of valid polynomials

In this subsection, we emphasise specific examples with weight functions that are λ-multiplicative. Let $G = (V, E)$ be any graph, K be any vertex cut that

separates vertex subsets V_1 and V_2, and (E_1, E_2) be any appropriate partition. We define G' to be the disjoint union of graphs $(V_1 \cup K, E_1)$ and $(V_2 \cup K, E_2)$. We could imagine forming G' from G by splitting each vertex in K, taking incident edges in E_1 with one copy of the vertex and those in E_2 with the other. It is trivial to verify multiplicativity with respect to disjoint union for each of the weight functions considered below. Therefore, to establish λ-multiplicativity for these weight functions w, it will suffice to verify that $\hat{\lambda}^{-|K|} \leq w(G')/w(G) \leq \hat{\lambda}^{|K|}$.

First, we observe that the partition function of the *random cluster model* for $q, \mu > 0$ satisfies the condition. Recalling (2), the relevant weight function is $w((V, S)) := q^{\kappa(S)}\mu^{|S|}$. To handle the $\mu^{|S|}$ factor, note that the graphs G and G' have the same number of edges. For the $q^{\kappa(S)}$ factor, the number of components in G' can be at most $\kappa(G)+|K|$ since G' can be obtained by splitting $|K|$ vertices of G. Thus, w is λ-multiplicative if we take $\lambda := q$.

This can also be seen in the context of the *Tutte polynomial* when $x, y > 1$. Recalling (3), the relevant weight function is $w((V, S)) := (x-1)^{r(E)-r(S)}(y-1)^{|S|-r(S)}$. As before, it is easy to take care of the $(x-1)^{r(E)}(y-1)^{|S|}$ factor. For the remaining $((x-1)(y-1))^{-r(S)}$ factor, it is enough to observe that the incidence matrix of G may be obtained from the incidence matrix of G' as follows. The matrix for G' has two rows for each of the vertices in K, one from $(V_1 \cup K, E_1)$ and one from $(V_2 \cup K, E_2)$. If we replace one of these two rows with the sum of the two rows, we do not alter the rank; if we then delete the other of the two rows, we change the rank by at most 1. Doing this for each vertex in K, we obtain the incidence matrix for G, at a total change in the rank r of the incidence matrix of at most $|K|$. Thus, w is λ-multiplicative if we take $\lambda := (x-1)(y-1)$.

Next, we see that the *adjacency-rank polynomial* of Ge and Štefankovič satisfies the condition if $q, \mu > 0$. Recalling (4), the relevant weight function is $w((V, S)) := q^{\mathrm{rk}_2(S)}\mu^{|S|}$. As before, it is simple to handle the $\mu^{|S|}$ factor. For the $q^{\mathrm{rk}_2(S)}$ factor, we note that the adjacency matrix of G may be formed from the adjacency matrix of G' by $|K|$ row additions, followed by $|K|$ column additions and finally the deletion of $|K|$ rows and $|K|$ columns. Since we must delete both rows and columns, the rank rk_2 of the adjacency matrix may change by up to $2|K|$. Thus, in this case, w is λ-multiplicative when taking $\lambda := q^2$.

Now, consider the *multivariate Tutte polynomial* as formulated by Sokal [40], defined for any graph $G = (V, E)$ and parameters $q, \boldsymbol{v} = \{v_e\}_{e \in E}$ by

$$Z_{Tutte}(G; q, \boldsymbol{v}) := \sum_{S \subseteq E} q^{\kappa(S)} \prod_{e \in S} v_e. \tag{7}$$

Under this expansion, $w := q^{\kappa(S)} \prod_{e \in S} v_e$ is an edge subset weight function if $q > 0$ and $v_e > 0$ for any $e \in E$ are fixed. We can handle the $q^{\kappa(S)}$ factor as we did for the random cluster model partition function. For the $\prod_{e \in S} v_e$ factor, observe that G and G' have the same set of edges. Thus, w is λ-multiplicative when taking $\lambda := q$.

Last, we discuss the *U-polynomial* of Noble and Welsh [37], defined for any graph $G = (V, E)$ and parameters $y, \boldsymbol{x} = \{x_i\}_{i=1}^{|V|}$ by

$$U(G; \boldsymbol{x}, y) := \sum_{S \subseteq E} (y-1)^{|S|-r(S)} \prod_{i=1}^{|V|} {x_i}^{\kappa(i,S)}, \tag{8}$$

where $\kappa(i, S)$ denotes the number of components of order i in (V, S). If $y > 1$ and $x_i > 0$ for all i, then $w((V, S)) := (y-1)^{|S|-r(S)} \prod_{i=1}^{|V|} {x_i}^{\kappa(i,S)}$ gives an edge subset weighting. The $(y-1)^{|S|-r(S)}$ factor can be handled as above. For the $\prod_{i=1}^{|V|} {x_i}^{\kappa(i,S)}$ factor, observe that $\sum_i |\kappa(i, G) - \kappa(i, G')|$ is at most $3|K|$, since, if we obtain G' by splitting the vertices in K, each time we split a vertex we either change the size of a single component or split a single component into two smaller components. Thus, taking $x' := \max_i \max\{x_i, x_i^{-1}\}$ and $y' := \max\{y-1, (y-1)^{-1}\}$, we see that w is λ-multiplicative when taking $\lambda := y' {x'}^3$.

2.3 Glauber dynamics for edge subsets

In this subsection, we define the Markov chain associated with the edge subset expansion formula for $\mathcal{P}$. From the formulation in (1), the *single bond flip chain* $\mathcal{M}$ on a given graph $G = (V, E)$ is defined as follows. We start with an arbitrary subset $X_0 \subseteq E$ and repeatedly generate X_{t+1} from X_t by running the following experiment.

1. Pick an edge $e \in E$ uniformly at random and let $S = X_t \oplus \{e\}$.
2. Set $X_{t+1} = S$ with probability $\frac{1}{2} \min\{1, w((V, S))/w((V, X_t))\}$ and $X_{t+1} = X_t$ with the remaining probability.

By convention, we denote the state space of $\mathcal{M}$ by Ω (i.e. $\Omega = 2^E$) and its transition probability matrix by P.

The term *rapidly mixing* applies to a Markov chain that quickly converges to its stationary distribution. We make this precise here. The *total variation distance* $\|\nu - \nu'\|_{TV}$ between two probability distributions ν and ν' is defined by $\|\nu - \nu'\|_{TV} = \frac{1}{2} \sum_{H \in \Omega} |\nu(H) - \nu'(H)|$. For $\varepsilon > 0$, the *mixing time* of a Markov chain $\mathcal{M}$ (with state space Ω, transition matrix P and stationary distribution π) is defined as

$$\tau(\varepsilon) := \max_{H \in \Omega} \{\min\{t \mid \|P^t(H, \cdot) - \pi(\cdot)\|_{TV} \leq \varepsilon\}\}.$$

In this paper, we shall say that a chain $\mathcal{M}$ *mixes rapidly* if, for any fixed ε, $\tau(\varepsilon)$ is (upper) bounded by a polynomial in the number of vertices of the input graph.

3 Results

We are now prepared for a precise statement of the main theorem.

Theorem 1. *Let $G = (V, E)$ where $|V| = n$. If w is λ-multiplicative for some $\lambda > 0$, then the mixing time of $\mathcal{M}$ on G satisfies*

$$\tau(\varepsilon) = O\left(n^{4+4(\mathrm{tw}(G)+1)|\log\lambda|}\log(1/\varepsilon)\right)$$

(where $\mathrm{tw}(G)$ denotes the tree-width of G).

In Subsection 2.2, we noted some examples of polynomials with λ-multiplicative weight functions; thus, Theorem 1 implies the following.

Corollary 1. *Let $G = (V, E)$ where $|V| = n$. In the following list, we state conditions on the parameters which guarantee rapid mixing of the single bond flip chain on G associated with the stated polynomial and weighting. We also state the mixing time bound.*

1. *For fixed $q, \mu > 0$ and the weighting (2) of $Z_{RC}(G; q, \mu)$, the mixing time satisfies*

   $$\tau(\varepsilon) = O\left(n^{4+4(\mathrm{tw}(G)+1)|\log q|}\log(1/\varepsilon)\right).$$

   Equivalently, for fixed $x, y > 1$ and the weighting (3) of $T(G; x, y)$, the mixing time satisfies

   $$\tau(\varepsilon) = O\left(n^{4+4(\mathrm{tw}(G)+1)|\log((x-1)(y-1))|}\log(1/\varepsilon)\right).$$

2. *For fixed $q, \mu > 0$ and the weighting (4) of $R_2(G; q, \mu)$, the mixing time satisfies*

   $$\tau(\varepsilon) = O\left(n^{4+8(\mathrm{tw}(G)+1)|\log q|}\log(1/\varepsilon)\right).$$

3. *For fixed $q > 0$ and $v_e > 0$ for all e and the weighting (7) of $Z(G; q, \boldsymbol{v})$, the mixing time satisfies*

   $$\tau(\varepsilon) = O\left(n^{4+4(\mathrm{tw}(G)+1)|\log q|}\log(1/\varepsilon)\right).$$

4. *For fixed $y > 1$ and $x_i > 0$ for all i and the weighting (8) of $U(G; \boldsymbol{x}, \mu)$, the mixing time satisfies*

   $$\tau(\varepsilon) = O\left(n^{4+4(\mathrm{tw}(G)+1)\left|\log\left(y'x'^3\right)\right|}\log(1/\varepsilon)\right)$$

   where $x' = \max_i \max\{x_i, x_i^{-1}\}$ and $y' = \max\{y-1, (y-1)^{-1}\}$.

Here, we remark that Ge and Štefankovič obtained part 1 above and showed part 2 above in the special case of trees. Parts 2–4 directly extend these findings, and our main theorem considerably broadens the scope of mixing time bounds for subset Glauber dynamics on graphs of bounded tree-width.

4 Proof outline

Due to page restrictions, the detailed proof of Theorem 1 has been postponed to a full journal version and can be found on arXiv, but we give a brief outline.

Although our main result is stated in terms of tree-width, we do not treat tree-width directly but instead use linear-width, a more restrictive width parameter introduced by Thomas [42], which is nearly equal to path-width pw [15]. This strategy was also employed by Ge and Štefankovič in the two specific cases mentioned above. For any graph $G = (V, E)$, an ordering $(e_1, \ldots, e_m)$ of E has linear-width at most ℓ, if, for each $i \in \{1, \ldots, m\}$, there are at most ℓ vertices that are incident to both an edge in $\{e_1, \ldots, e_{i-1}\}$ and an edge in $\{e_i, \ldots, e_m\}$. The *linear-width* $\mathrm{lw}(G)$ of $G = (V, E)$ is the smallest integer ℓ such that there is an ordering of E with linear-width at most ℓ. The motive for using linear-width is that it implies an ordering of the edges which we can then use to define canonical paths between pairs of edge subsets. Then we show that λ-multiplicativity is the general condition under which we can bound the congestion of these canonical paths. The use of canonical paths is a standard technique for obtaining a bound on the mixing time of MCMC methods — see the lecture notes of Jerrum [26] for an expository account of this approach.

5 Vertex subset mixing for bounded tree-width

Until now, we had been considering edge subsets (subgraphs) and Glauber transitions which change one edge at a time. In this section, we consider vertex subsets (induced subgraphs) and transitions that involve one vertex at a time — each such transition can affect many edges, up to the maximum degree of G.

A *vertex subset expansion formula for* $\mathcal{P}$ is written as follows: for any simple graph $G = (V, E)$,

$$\mathcal{P}(G) = \sum_{S \subseteq V} w(G[S]) \tag{9}$$

for some graph function w, where $G[S]$ denotes the subgraph of G induced by S. If the function w is non-negative, we refer to (9) as an *vertex subset weighting for* $\mathcal{P}$ and to w as its *weight function*. From such a weighting, we define the *single site flip chain* $\mathcal{M}'$ on a given graph $G = (V, E)$ as follows. We start with an arbitrary subset $X_0 \subseteq V$ and repeatedly generate X_{t+1} from X_t by running the following experiment.

1. Pick a vertex $v \in V$ uniformly at random and let $S = X_t \oplus \{v\}$.
2. Set $X_{t+1} = S$ with probability $\frac{1}{2} \min\{1, w(G[S])/w(G[X_t])\}$ and $X_{t+1} = X_t$ with the remaining probability.

For fixed $\lambda > 0$, we say that the weight function w in (9) is *vertex* λ-*multiplicative*, if for any $G = (V, E)$ and K a vertex cut that separates sets

V_1 and V_2 with respect to G, we have

$$\hat{\lambda}^{-|K|} \leq \frac{w(G[V_1])w(G[V_2 \cup K])}{w(G)} \leq \hat{\lambda}^{|K|}. \tag{10}$$

The main result of this section is the following.

Theorem 2. *Let $G = (V, E)$ where $|V| = n$. If w is vertex λ-multiplicative for some $\lambda > 0$, then the mixing time of $\mathcal{M}'$ on G satisfies*

$$\tau(\varepsilon) = O\left(n^{2+4(\mathrm{tw}(G)+1)|\log\lambda|}\log(1/\varepsilon)\right).$$

Again, due to space limitations, we have omitted the proof, but note that it follows a pattern similar to what is described in Section 4, with the exception that instead of linear-width it is convenient to work with *vertex-separation* (a closely related width parameter, shown by Kinnersley [30] to be equal to path-width).

Recalling (5), for fixed $x, y > 1$, $w(G[S]) := (x-1)^{\mathrm{rk}_2(S)}(y-1)^{|V|-\mathrm{rk}_2(S)}$ gives a vertex subset weighting for $q(G; x, y)$. With arguments similar to those given in Subsection 2.2, it can be verified that this weight function is vertex λ-multiplicative. By Theorem 2, it follows that a natural Markov chain derived from the bivariate interlace polynomial — a chain that has not been studied extensively, as far as we are aware — mixes rapidly on tree-width-bounded graphs.

Corollary 2. *Let $G = (V, E)$ where $|V| = n$. If $x, y > 1$ are fixed, then for the single site flip chain on G associated with the weighting (5) of $q(G; q, \mu)$, the mixing time satisfies*

$$\tau(\varepsilon) = O\left(n^{2+8(\mathrm{tw}(G)+1)|\log((x-1)/(y-1))|}\log(1/\varepsilon)\right).$$

6 Conclusion

In this work, we have developed a new general framework of graph polynomials and Markov chains defined via subset expansion formulae for these polynomials, and demonstrated that their dynamics mix rapidly for graphs of bounded tree-width. On a graph G with n vertices, we have shown a mixing time of order $n^{O(1)}e^{O(\mathrm{pw}(G))} = n^{O(\mathrm{tw}(G))}$. Our results apply to many of the most prominent and well-known polynomials in the field. The mixing times of our processes have, respectively, exponential and super-exponential dependencies upon path-width and tree-width. We ask if this could be improved, in particular, to achieve something akin to fixed-parameter tractability in terms of tree-width.

References

1. N. Alon, A. Frieze, and D. Welsh. Polynomial time randomized approximation schemes for Tutte-Gröthendieck invariants: the dense case. *Random Structures Algorithms*, 6(4):459–478, 1995.

2. A. Andrzejak. An algorithm for the Tutte polynomials of graphs of bounded treewidth. *Discrete Math.*, 190(1-3):39–54, 1998.
3. R. Arratia, B. Bollobás, and G. B. Sorkin. A two-variable interlace polynomial. *Combinatorica*, 24(4):567–584, 2004.
4. N. Berger, C. Kenyon, E. Mossel, and Y. Peres. Glauber dynamics on trees and hyperbolic graphs. *Probab. Theory Related Fields*, 131(3):311–340, 2005.
5. G. D. Birkhoff. A determinant formula for the number of ways of coloring a map. *Ann. of Math. (2)*, 14(1-4):42–46, 1912/13.
6. M. Bläser and C. Hoffmann. Fast evaluation of interlace polynomials on graphs of bounded treewidth. To appear in *Algorithmica*. doi:10.1007/s00453-010-9439-4.
7. M. Bläser and C. Hoffmann. On the complexity of the interlace polynomial. In S. Albers and P. Weil, editors, *25th International Symposium on Theoretical Aspects of Computer Science (STACS 2008)*, volume 1 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 97–108, Dagstuhl, Germany, 2008. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
8. H. L. Bodlaender. Treewidth: characterizations, applications, and computations. In *Graph-theoretic concepts in computer science*, volume 4271 of *Lecture Notes in Comput. Sci.*, pages 1–14. Springer, Berlin, 2006.
9. M. Bordewich. Approximating the number of acyclic orientations for a class of sparse graphs. *Combin. Probab. Comput.*, 13(1):1–16, 2004.
10. B. Courcelle. A multivariate interlace polynomial and its computation for graphs of bounded clique-width. *Electron. J. Combin.*, 15(1):Research Paper 69, 36, 2008.
11. A. Dembo and A. Montanari. Ising models on locally tree-like graphs. *Ann. Appl. Probab.*, 20(2):565–592, 2010.
12. J. Ding, E. Lubetzky, and Y. Peres. Mixing time of critical Ising model on trees is polynomial in the height. *Comm. Math. Phys.*, 295(1):161–207, 2010.
13. J. A. Ellis-Monaghan and C. Merino. Graph polynomials and their applications I: The Tutte polynomial. In M. Dehmer, editor, *Structural Analysis of Complex Networks*, pages 219–255. Birkhäuser Boston, 2011.
14. J. A. Ellis-Monaghan and C. Merino. Graph polynomials and their applications II: Interrelations and interpretations. In M. Dehmer, editor, *Structural Analysis of Complex Networks*, pages 257–292. Birkhäuser Boston, 2011.
15. F. V. Fomin and D. M. Thilikos. A 3-approximation for the pathwidth of Halin graphs. *J. Discrete Algorithms*, 4(4):499–510, 2006.
16. Q. Ge and D. Štefankovič. A graph polynomial for independent sets of bipartite graphs. *CoRR*, abs/0911.4732, 2009.
17. R. J. Glauber. Time-dependent statistics of the Ising model. *J. Mathematical Phys.*, 4:294–307, 1963.
18. L. A. Goldberg and M. Jerrum. Personal communication, 2010.
19. L. A. Goldberg and M. Jerrum. Approximating the partition function of the ferromagnetic potts model. In S. Abramsky, C. Gavoille, C. Kirchner, F. M. auf der Heide, and P. G. Spirakis, editors, *ICALP (1)*, volume 6198 of *Lecture Notes in Computer Science*, pages 396–407. Springer, 2010.
20. L. A. Goldberg, M. Jerrum, and M. Karpinski. The mixing time of Glauber dynamics for coloring regular trees. *Random Structures Algorithms*, 36(4):464–476, 2010.
21. G. Grimmett. *The random-cluster model*, volume 333 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006.
22. W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

23. P. Hliněný, S.-i. Oum, D. Seese, and G. Gottlob. Width Parameters Beyond Tree-width and their Applications. *The Computer Journal*, 51(3):326–362, 2008.
24. E. Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik A Hadrons and Nuclei*, 31:253–258, 1925. 10.1007/BF02980577.
25. F. Jaeger, D. L. Vertigan, and D. J. A. Welsh. On the computational complexity of the Jones and Tutte polynomials. *Math. Proc. Cambridge Philos. Soc.*, 108(1):35–53, 1990.
26. M. Jerrum. *Counting, sampling and integrating: algorithms and complexity*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.
27. M. Jerrum and A. Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.*, 22(5):1087–1116, 1993.
28. V. F. R. Jones. A polynomial invariant for knots via von Neumann algebras. *Bull. Amer. Math. Soc. (N.S.)*, 12(1):103–111, 1985.
29. D. R. Karger. A randomized fully polynomial time approximation scheme for the all-terminal network reliability problem. *SIAM Rev.*, 43(3):499–522, 2001.
30. N. G. Kinnersley. The vertex separation number of a graph equals its path-width. *Inform. Process. Lett.*, 42(6):345–350, 1992.
31. J. A. Makowsky. From a zoo to a zoology: towards a general theory of graph polynomials. *Theory Comput. Syst.*, 43(3-4):542–562, 2008.
32. J. A. Makowsky and J. P. Mariño. Farrell polynomials on graphs of bounded tree width. *Adv. in Appl. Math.*, 30(1-2):160–176, 2003. Formal power series and algebraic combinatorics (Scottsdale, AZ, 2001).
33. F. Martinelli, A. Sinclair, and D. Weitz. Fast mixing for independent sets, colorings, and other models on trees. *Random Structures Algorithms*, 31(2):134–172, 2007.
34. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
35. S. D. Noble. Evaluating the Tutte polynomial for graphs of bounded tree-width. *Combin. Probab. Comput.*, 7(3):307–321, 1998.
36. S. D. Noble. Evaluating a weighted graph polynomial for graphs of bounded tree-width. *Electron. J. Combin.*, 16(1):Research Paper 64, 14, 2009.
37. S. D. Noble and D. J. A. Welsh. A weighted graph polynomial from chromatic invariants of knots. *Ann. Inst. Fourier (Grenoble)*, 49(3):1057–1087, 1999. Symposium à la Mémoire de François Jaeger (Grenoble, 1998).
38. R. B. Potts. Some generalized order-disorder transformations. *Proc. Cambridge Philos. Soc.*, 48:106–109, 1952.
39. D. Randall. Rapidly mixing Markov chains with applications in computer science and physics. *Computing in Science Engineering*, 8(2):30–41, March-April 2006.
40. A. D. Sokal. The multivariate Tutte polynomial (alias Potts model) for graphs and matroids. In *Surveys in combinatorics 2005*, volume 327 of *London Math. Soc. Lecture Note Ser.*, pages 173–226. Cambridge Univ. Press, Cambridge, 2005.
41. P. Tetali, J. C. Vera, E. Vigoda, and L. Yang. Phase transition for the mixing time of the Glauber dynamics for coloring regular trees. In M. Charikar, editor, *SODA*, pages 1646–1656. SIAM, 2010.
42. R. Thomas. Tree-decompositions of graphs, 1996. Lecture notes; available at `http://www.math.gatech.edu/~thomas/tree.ps`.
43. W. T. Tutte. A contribution to the theory of chromatic polynomials. *Canadian J. Math.*, 6:80–91, 1954.
44. D. J. A. Welsh. *Complexity: knots, colourings and counting*, volume 186 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1993.