A graph polynomial for independent sets of bipartite graphs*

Qi Ge and Daniel Štefankovič

Department of Computer Science
University of Rochester
Rochester, NY 14627, USA
{qge,stefanko}@cs.rochester.edu

Abstract

We introduce a new graph polynomial that encodes interesting properties of graphs, for example, the number of matchings, the number of perfect matchings, and, for bipartite graphs, the number of independent sets (#BIS).

We analyze the complexity of exact evaluation of the polynomial at rational points and show a dichotomy result—for most points exact evaluation is #P-hard (assuming the generalized Riemann hypothesis) and for the rest of the points exact evaluation is trivial.

We propose a natural Markov chain to approximately evaluate the polynomial for a range of parameters. We prove an upper bound on the mixing time of the Markov chain on trees. As a by-product we show that the "single bond flip" Markov chain for the random cluster model is rapidly mixing on constant tree-width graphs.

1998 ACM Subject Classification F.2.2, G.2.1, G.2.2, G.3

Keywords and phrases graph polynomials, #P-complete, independent sets, approximate counting problems, Markov chain Monte Carlo

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2010.240

1 Introduction

Graph polynomials are a well-developed area useful for analyzing properties of graphs (see, e. g., [7, 8] and [18]). Arguably the most intriguing graph polynomial is the Tutte polynomial [24, 25]. The partition function of the random cluster model from statistical mechanics provides a particularly simple definition: for a graph $G = (V, E)$ let

$$Z(G; q, \mu) = \sum_{S \subseteq E} q^{\kappa(S)} \mu^{|S|}, \tag{1}$$

where $\kappa(S)$ is the number of connected components of the graph (V, S). It is well-known that the Tutte polynomial is obtained from Z by a simple transformation, see, e. g., [28]. The Tutte polynomial includes many graph polynomials as special cases, e. g., the chromatic polynomial, the flow polynomial, and the Potts model (see, e. g., [28]).

Now we define our graph polynomial.

▶ **Definition 1.** The R_2-polynomial of a graph $G = (V, E)$ is

$$R_2(G; q, \mu) = \sum_{S \subseteq E} q^{\mathrm{rk}_2(S)} \mu^{|S|}, \tag{2}$$

* Research supported, in part, by NSF grant CCF-0910584.

© Qi Ge and Daniel Štefankovič;
licensed under Creative Commons License NC-ND

IARCS Int'l Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2010).
Editors: Kamal Lodaya, Meena Mahajan; pp. 240–250

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

where $\mathrm{rk}_2(S)$ is the rank of the adjacency matrix of (V,S) over $\mathbb{F}_2$ (the field with 2 elements).

The most interesting fact about the R_2 polynomial is that for bipartite graphs it encodes the number of independent sets (see Theorem 4 below). We are not aware of any other graph polynomial that encodes the number of independent sets in a non-obvious manner. (The independence polynomial of graph G is $I(G;x)=\sum_k s_k x^k$, where s_k is the number of independent sets of G of size k; here, obviously $I(G,1)$ counts the number of independent sets of G.)

To illustrate a difference between the random cluster polynomial and the R_2-polynomial we provide a few small examples. Note that P_4 and claw graph (one vertex attached to 3 other vertices) have the same random cluster polynomial whereas C_3 and claw graph have the same R_2-polynomial.

	Random cluster polynomial	R_2 polynomial
claw graph	$(\mu+q)^3q$	$(\mu^3+3\mu^2+3\mu)q^2+1$
path P_4	$(\mu+q)^3q$	$(\mu^3+\mu^2)q^4+(2\mu^2+3\mu)q^2+1$
cycle C_3	$q^3+3\mu q^2+(\mu^3+3\mu^2)q$	$(\mu^3+3\mu^2+3\mu)q^2+1$

2 Our results

Now we look at how $R_2(G;q,\mu)$ encodes some properties of graphs.

▶ **Lemma 2.** *Substituting $q=\mu^{-1/2}$ into equation (2) we define*

$$P(G;\mu):=R_2(G;\mu^{-1/2},\mu)=\sum_{S\subseteq E(G)}\mu^{|S|-\mathrm{rk}_2(S)/2}.$$

Then $P(G;0)$ is the number of matchings in G.

Proof. Note that $\mathrm{rk}_2(S)\le 2|S|$ (since adding an edge to S changes two entries in the adjacency matrix and hence can change rank by at most two), and $\mathrm{rk}_2(S)<2|S|$ if S is not a matching (since rank of the adjacency matrix of a star is $2<2|S|$, and adding further edges preserves the strict inequality). ◀

▶ **Lemma 3.** *Let*

$$P(G;t,\mu):=t^{|V|}R_2(G;1/t,\mu)\quad\text{and}\quad P_2(G;\mu):=\mu^{-|V|/2}P(G;0,\mu).$$

Then $P_2(G;0)$ is the number of perfect matchings of G.

Proof. Note that only subsets with full rank adjacency matrix contribute to $P(G;0,\mu)$, and then only the minimal cardinality subsets with full rank adjacency matrix contribute to $P_2(G;0)$ (these subsets are exactly the perfect matchings). ◀

From now on we focus solely on bipartite graphs. For a *bipartite* graph $G=(U\cup W,E)$ we let

$$R_2'(G;\lambda,\mu)=\sum_{S\subseteq E}\lambda^{\mathrm{rk}_2(S)}\mu^{|S|},\tag{3}$$

where $\mathrm{rk}_2(S)$ is the rank of the *bipartite* adjacency matrix of $(U\cup W,S)$. Note that

$$R_2(G;\lambda,\mu)=R_2'(G;\lambda^2,\mu),\tag{4}$$

since the adjacency matrix contains "two copies" of the bipartite adjacency matrix (one of them transposed). (The reason for definition (3) is that we prefer to operate with bipartite adjacency matrix for bipartite graphs.)

In Section 3 we prove that R'_2 counts the number of independent sets in bipartite graphs.

▶ **Theorem 4.** *Let $G = (U \cup W, E)$ be a bipartite graph. The number of independent sets of G is given by $2^{|U|+|W|-|E|} R'_2(G; 1/2, 1)$.*

Exact evaluation of the polynomial $R'_2(G; \lambda, \mu)$ is #P-hard at a variety of rational points (λ, μ) assuming the validity of the generalized Riemann hypothesis (GRH). The result is summarized in the following theorem.

▶ **Theorem 5.** *Exact evaluation of R'_2 at rational point (λ, μ) is*

- *polynomial-time computable when $\lambda \in \{0, 1\}$ or $\mu = 0$ or $(\lambda, \mu) = (1/2, -1)$;*
- *#P-hard when $\lambda \notin \{0, 1, 1/2\}$ and $\mu \neq 0$, assuming GRH;*
- *#P-hard when $\lambda = 1/2$ and $\mu \notin \{0, -1\}$.*

▶ Remark. For the non-bipartite case we have the following classification. Exact evaluation of R_2 at rational point (λ, μ) is polynomial-time computable when $\mu = 0$ or $\lambda \in \{-1, 0, 1\}$; the $\lambda = -1$ case follows from the fact that a skew-symmetric matrix with zero diagonal has even rank over any field (the zero diagonal condition is redundant for fields of characteristic $\neq 2$). For any other rational λ and μ we get #P-hardness of evaluating the R_2 polynomial from Theorem 5 and (4) (again assuming GRH). (Note that $(\lambda, \mu) \mapsto (\lambda^2, \mu)$ never maps to the easy case $(1/2, -1)$, since λ is rational. It would be nice to have hardness classification of evaluating R_2 and R'_2 for, say, algebraic λ and μ.)

Because of the hardness of exact evaluation of R'_2, we turn to approximate evaluation of $R'_2(G; \lambda, \mu)$.

We now define the sampling problem associated with R'_2.

Rank Weighed Subgraphs with $\lambda, \mu \geq 0$, (RWS(λ, μ))
Instance: a bipartite graph $G = (U \cup W, E)$,
Output: $S \subseteq E$ with probability of $S \propto \lambda^{\mathrm{rk}_2(S)} \mu^{|S|}$.

The "single bond flip" chain is a natural approach to sampling from RWS(λ, μ).

▶ **Definition 6.** *Single bond flip* chain is defined as follows: pick an edge $e \in E$ at random and let $S = X_t \oplus \{e\}$. Set $X_{t+1} = S$ with probability

$$(1/2) \min\{1, \lambda^{\mathrm{rk}_2(S) - \mathrm{rk}_2(X_t)} \mu^{|S| - |X_t|}\}$$

and $X_{t+1} = X_t$ with the remaining probability.

In each step of the single bond flip chain, we have to compute the rank of a matrix over $\mathbb{F}_2$ (corresponding to S) which differs from the current matrix (corresponding to X_t) in a single entry. One can use dynamic matrix rank problem algorithms to perform this computation in $O(n^{1.575})$ arithmetic operations per step [9].

Instead of flipping one edge in a step, we can have another Markov chain which flips a random subset of edges adjacent to a single vertex. It seems likely that the new chain can generate good random samples faster than the single bond flip chain—a step of the new chain can be performed in $O(n^2)$ arithmetic operations (using "rank one update" for the dynamic matrix rank problem [9]).

We consider the following question.

▶ Question 1. For which classes of bipartite graphs does the single bond flip chain mix?

In Section 5 we prove that for fixed $\lambda, \mu > 0$ the single bond flip chain mixes, in time polynomial in the number of vertices, for trees. The next theorem is motivated by Question 1, (R'_2 polynomial can be evaluated in polynomial time on trees).

▶ **Theorem 7.** *For every fixed $\lambda, \mu > 0$, the mixing time $\tau(\varepsilon)$ of the single bond flip chain for a tree on n vertices is*

$$\tau(\varepsilon) = O\left(n^{3+|\log_2 \lambda|}(|\log\lambda| + |\log\mu| + \log(1/\varepsilon))\right).$$

▶ Remark. Goldberg and Jerrum [12] recently showed that there exist bipartite graphs for which the single bond flip chain needs exponential time to mix for $\lambda = 1/2$ and $\mu = 1$ (which is the most interesting setting of λ and μ). Question 1 is still relevant—there may exist interesting classes of graphs for which the chain mixes.

As a by-product of our techniques, we show that single bond flip Markov chain for the random cluster model is rapidly mixing if $q, \mu > 0$ and G has constant tree-width (the condition $q, \mu > 0$ is equivalent to $x, y > 1$ for the Tutte polynomial $T(G; x, y)$).

Due to page limitation, we omit most of the proofs. Refer to [10] for a full version.

3 Independent sets in bipartite graphs

The problem of counting independent sets (#IS) in a graph is of interest in both computer science and statistical physics (independent sets are a special case of the so-called hard-core model, see, e. g., [1]). Exact computation of #IS is #P-complete even for 3-regular planar bipartite graphs [26, 29]. Fully polynomial randomized approximation scheme (FPRAS) is known for graphs with maximum degree $\Delta \leq 5$, [17, 6, 27]. Unless NP=RP, an FPRAS does not exist for graphs with $\Delta \geq 6$, [3, 21] .

Now we focus on the problem of counting independent sets in bipartite graphs (#BIS). While for exact counting the complexity of #BIS and #IS is the same, the situation looks very different for approximate counting, for example, no inapproximability result is known for #BIS. Dyer et al. [4] show that #BIS is complete w.r.t. approximation-preserving reductions (AP-reductions) in a sub-class of #P. Many problems were shown to be equivalent (w.r.t. AP-reductions) to #BIS, for example, #Downsets, #1p1nSat [4], computing the partition function of a ferromagnetic Ising model with local fields [11], and counting the number of satisfying assignments of a class of Boolean CSP instances [5]. A pertinent negative result for #BIS is that Glauber dynamics (or more generally, any chain whose states are independent sets and that flips at most $0.35n$ vertices in one step) cannot be used to efficiently sample random independent sets in a random 6-regular bipartite graphs on $n + n$ vertices [3].

The rest of this section is devoted to proving Theorem 4. It will be convenient to work with matrices instead of graphs. For two zero-one matrices A, B we say $B \leq A$ if B corresponds to a subgraph of A, formally

▶ **Definition 8.** Let A, B be zero-one $n_1 \times n_2$ matrices. We say $B \leq A$ if $A_{ij} = 0$ implies $B_{ij} = 0$, for all $i \in [n_1]$ and $j \in [n_2]$. Let $\mathcal{C}_A$ be the set of zero-one $n_1 \times n_2$ matrices B such that $B \leq A$.

Let $\#_1(A)$ denote the number of ones in A (that is, the number of edges in the corresponding graph). The RWS problem rephrased for matrices is:

Rank Weighed Matrices with $\lambda, \mu \geq 0$ (RWM(λ, μ))

Instance: an $n_1 \times n_2$ matrix A.
Output: $B \in \mathcal{C}_A$ with probability of $B \propto \lambda^{\mathrm{rk}_2(B)}\mu^{\#_1(B)}$.

The problem of sampling independent sets in bipartite graphs is:

Bipartite Independent Sets (BIS)
Instance: a bipartite graph $G = (U \cup W, E)$.
Output: a uniformly random independent set of G.

Before we show a connection between BIS and RWM$(1/2, 1)$ we remark that to sample bipartite independent sets it is enough to sample a subset of one side, say U, from the correct (marginal) distribution. We now describe this distribution in a setting which will be advantageous for the proof of Theorem 4.

We will represent an independent set by a pair of (indicator) vectors u, v (where $u \in \mathbb{F}_2^{n_1}$ and $v \in \mathbb{F}_2^{n_2}$).

▶ **Definition 9.** We say that two vectors $\alpha, \beta \in \mathbb{F}_2^n$ *share a one* if there exists $i \in [n]$ such that $\alpha_i = \beta_i = 1$.

We will use the following simple fact.

▶ Observation 1. Let $\alpha, \beta \in \mathbb{F}_2^n$. Let d be the number of ones in β. If α, β share a one then there are 2^{d-1} vectors $\beta' \leq \beta$ such that $\alpha^{\mathrm{T}}\beta' \equiv 0 \mod 2$. If α, β do not share a one then there are 2^d vectors $\beta' \leq \beta$ such that $\alpha^{\mathrm{T}}\beta' \equiv 0 \mod 2$.

Let $u \in \mathbb{F}_2^{n_1}$ be a vector. We would like to count the number of $v \in \mathbb{F}_2^{n_2}$ such that u, v is an independent set. Note that u, v is an independent set iff $v_j = 0$ for every $j \in [n_2]$ such that u and j-th column of A share a one. Let k be the number of columns of A that do not share a one with u. Then we have

$$u \in \mathbb{F}_2^{n_1} \text{ occurs in } 2^k \text{ independent sets.} \tag{5}$$

Thus to sample independent sets in a bipartite graph G with $n_1 \times n_2$ bipartite adjacency matrix A it is enough to sample $u \in \mathbb{F}_2^{n_1}$ with the probability of u proportional to 2^k, where k is the number of columns of A that do not share a one with u. We will call this distribution on u the marginal BIS distribution.

The following lemma shows a tight connection between BIS and RWM$(1/2, 1)$—given a sample from one distribution it is trivial to obtain a sample from the other one.

▶ **Lemma 10.** *Let G be a bipartite graph with bipartite adjacency matrix A.*

- *Let u, v be a uniformly random independent set of G. Let B be a uniformly random matrix from the following set $\{D \in \mathcal{C}_A \mid u^{\mathrm{T}}D \equiv 0 \mod 2\}$. Then B is from the* RWM$(1/2, 1)$*-distribution.*
- *Let $B \in \mathcal{C}_A$ be a random matrix from the* RWM$(1/2, 1)$*-distribution. Let $u \in \mathbb{F}_2^{n_1}$ be a uniformly random vector from the left null space of B (that is, $\{\beta \in \mathbb{F}_2^{n_1} \mid \beta^{\mathrm{T}}B \equiv 0 \mod 2\}$). Then u is from the marginal* BIS *distribution.*

Proof. Let Q be the set of u, B pairs such that $u^{\mathrm{T}}B \equiv 0 \mod 2$ and $B \leq A$. Let ψ be the uniform distribution on Q. Note that ψ marginalized over u yields the RWM$(1/2, 1)$-distribution on $B \leq A$, here we are using the fact that a d-dimensional space (in this case the left null space of B) over $\mathbb{F}_2$ has 2^d elements. Formally,

$$P(B) = \sum_{u : u^{\mathrm{T}}B \equiv 0 \mod 2} \frac{1}{|Q|} = \frac{2^{n_1 - \mathrm{rk}_2(B)}}{|Q|} = \frac{2^{-\mathrm{rk}_2(B)}}{R'_2(G; 1/2, 1)}. \tag{6}$$

Next we show that ψ marginalized over B yields the marginal BIS distribution. We compute the number of $B \leq A$ such that $u^{\mathrm{T}}B \equiv 0 \mod 2$. Let us use the same k as in (5), that is, k is the number of columns of A that do not share a one with u.

Note that the columns of B can be chosen independently and only if the column and u share a one is the number of choices (for that column) halved. Let $\#_1(A)$ be the number of ones in A. Thus

$$\text{there are } 2^{\#_1(A)-(n_2-k)} \text{ choices of } B \leq A \text{ such that } u^{\mathrm{T}}B \equiv 0 \mod 2. \tag{7}$$

Note that for fixed u the counts in (5) and (7) differ by a factor of $2^{\#_1(A)-n_2}$ (which is independent of u). Thus ψ marginalized over B yields the marginal BIS distribution on u. Formally

$$P(u) = \frac{2^{\#_1(A)-(n_2-k)}}{|Q|} = \frac{2^k}{\#\mathrm{BIS}(G)}. \tag{8}$$

Note that this proves both claims of the lemma since in both cases the u, B pair is from ψ (by first sampling from a marginal and then sampling the remaining variable) and the conclusion in both claims is a statement about marginal (of the remaining variable). ◀

Theorem 4 now follows from the proof of Lemma 10.

Proof of Theorem 4. Let Q be the set from the proof of Lemma 10. From (6) we obtain

$$|Q| = R_2'(G; 1/2, 1)2^{n_1}. \tag{9}$$

From (8) we have that the number of independent sets of G is given by

$$\#\mathrm{BIS}(G) = \frac{|Q|}{2^{\#_1(A)-n_2}}. \tag{10}$$

Combining (9) and (10) we obtain the theorem. ◀

We do not know a good combinatorial interpretation for the mod-2 rank of B for general graphs. For forests (which are, of course, always bipartite) we have the following characterization.

▶ **Lemma 11.** *Let $G = (V, E) = (U \cup W, E)$ be a forest with bipartite adjacency matrix A. Then $\mathrm{rk}_2(A)$ is the size of maximum matching in G.*

Proof. Let $a \in V$ be a leaf of G and let $e = \{a, b\} \in E$ be the edge adjacent to a. Note that b is matched in every maximum matching M (otherwise one could add e to M). Thus removing b and all adjacent edges decreases the size of maximum matching by 1.

Now we argue that removing b (and all adjacent edges) also decreases rank (over $\mathbb{F}_2$) by 1. W.l.o.g. assume that b corresponds to the first row and a corresponds to the first column. Removing b (and all adjacent edges) corresponds to removing the first row of A. Note that this decreases rank by at most 1 and it does decrease it by 1, since the only non-zero entry in the first column is in the first row. ◀

4 The linear-width of a graph

For the proof of Theorem 7 we will use the linear-width of a graph, a concept which was first defined by Thomas [23]. In this section we prove a bound on linear-width in terms of tree-width.

The *linear-width* of a graph $G = (V, E)$ is the smallest integer ℓ such that the edges of G can be arranged in a linear order $e_1, \ldots, e_m$ in such a way that, for every $i \in [m]$, there are at most ℓ vertices that have an adjacent edge in $\{e_1, \ldots, e_{i-1}\}$ and an adjacent edge in $\{e_i, \ldots, e_m\}$. It is known that computing the linear-width of a graph is NP-complete [22]. For paths and cycles the linear-width is easy to compute.

▶ **Example 12.** The linear-width of a path is 1. The linear-width of a cycle is 2.

Let $e_1, \ldots, e_m$ be a permutation of the edges of $G = (V, E)$. We say that a vertex $v \in V$ is *dangerous* w.r.t. $i \in [m]$, if there exist two edges e_j, e_k adjacent to v such that $j < i \leq k$. Let D_i be the set of vertices which are dangerous w.r.t. i. Note that the linear-width of G is the minimum value of $\max_i |D_i|$ optimized over all permutations of the edges.

Now we give an upper bound on the linear-width for trees.

▶ **Lemma 13.** *Let $T = (V, E)$ be a tree on n vertices. The linear-width of T is at most $\lfloor \log_2 n \rfloor$.*

For general graphs we will show a generalization of Lemma 13: a bound on the linear-width of G in terms of the tree-width of G. We now define tree-width (see, e. g., [16] for a nice treatment).

Given a graph $G = (V, E)$, a *tree decomposition* of G is a pair $(T, \{U_h\}_{h \in V_T})$ where $T = (V_T, E_T)$ is a tree and $U_h \subseteq V$ satisfy: (i) each edge of G is in at least one subgraph induced by U_h; and (ii) for any three vertices t_1, t_2, t_3 of T such that t_2 is in the path between t_1 and t_3 in T we have $U_{t_1} \cap U_{t_3} \subseteq U_{t_2}$. The *width* of a decomposition is $\max_{h \in V_T} |U_h| - 1$. The *tree-width* of G (denoted $\mathrm{tw}(G)$) is the minimum width optimized over all tree decompositions.

▶ **Lemma 14.** *Let $G = (V, E)$ be a graph. Then*

$$\text{linear-width}(G) \leq (\mathrm{tw}(G) + 1)(\lfloor \log_2 n \rfloor + 1).$$

5 Analysis of the single bond flip chain for trees

Given a tree $G = (V, E)$, let Ω be the set of $2^{|E|}$ subsets of E. By Lemma 11, for every $H \subseteq E$, we know that $\mathrm{rk}_2(H)$ is the size of maximum matching of the subgraph (V, H). Let $w(H)$ be the size of maximum matching in a graph (V, H). Let P be the transition matrix of the single bond flip Markov chain $\mathcal{M}$ from definition 6. It's easy to see that $\mathcal{M}$ is ergodic with unique stationary distribution π such that $\pi(H) \propto \lambda^{w(H)} \mu^{|E|}$.

The goal of this section is to prove Theorem 7.

5.1 The canonical paths

We will bound the mixing time of our chain $\mathcal{M}$ using the canonical paths method, introduced in [2, 20, 15]. Now we go over the basic definitions for Markov chains, see, e. g., [14] for a comprehensive background.

▶ **Definition 15.** The *total variation distance* of two probability distribution ν and ν' on Ω is

$$\|\nu - \nu'\|_{TV} = \frac{1}{2}\sum_{H\in\Omega} |\nu(H) - \nu'(H)| = \max_{\mathcal{S}\subseteq\Omega} |\nu(\mathcal{S}) - \nu'(\mathcal{S})|.$$

▶ **Definition 16.** The *mixing time* from initial state H, $\tau_H(\varepsilon)$, is defined as

$$\tau_H(\varepsilon) = \min\{t : \|P^t(H,\cdot) - \pi\|_{TV} \le \varepsilon\},$$

and the *mixing time* $\tau(\varepsilon)$ of the chain is defined as $\tau(\varepsilon) = \max_{H\in\Omega}\{\tau_H(\varepsilon)\}$.

Let $\sigma = e_1, \ldots, e_m$ be an ordering of the edges of $G = (V, E)$ (we will usually use the orderings supplied by Lemma 13 or Lemma 14). Given any pair $I, F \in \Omega$, let $I \oplus F$ be the symmetric difference of I and F (that is, the set of edges which are in either I or F but not in both). We define a canonical path $\gamma_{I,F}$ between I and F as follows. Let $e_{i_1}, \ldots, e_{i_k}$ be the edges from $I \oplus F$ ordered according to σ (that is, $i_1 < i_2 < \cdots < i_k$). Let

$$\gamma_{I,F} = (H_0, H_1, \ldots, H_k), \tag{11}$$

where $H_0 = I$, $H_k = F$ and $H_j = H_{j-1} \oplus \{e_{i_j}\}$.

▶ **Lemma 17.** *Let $G = (V, E)$ be a graph. Let $\sigma = e_1, \ldots, e_m$ be an ordering on E with linear-width ℓ. Let I, F be subsets of E and let H be on the canonical path* (11) *(that is, $H = H_j$ for some $j \in \{0, \ldots, k\}$). Then*

$$|w(I) + w(F) - w(H) - w(C)| \le \ell,$$

where $C = I \oplus F \oplus H$, (and $w(S)$ is the size of the maximum matching in (V, S)).

Proof. Let $Q = \{e_1, \ldots, e_{i_j}\}$. Note that $H = (F \cap Q) \cup (I \cap Q^c)$, where Q^c is the complement of Q (that is, $E \setminus Q$). Similarly, $C = (I \cap Q) \cup (F \cap Q^c)$.

Let D be the set of dangerous vertices w.r.t. e_{i_j+1}. Let M_I and M_F be maximum matchings of I and F, respectively. Let

$$M_H = (M_F \cap Q) \cup (M_I \cap Q^c) \quad \text{and} \quad M_C = (M_I \cap Q) \cup (M_F \cap Q^c).$$

Note that all vertices of M_H with degree ≥ 2 are in D (a vertex which is not D has all adjacent edges (in G) from Q or from Q^c and hence the adjacent edges (in M_H) agree with M_I or M_F). The same is true for M_C. Moreover if a vertex $v \in D$ has degree 2 in M_H then it has degree 0 in M_C. Thus by removing $\le |D|$ edges from M_H and M_C we can turn both of them into matchings. Thus

$$w(H) + w(C) \ge w(I) + w(F) - |D| \ge w(I) + w(F) - \ell. \tag{12}$$

Note that a canonical path from $I' := H$ to $F' := C$ passes through $H' := I$ (with $C' := I' \oplus F' \oplus H' = F$). Thus

$$w(I) + w(F) = w(H') + w(C') \ge w(I') + w(F') - \ell = w(H) + w(C) - \ell. \tag{13}$$

Combining (12) and (13) we get the lemma. ◀

5.2 The congestion of $\mathcal{M}$

Now we analyze the congestion of the collection $\Gamma = \{\gamma_{I,F} \,|\, I, F \in \Omega\}$ where $\gamma_{I,F}$ are canonical paths defined in (11). For each transition (H, H') such that $P(H, H') > 0$, let $cp(H, H')$ be the set of pairs (I, F) such that $(H, H') \in \gamma_{I,F}$. The congestion of Γ on (H, H') is (see, e. g., [14])

$$\varrho_{(H,H')} = \frac{1}{P(H,H')} \sum_{I,F:(H,H')\in\gamma_{I,F}} \frac{\pi(I)\pi(F)}{\pi(H)} |\gamma_{I,F}|, \tag{14}$$

where $|\gamma_{I,F}|$ is the length of $\gamma_{I,F}$. The congestion of Γ is defined as

$$\varrho := \max_{\substack{(H,H'):\\ P(H,H')>0}} \varrho_{(H,H')}.$$

We will use the following connection between the congestion and the mixing time.

▶ **Theorem 18** ([2, 20]). *$\tau_H(\varepsilon) \le \varrho(\log(1/\pi(H)) + \log(1/\varepsilon))$ for each starting state $H \in \Omega$.*

At the end of this section we prove the following bound on the congestion of Γ.

▶ **Lemma 19.** *Let $G = (V, E)$ be a graph. Let $\sigma = e_1, \ldots, e_m$ be an ordering on E with linear-width ℓ. For every (H, H') such that $P(H, H') > 0$, and for every $\lambda, \mu > 0$ we have $\varrho_{(H,H')} \le 2|E|^2 \bar{\lambda}^\ell$, where $\bar{\lambda} = \max\{\lambda, 1/\lambda\}$.*

We can now prove Theorem 7.

Proof of Theorem 7. Since $G = (V, E)$ is a tree, by Lemma 13, we have $\ell \le \lfloor \log_2 n \rfloor$, by Lemma 19, we have

$$\varrho \le 2|E|^2 \bar{\lambda}^\ell \le 2|E|^2 n^{|\log_2 \lambda|} \le 2n^{2+|\log_2 \lambda|}.$$

Theorem 7 now follows from Theorem 18. ◀

Now we bound the congestion of our canonical paths.

Proof of Lemma 19. We will bound $\varrho_{(H,H')}$ for every (H, H') such that $P(H, H') > 0$. Let $\hat{H} = H$ if $\pi(H) \le \pi(H')$ and $\hat{H} = H'$ otherwise. Note that

$$\frac{\pi(\hat{H})}{2|E|} = \pi(H)P(H, H') = \pi(H')P(H', H), \tag{15}$$

since $\mathcal{M}$ is reversible. We define a mapping $f : cp(H, H') \to \Omega$ such that $f(I, F) = I \oplus F \oplus \hat{H}$ for every pair $(I, F) \in cp(H, H')$.

First, note that f is an injection. Given $J \in \Omega$ we can determine the unique I, F such that $f(I, F) = J$, by first computing $J \oplus \hat{H}$, and the using the ordering σ on the edges of G to recover I and F.

Note that

$$|I| + |F| = |\hat{H}| + |f(I, F)|, \tag{16}$$

and

$$|w(I) + w(F) - w(\hat{H}) - w(f(I, F))| \le \ell, \tag{17}$$

where (17) follows from Lemma 17.

Let $L = \sum_J \lambda^{w(J)}\mu^{|J|}$. We have the following upper bound on $\varrho_{(H,H')}$. By (14) and (15), we have

$$\begin{aligned}
\varrho_{(H,H')} &= 2|E| \sum_{(I,F)\in cp(H,H')} \frac{\pi(I)\pi(F)}{\pi(\hat{H})}|\gamma_{I,F}| \\
&= 2|E|^2 \sum_{(I,F)\in cp(H,H')} \frac{\lambda^{w(I)+w(F)-w(\hat{H})}\mu^{|I|+|F|-|\hat{H}|}}{L} \\
&\le 2|E|^2\bar{\lambda}^{\ell} \sum_{(I,F)\in cp(H,H')} \frac{\lambda^{w(f(I,F))}\mu^{|f(I,F)|}}{L} \qquad (18)\\
&\le 2|E|^2\bar{\lambda}^{\ell}, \qquad (19)
\end{aligned}$$

where (18) follows from (16) and (17), and (19) follows from the fact that f is an injection from $cp(H,H')$ to Ω. ◀

6 Conclusions

We conclude with an observation that a generalization of RWM(λ,μ) does not have an FPRAS (unless NP=RP) and a few questions.

Let A be an $m \times n$ matrix whose entries are zeros, ones, and indeterminates, where each indeterminate occurs once. A *completion* of A is a substitution of $0, 1$ to all the indeterminates in A. We denote $\mathcal{C}_A$ to be the set of all completions of A. Let $\mathrm{rk}_2(B)$ be the rank of B over $\mathbb{F}_2$. Can we sample B from $\mathcal{C}_A$ with the probability of B proportional to $\lambda^{\mathrm{rk}_2(B)}$? Note that this problem is a generalization of the RWM$(\lambda, 1)$ problem. It turns out that finding the minimum rank completion of a matrix is NP-hard (Proposition 2.1, [19]) and hence a sampler is unlikely (unless NP=RP), since for $\lambda = 2^{-n^2}$ a random completion will be the minimum rank completion (with constant probability). The sampling problem could be easy for sufficiently large λ (the problem of finding maximum rank completion is in P, see, e. g., Section 4.1 of [13]).

▶ Question 2. What other interesting properties are encoded by the polynomial?

▶ Question 3. Can one sample maximum rank completions of a matrix?

▶ Question 4. Is the exact evaluation of the polynomial easy for bounded tree-width graphs?

References

1 Rodney J. Baxter. *Exactly solved models in statistical mechanics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1989. Reprint of the 1982 original.

2 Persi Diaconis and Daniel Stroock. Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.*, 1(1):36–61, 1991.

3 Martin Dyer, Alan Frieze, and Mark Jerrum. On counting independent sets in sparse graphs. *SIAM J. Comput.*, 31(5):1527–1541 (electronic), 2002.

4 Martin Dyer, Leslie Ann Goldberg, Catherine Greenhill, and Mark Jerrum. The relative complexity of approximate counting problems. *Algorithmica*, 38(3):471–500, 2004. Approximation algorithms.

5 Martin Dyer, Leslie Ann Goldberg, and Mark Jerrum. An approximation trichotomy for boolean #CSP. *Journal of Computer and System Sciences*, In Press, Corrected Proof, 2009.

6 Martin Dyer and Catherine Greenhill. On Markov chains for independent sets. *J. Algorithms*, 35(1):17–49, 2000.

7 Joanna Ellis-Monaghan and Criel Merino. Graph polynomials and their applications I: the Tutte polynomial. *arXiv*, 0803.3079, Jun 2008.

8 Joanna Ellis-Monaghan and Criel Merino. Graph polynomials and their applications II: interrelations and interpretations. *arXiv*, 0806.4699, Jun 2008.

9 Gudmund Skovbjerg Frandsen and Peter Frands Frandsen. Dynamic matrix rank. In *Automata, languages and programming. Part I*, volume 4051 of *Lecture Notes in Comput. Sci.*, pages 395–406. Springer, Berlin, 2006.

10 Qi Ge and Daniel Štefankovič. A graph polynomial for independent sets of bipartite graphs. *arXiv*, 0911.4732, Jan 2010.

11 Leslie Ann Goldberg and Mark Jerrum. The complexity of ferromagnetic Ising with local fields. *Combin. Probab. Comput.*, 16(1):43–61, 2007.

12 Leslie Ann Goldberg and Mark Jerrum. Personal communication, 2010.

13 Nicholas J. A. Harvey, David R. Karger, and Kazuo Murota. Deterministic network coding by matrix completion. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 489–498 (electronic), New York, 2005. ACM.

14 Mark Jerrum. *Counting, sampling and integrating: algorithms and complexity*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.

15 Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM J. Comput.*, 18(6):1149–1178, 1989.

16 Jon Kleinberg and Eva Tardos. *Algorithm Design*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2005.

17 Michael Luby and Eric Vigoda. Fast convergence of the Glauber dynamics for sampling independent sets. *Random Structures Algorithms*, 15(3-4):229–241, 1999. Statistical physics methods in discrete probability, combinatorics, and theoretical computer science (Princeton, NJ, 1997).

18 Johann A. Makowsky. From a zoo to a zoology: towards a general theory of graph polynomials. *Theory Comput. Syst.*, 43(3-4):542–562, 2008.

19 René Peeters. Orthogonal representations over finite fields and the chromatic number of graphs. *Combinatorica*, 16(3):417–431, 1996.

20 Alistair Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combin. Probab. Comput.*, 1(4):351–370, 1992.

21 Allan Sly. Computational transition at the uniqueness threshold. *arXiv*, 1005.5584, May 2010.

22 Dimitrios M. Thilikos. Algorithms and obstructions for linear-width and related search parameters. *Discrete Appl. Math.*, 105(1-3):239–271, 2000.

23 Robin Thomas. Tree-decompositions of graphs (lecture notes), 1996.

24 William T. Tutte. A ring in graph theory. *Proc. Cambridge Philos. Soc.*, 43:26–40, 1947.

25 William T. Tutte. A contribution to the theory of chromatic polynomials. *Canadian J. Math.*, 6:80–91, 1954.

26 Salil P. Vadhan. The complexity of counting in sparse, regular, and planar graphs. *SIAM J. Comput.*, 31(2):398–427 (electronic), 2001.

27 Dror Weitz. Counting independent sets up to the tree threshold. In *STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 140–149. ACM, New York, 2006.

28 Dominic J. A. Welsh. *Complexity: knots, colourings and counting*, volume 186 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1993.

29 Mingji Xia, Peng Zhang, and Wenbo Zhao. Computational complexity of counting problems on 3-regular planar graphs. *Theoret. Comput. Sci.*, 384(1):111–125, 2007.