On Sampling Colorings of Bipartite Graphs†

R. Balasubramanian and C.R. Subramanian

The Institute of Mathematical Sciences, Taramani, Chennai - 600 113, India.
Email: {balu,crs}@imsc.res.in

received Feb 12, 2004, revised Jan 19, 2005, accepted Dec 21, 2005.

We study the problem of efficiently sampling k-colorings of bipartite graphs. We show that a class of markov chains cannot be used as efficient samplers. Precisely, we show that, for any k, $6 \leq k \leq n^{1/3-\epsilon}$, $\epsilon > 0$ fixed, *almost every* bipartite graph on $n + n$ vertices is such that the mixing time of any ergodic markov chain asymptotically uniform on its k-colorings is exponential in n/k^2 (if it is allowed to only change the colors of $O(n/k)$ vertices in a single transition step). This kind of exponential time mixing is called *torpid mixing*. As a corollary, we show that there are (for every large n) bipartite graphs on $2n$ vertices with $\Delta(G) = \Omega(\ln n)$ such that for every k, $6 \leq k \leq \Delta/(6 \ln \Delta)$, each member of a large class of chains mixes torpidly.

We also show that these negative results hold true for H-colorings of bipartite graphs provided H contains a spanning complete bipartite subgraph. We also present explicit examples of colorings (k-colorings or H-colorings) which admit 1-cautious chains that are ergodic and are shown to have exponential mixing time.

While, for fixed k or fixed H, such negative results are implied by the work of [2], our results are more general in that they allow k or H to vary with n.

Keywords: Graph Colorings, Markov Chains, Analysis of Algorithms

1 Introduction

We consider the problem of efficiently sampling uniformly at random the k-colorings of a graph. This problem is related to the corresponding problem of counting the number of such structures which arise in statistical physics (see the survey of [9]). The Monte Carlo Markov chain (MCMC) approach has been very successful in obtaining efficient samplers for this problem under various assumptions about the input graph. This naturally leads to the question of how far can the MCMC approach take us. In this paper, we present some negative results on obtaining efficient samplers for colorings based on markov chains.

Several heuristics have been proposed and analyzed for efficiently sampling k-colorings [8, 13, 1, 5, 16]. The currently best known positive result is due to Vigoda [16] who showed that the chains associated with the so-called (WSK) heuristic and the Glauber dynamics both mix in polynomial time for arbitrary G provided $k \geq (11/6)\Delta(G)$ where $\Delta(G)$ denotes the maximum degree of G. See [3, 11, 6, 7] for further improvements on special classes of graphs.

†This research was partially supported by the EU ESPRIT LTR Project No. 20244 (ALCOM-IT), WP 3.3.

For arbitrary G, the focus has been on $k \geq \Delta$ because, for $k < \Delta$, it is not even guaranteed that G has a k-coloring and moreover determining if it has such a coloring is NP-complete. But, a bipartite G always has a k-coloring for $k \geq 2$ and a natural question is whether MCMC approach succeeds for $k < \Delta$. In view of the previous successes, it appears to be likely. However, no such positive result has appeared so far and only some negative results have been obtained.

For example, recently, Luczak and Vigoda [10] showed that WSK heuristic has exponential mixing time for a family of Δ-regular (constant Δ) bipartite graphs $G(k,n)$ when $3 \leq k \leq \Delta/20 \log \Delta$. Such chains with exponential mixing times are said to be *torpidly mixing*.

Similarly, the negative results of Cooper, Dyer and Frieze [2] on sampling H-colorings imply that for every *fixed* $k \geq 3$, there exist constants $\Delta = \Delta(k)$, $\delta = \delta(k) > 0$, $\tau = \tau(k) > 1$ and Δ-regular bipartite graphs G (on $2n$ vertices) such that for every δn-cautious ergodic chain $\mathcal{M}$ which is uniform on k-colorings of G, its mixing time is at least τ^n. Here, by a δn-cautious chain, we mean a chain which does not change the colors of more than δn vertices in a single transition. But these results are true only for fixed values of k and the arguments do not seem to carry over when k is allowed to grow with n.

In this paper, we further strengthen these results by establishing that, for any $s = \Omega(\ln n)$ and any k, $6 \leq k \leq s/(6 \ln s)$, for almost every bipartite graph on $2n$ vertices with $\Delta(G) \approx s$, the mixing time of every member of a large class of chains for sampling k-colorings is exponentially large in n. This result is derived as a corollary of a more general statement (given below) we prove. We use $\mathcal{C}_k(G)$ to denote the set of all (labelled) k-colorings of G.

Theorem 1.1 *Define $d = d(k) = 1/8k$ and $\gamma = \gamma(k) = 1/(385k^2)$. The following is true for every sufficiently large n and for every k, $6 \leq k \leq n^{1/3-\epsilon}$, $\epsilon > 0$ fixed but arbitrary :*

Let $G \in \mathcal{G}(n,n,p)$ with $np \geq 6k(\ln k)$. Then, with probability at least $1 - e^{-n/(k^2 \ln n)}$, G satisfies the following : For any dn-cautious ergodic markov chain $\mathcal{M}(G)$ which is asymptotically uniform on $\mathcal{C}_k(G)$, its mixing time is $\Omega(e^{\gamma n})$.

In this, by the phrase "$G \in \mathcal{G}(n,n,p)$", we mean a random bipartite graph formed between two sets A and B of equal size n. Each edge joining A and B is chosen independently with probability p. Note that np is essentially the average degree of a vertex in G.

Note that Theorem 1.1 allows k to grow with n and does not require it to be fixed. In fact, we can even suitably choose the random model so that G is a random element drawn from specific classes (see Section 5 for more details). Thus, for example, we can assume that G is a random graph of maximum degree $\approx s$ where s is any suitably large value. Similarly, we can assume that $G = K_{n,n}$. In particular, any $(n/8k)$-cautious chain (this includes the well-known Glauber dynamics) mixes torpidly on $\mathcal{C}_k(G)$ if $G = K_{n,n}$ and k lies in the stated range. (i)

The overall approach of our proof is similar to that employed by Dyer, Frieze and Jerrum [4] and is based on showing that almost surely, $\mathcal{C}_k(G)$ is such that one can obtain an exponentially small upper bound on the conductance of any dn-cautious chain. Sections 3 and 4 contain the proof of this theorem.

It is easy to check that Theorem 1.1 holds even if we redefine $\mathcal{C}_k(G)$ to the set of those k-colorings of G in which some color is not used on any vertex of some partite set (A or B). For such colorings, even such simple chains as Glauber dynamics is ergodic on any bipartite graph. See Section 6 for further discussion.

We also present some explicit examples of colorings for which the 1-cautious chain Glauber dynamics can be shown both to be ergodic and to require exponential mixing time. These explicit examples are

---

(i) Actually, for $K_{n,n}$, one can show, by slightly modifying an argument in [10], that Glauber dynamics is torpidly mixing for any $k \leq n^{1-\epsilon}$, $\epsilon > 0$ is fixed.

illustrations of the applicability of Theorem 1.1. These examples are presented in Section 7.

1.1 Generalizations to H-colorings

A careful study of the proof of Theorem 1.1 shows that the proof can be modified to prove a similar theorem for a more general scenario of sampling H-colorings of bipartite graphs. The details are given below.

Given graphs $G = (V, E)$ and $H = (U, F)$, a H-coloring of G is a mapping $c : V \rightarrow U$ such that $(c(x), c(y))$ is an edge of H whenever (x, y) is an edge of G. The standard proper k-coloring is a K_k-coloring where K_k is a simple complete graph on k vertices. We use $\mathcal{C}_H(G)$ to denote the set of all H-colorings of G. Note that if H has at least one edge, then any bipartite G admits a H-coloring and if H has no edges, then a bipartite G admits a H-coloring if and only if G also has no edges. Hence a H-coloring of a bipartite G (for arbitrary H) (or equivalently a starting point in $\mathcal{C}_H(G)$) can be found in polynomial time. Using the arguments of the proof of Theorem 1.1, we can prove the following generalization :

Theorem 1.2 *Define $d = d(k) = 1/8k$ and $\gamma = \gamma(k) = 1/(385k^2)$. The following is true for every sufficiently large n and for every k, $6 \leq k \leq n^{1/3-\epsilon}$, $\epsilon > 0$ fixed but arbitrary :*

Let $G \in \mathcal{G}(n, n, p)$ with $np \geq 6k(\ln k)$ and let H be any simple (without self-loops) graph on k vertices such that H has a $K_{\lceil k/2 \rceil, \lfloor k/2 \rfloor}$ as a subgraph. Then, with probability at least $1 - e^{-n/(k^2 \ln n)}$, G satisfies the following : For any dn-cautious ergodic markov chain $\mathcal{M}(G)$ which is asymptotically uniform on $\mathcal{C}_H(G)$, its mixing time is $\Omega(e^{\gamma n})$.

A sketch of the proof of this theorem is given in Subsection 3.1. Note that H is not required to be fixed and H and its size can vary with n. The results of [2] require H to be fixed.

2 Preliminaries

A discrete-time Markov chain M over a finite space Ω and with transition matrix P is said to be *ergodic* if there is a limiting distribution $\pi : \Omega \mapsto [0, 1]$ such that for all $x, y \in \Omega$, $Lim_{t\rightarrow\infty} P^t(x, y) = \pi(y)$. Two conditions which together are sufficient to ensure ergodicity are (1) M is *aperiodic*, i.e., $\gcd\{t : P^t(x, y) > 0\} = 1$ for all $x, y \in \Omega$, and (2) M is *irreducible*, i.e., for each $x, y \in \Omega$, there is a finite t such that $P^t(x, y) > 0$. But if M is ergodic and is asymptotically uniform over Ω, then M is also irreducible and aperiodic. The mixing time $\tau(\epsilon)$, $\epsilon > 0$, of M is the least integer $t > 0$ such that the $\mathcal{L}_1$ distance between the t-th step distribution and π is at most 2ϵ, irrespective of the starting point. For our negative results, it is enough to prove a lower bound on the value of $\tau(1/e)$.

We use the following notion introduced in [4]. For a graph G, let $\mathcal{M}(G)$ be any such chain (on $\mathcal{C}_k(G)$) with transition matrix P_G. For any $b \leq n$, we say that $\mathcal{M}(G)$ is *b-cautious* if, for any two k-colorings $\sigma_1, \sigma_2 \in \mathcal{C}_k(G)$, $\quad h(\sigma_1, \sigma_2) > b \;\Rightarrow P_G(\sigma_1, \sigma_2) = 0$. Here, $h(\sigma_1, \sigma_2)$ refers to their hamming distance (number of vertices in which σ_1 and σ_2 differ).

3 Proof of Theorem 1.1

The proof is based on establishing an upper bound on the conductance of the chains for sampling k-colorings. Its major outline is similar to that employed in [4] and is given below.

Broad outline of the proof

We prove by showing that, almost surely, G satisfies : For every dn-cautious ergodic chain $\mathcal{M}(G)$, shortly $\mathcal{M}$, with asymptotic uniform distribution π on $\Omega \doteq \mathcal{C}_k(G)$ and whose transition matrix is P :

$$\text{there exists } S \subseteq \mathcal{C}_k(G) \text{ with } 0 < |S| \leq 0.5|\Omega| \text{ and } \psi_S \leq e^{-\gamma n} \text{ where } \psi_S \doteq \frac{\sum_{i \in S, j \notin S} \pi(i)P(i,j)}{\pi(S)}.$$

The quantity ψ_S is called the *conductance* of the set S. It then follows from a well-known fact [14, 15] that $\tau(1/e) = \Omega(e^{\gamma n})$.

Before we proceed with the proof, we introduce some definitions and assumptions which will be used later. Let $V_1, V_2, \ldots, V_k$ denote the k (ordered) color classes of an arbitrary k-coloring of $G \in \mathcal{G}(n,n,p)$. For every $C \in \mathcal{C}_k(G)$, we can associate a $2k$-vector $(\overline{\alpha}, \overline{\beta})$ denoting a sequence $(\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_k)$ of nonnegative real values defined by $\alpha_i n = |V_i \cap A|$ and $\beta_i n = |V_i \cap B|$ for each $i \leq k$. Note that $\sum_i \alpha_i = \sum_i \beta_i = 1$.

Proof: Define for any k : $c_1(k) = \frac{1}{4k}$, $c_2(k) = \frac{1}{2k}$ and $c_3(k) = \frac{1}{384k^2}$. Note that $2d(k) = c_2(k) - c_1(k)$.

We introduce two specific k-tuples $\overline{\gamma}$ and $\overline{\delta}$ of reals which will be used in the definition of the set S. For even k, $\overline{\gamma}$ is defined as $\gamma_i = 2/k$ for $i \leq k/2$ and $\gamma_i = 0$ for $k/2 < i \leq k$. $\overline{\delta}$ is defined by $\delta_i = 0$ for $i \leq k/2$ and $\delta_i = 2/k$ for $k/2 < i \leq k$. Thus, for even k, the point $(\overline{\gamma}, \overline{\delta})$ looks like

$$\begin{pmatrix} \overline{\gamma} \\ \overline{\delta} \end{pmatrix} = \begin{pmatrix} \frac{2}{k}, \ldots, \frac{2}{k}, 0, \ldots, 0 \\ 0, \ldots, 0, \frac{2}{k}, \ldots, \frac{2}{k} \end{pmatrix}$$

For odd k, $\gamma_i = 2/(k+1)$ and $\delta_i = 0$ for $i \leq (k+1)/2$ and and $\gamma_i = 0$ and $\delta_i = 2/(k-1)$ for $i > (k+1)/2$.

We provide the proof for even k. For odd k, the arguments are almost identical. Define S_1 to be the set of all k-colorings with part sizes corresponding to some $(\overline{\alpha}, \overline{\beta})$ with $||(\overline{\gamma}, \overline{\delta}) - (\overline{\alpha}, \overline{\beta})||_1 \leq c_2$ where $(\overline{\gamma}, \overline{\delta})$ and c_2 are defined before. Here, $||(\overline{\gamma}, \overline{\delta}) - (\overline{\alpha}, \overline{\beta})||_1$ denotes the $L1$-norm defined as $\sum_{1 \leq i \leq k} |\gamma_i - \alpha_i| + |\delta_i - \beta_i|$. Similarly, define S_2 to be the set of all k-colorings whose corresponding vectors $(\overline{\alpha}, \overline{\beta})$ satisfy $||(\overline{\delta}, \overline{\gamma}) - (\overline{\alpha}, \overline{\beta})||_1 \leq c_2$. Both S_1 and S_2 are non-empty. Also, since $||(\overline{\gamma}, \overline{\delta}) - (\overline{\delta}, \overline{\gamma})||_1 = 2$, it follows from triangle inequality of the $\mathcal{L}_1$ metric that $S_1 \cap S_2 = \emptyset$. Define S to be S_1 if $|S_1| \leq |S_2|$ and define S to be S_2 otherwise. It follows that

Claim 3.1 $|S| \leq |\Omega|/2$.

Note that this claim holds true with probability 1, i.e., the claim is true for every possible value of G.

Without loss of generality, assume that $S = S_1$. For any chain $\mathcal{M}$ with transition matrix P, the boundary ∂S of S is defined as

$$\partial S \;=\; \{i \in S \;:\; P(i,j) > 0 \text{ for some } j \in \overline{S}\}$$

It then follows (as noticed in [10]) that

$$\psi_S \;\leq\; \left(\sum_{i \in \partial S} \pi(i) \right) / \pi(S) \;=\; \pi(\partial S)/\pi(S) \;=\; |\partial S|/|S|$$

Hence it suffices to show that, for almost every G, for every dn-cautious $\mathcal{M}$, we have $|\partial S|/|S| = O(e^{-\gamma n})$.

For an arbitrary coloring C, let $(\overline{\alpha}_C, \overline{\beta}_C)$ denote its corresponding vector. Consider any two k-colorings C, D of G. The proof of the following claim is given later.

Claim 3.2 $h(C, D) \geq 0.5n||(\overline{\alpha}_C, \overline{\beta}_C) - (\overline{\alpha}_D, \overline{\beta}_D)||_1$.

We claim that for every ergodic dn-cautious chain $\mathcal{M}(G)$,

$$\partial S \ \subseteq \ \{C \in S \ : \ c_1 \leq ||(\overline{\gamma}, \overline{\delta}) - (\overline{\alpha}_C, \overline{\beta}_C)||_1 \leq c_2\}.$$

To see this, consider any arbitrary $C \in \partial S$. Then, there exists a $D \notin S$ such that $h(C, D) \leq dn$ and hence

$$||(\overline{\alpha}_C, \overline{\beta}_C) - (\overline{\alpha}_D, \overline{\beta}_D)||_1 \quad \leq \quad 2h(C, D)/n \leq 2d = c_2 - c_1.$$

As a result, it follows that $||(\overline{\gamma}, \overline{\delta}) - (\overline{\alpha}_C, \overline{\beta}_C)||_1 \ \geq \ c_1$. For otherwise, we will have

$$\begin{aligned} ||(\overline{\gamma}, \overline{\delta}) - (\overline{\alpha}_D, \overline{\beta}_D)||_1 \quad &\leq \quad ||(\overline{\gamma}, \overline{\delta}) - (\overline{\alpha}_C, \overline{\beta}_C)||_1 + ||(\overline{\alpha}_C, \overline{\beta}_C) - (\overline{\alpha}_D, \overline{\beta}_D)||_1 \\ &< \quad c_1 + (c_2 - c_1) \ = \ c_2 \end{aligned}$$

which violates our assumption that $D \notin S$. Hence our claim is true.

Let $n_{max}(G)$ denote the number of k-colorings with part sizes corresponding to $(\overline{\gamma}, \overline{\delta})$. Let $n_{bd}(G)$ denote the number of k-colorings with part sizes corresponding to some $(\overline{\alpha}, \overline{\beta})$ with $c_1 \leq ||(\overline{\gamma}, \overline{\delta}) - (\overline{\alpha}, \overline{\beta})||_1 \leq c_2$. Thus, $|\partial S| \leq n_{bd}(G)$ and $|S| \geq n_{max}(G)$.

It is shown later (Theorem 3.1) that, with probability at least $1 - e^{-n/(k^2 \ln n)}$, G satisfies

$$\frac{|\partial S|}{|S|} \quad \leq \quad \frac{n_{bd}(G)}{n_{max}(G)} \quad < \quad e^{-\gamma n}$$

This completes the proof of Theorem 1.1. □

Proof: (of Claim 3.2) Let $C_i(A)$ denote the set of vertices in A colored (by C) with i. $C_i(B)$ is similarly defined. We can write $h(C, D) = h_A(C, D) + h_B(C, D)$ where $h_A(C, D)$ and $h_B(C, D)$ denote respectively the contributions of A and B to $h(C, D)$. We use $\oplus$ for the symmetric difference operation on sets. The claim follows from the following lower bound on $h_A(C, D)$ and a similar one for $h_B(C, D)$.

$$\begin{aligned} h_A(C, D) \quad &= \quad 0.5 \left(\sum_i \#(C_i(A) \oplus D_i(A)) \right) \quad \geq \quad 0.5 \left(\sum_i | \ \#(C_i(A)) - \#(D_i(A)) \ | \right) \\ &= \quad 0.5n \left(\sum_i | \ \alpha_i(C) - \alpha_i(D) \ | \right) \quad = \quad 0.5n \ ||\overline{\alpha}_C - \overline{\alpha}_D||_1. \end{aligned}$$

□

It only remains to prove Theorem3.1 stated formally below.

Theorem 3.1 *Let d, γ, k, G have the meaning stated in Theorem 1.1. Also, let $n_{max}G$ and $n_{bd}(G)$ have the meaning given in the proof of Theorem 1.1. Then, with probability at least $1 - e^{-n/(k^2 \ln n)}$,*

$$n_{bd}(G) < e^{-\gamma n} n_{max}(G)$$

Before we proceed with the proof of this theorem, we introduce and study some notions which are required in its derivation.

For each pair $(\overline{\alpha}, \overline{\beta})$ of two k-tuples of non-negative reals each summing to 1, let $\mathcal{C}_G(\overline{\alpha}, \overline{\beta})$ denote the set of all k-colorings of G with part sizes $|V_i \cap A| = \alpha_i n$ and $|V_i \cap B| = \beta_i n$ for each $i \leq k$. Let $N_G(\overline{\alpha}, \overline{\beta})$ denote the expected number of such k-colorings. First, we estimate the expectation $N_G(\overline{\alpha}, \overline{\beta})$ for each $(\overline{\alpha}, \overline{\beta})$ and bound these expections. We then apply Markov's inequality to get an upper bound on $n_{bd}(G)$. For a given $(\overline{\alpha}, \overline{\beta})$, note that

$$\begin{aligned} N_G(\overline{\alpha}, \overline{\beta}) &= \binom{n}{\alpha_1 n, \ \ldots, \alpha_k n} \binom{n}{\beta_1 n, \ \ldots, \beta_k n} (1-p)^{(\alpha_1\beta_1 + \ldots + \alpha_k\beta_k) n^2} \\ &= \left(\frac{n!}{(\alpha_1 n)! \ldots (\alpha_k n)!} \right) \left(\frac{n!}{(\beta_1 n)! \ldots (\beta_k n)!} \right) e^{a(\alpha_1\beta_1 + \ldots + \alpha_k\beta_k) n} \end{aligned} \tag{1}$$

where a is defined to be $a = a(n, p) = n \ln(1-p)$.[ii] Note that $a \leq 0$ and $|a| \geq np$ always. Using Stirling's approximation of factorials, it is easy to see that : For some large constant $D > 0$,

$$n^{-k} e^{\psi(\overline{\alpha}, \overline{\beta}) n} \leq \left(\frac{n!}{(\alpha_1 n)! \ldots (\alpha_k n)!} \right) \left(\frac{n!}{(\beta_1 n)! \ldots (\beta_k n)!} \right) \leq D^k e^{\psi(\overline{\alpha}, \overline{\beta}) n} \tag{2}$$

where $\psi(\overline{\alpha}, \overline{\beta})$ is defined by $\psi(\overline{\alpha}, \overline{\beta}) = \left(\sum_{i \leq k} -\alpha_i \ln \alpha_i \right) + \left(\sum_{i \leq k} -\beta_i \ln \beta_i \right)$.

Combining (1) and (2), we get

$$n^{-k} e^{\phi(\overline{\alpha}, \overline{\beta}) n} \leq N_G(\overline{\alpha}, \overline{\beta}) \leq D^k e^{\phi(\overline{\alpha}, \overline{\beta}) n} \tag{3}$$

where $\phi(\overline{\alpha}, \overline{\beta}) = \psi(\overline{\alpha}, \overline{\beta}) + a\,(\alpha_1\beta_1 + \ldots + \alpha_k\beta_k)$.[iii]

Hence if $k \leq n^{1/3}$ and if $\phi(\overline{\alpha}, \overline{\beta})$ is bounded below by a positive constant, then what essentially determines the behavior of $N_G(\overline{\alpha}, \overline{\beta})$ (for large values of n) is the value of $\phi(\overline{\alpha}, \overline{\beta})$. As we will see later, this is indeed true for each $(\overline{\alpha}, \overline{\beta})$ in our domain of interest. Analyzing ϕ, we get the following

Theorem 3.2 *Recall the assumptions mentioned and definitions of $c_1(k), c_2(k), c_3(k)$ and $(\overline{\gamma}, \overline{\delta})$ stated in the proof of Theorem 1.1. Define a to be $a = n\ln(1-p)$. It follows that $|a| \geq np \geq 6k \ln k$. Then,* for $k \geq 6$,

(i) For each $(\overline{\alpha}, \overline{\beta})$ with $c_1 \leq ||(\overline{\gamma}, \overline{\delta}) - (\overline{\alpha}, \overline{\beta})||_1 \leq c_2$, $\phi(\overline{\alpha}, \overline{\beta}) \leq \phi(\overline{\gamma}, \overline{\delta}) - c_3$.

(ii) $(\overline{\gamma}, \overline{\delta})$ is not a local maximum of ϕ. [iv]

[ii] $\ln \alpha$ denotes the logarithm of α to the natural base e.

[iii] Note that ϕ is a continuous function defined over the interior of the set ($\alpha_k = 1 - \sum_{i<k} \alpha_i$ and $\beta_k = 1 - \sum_{i<k} \beta_i$)

$$\mathcal{B} = \{(\alpha_1, \ldots, \alpha_{k-1}, \beta_1, \ldots, \beta_{k-1}) \ : \ 0 \leq \sum \alpha_i, \sum \beta_i \leq 1\}$$

and is symmetrical between the values of α_i, β_i, i.e., $\phi(\overline{\alpha}, \overline{\beta}) = \phi(\overline{\beta}, \overline{\alpha})$. ϕ can be extended to the boundary by taking the limits which are obtained by defining $0 \ln 0 = Lt_{x \to 0} x \ln x = 0$. For our purposes, each α_i (β_i) is a non-negative integral multiple of $1/n$. This is because the sizes of the color classes are integer valued.

[iv] Statement (ii) is only for the sake of information, it is not required in the analysis of mixing time.

Proof: See Section 4. $\square$

For even k, $\phi(\overline{\gamma},\overline{\delta}) = 2\ln(k/2)$ and for odd k, $\phi(\overline{\gamma},\overline{\delta}) = \ln(k+1)+\ln(k-1)-2\ln 2$. Using Theorem 3.2, we prove

Proof: (of Theorem 3.1) By definition, each k-coloring counted for $n_{max}(G)$ does not depend on any potential edge joining A and B and hence $n_{max}(G)$ is the same whatever the value of G is. Hence

$$\mathbf{Pr}(\ n_{max}(G) = E[n_{max}(G)]\ ) = 1$$

Thus, by (3), $n_{max}(G) \ \geq\ e^{\phi(\overline{\gamma},\overline{\delta})n}/n^k$. Also,

$$\begin{aligned} E[n_{bd}(G)] &= \sum_{c_1 \leq ||(\overline{\alpha},\overline{\beta})-(\overline{\gamma},\overline{\delta})||_1 \leq c_2} N_G(\overline{\alpha},\overline{\beta}) \\ &\leq \sum_{c_1 \leq ||(\overline{\alpha},\overline{\beta})-(\overline{\gamma},\overline{\delta})||_1 \leq c_2} D^k\ e^{\phi(\overline{\alpha},\overline{\beta})n} \quad \text{by} \quad (3) \\ &< (n+1)^{2k} D^k e^{-c_3 n} e^{\phi(\overline{\gamma},\overline{\delta})n} \quad \text{by Theorem 3.2} \\ &\leq (n+1)^{3k} D^k e^{-c_3 n} n_{max}(G) \\ &\leq e^{-c_3 n[1-o(1)]} n_{max}(G) \quad \text{using} \quad k = O(n^{1/3-\epsilon}). \end{aligned}$$

Hence, by Markov's inequality, with probability at least $1 - e^{-n/(k^2 \ln n)}$,

$$n_{bd}(G) \ \leq\ e^{n/(k^2(\ln n))}\ e^{-c_3 n[1-o(1)]} n_{max}(G) \ <\ e^{-\gamma n} n_{max}(G)$$

$\square$

3.1 Proof of Theorem 1.2

The proof of this generalization is essentially the same as the proof of Theorem 1.1. As before, we show that the conductance of the set S (where S is essentially the same as before) is exponentially small. We need to change the definition of $\phi(\overline{\alpha},\overline{\beta})$ as follows.

$$\phi(\overline{\alpha},\overline{\beta}) \ =\ \psi(\overline{\alpha},\overline{\beta}) + a\left(\sum_{i\in V(H)} \alpha_i\beta_i + \sum_{i\neq j,(i,j)\notin E(H)} (\alpha_i\beta_j + \alpha_j\beta_i)\right).$$

The definition of $\psi(\overline{\alpha},\overline{\beta})$ remains the same. Here, we assume that H is a labeled graph on k vertices $\{1,\ldots,k\}$. Because of the assumption that H contains a $K_{k/2,k/2}$ subgraph, $n_{max}(G)$ remains the same. Also, the expectation of $n_{bd}(G)$ decreases because of the extra additive term which takes negative values.

4 Analysis of ϕ at $(\overline{\gamma},\overline{\delta})$

Proof of Theorem 3.2i Recall that $|a| \geq 6k\ln k$. We prove the statement for even k. For odd k, the arguments are similar. Define $c_1(k) = \frac{1}{4k}$, $c_2(k) = \frac{1}{2k}$ and $c_3(k) = \frac{1}{384k^2}$. The point $(\overline{\gamma},\overline{\delta})$ is $(2/k,\ldots,0,\ldots,0,\ldots,2/k,\ldots)$. Consider any point $(\overline{\alpha},\overline{\beta})$ with $c_1 \leq ||(\overline{\gamma},\overline{\delta}) - (\overline{\alpha},\overline{\beta})||_1 \leq c_2$.

For $i \leq k/2$, write $\alpha_i = 2/k + \kappa_i$ and $\beta_i = \tau_i$. For $i > k/2$, write $\alpha_i = \tau_i$ and $\beta_i = 2/k + \kappa_i$. Each $\tau_i \geq 0$. Since $2 = \sum_i \alpha_i + \sum_i \beta_i = 2 + \sum_i \kappa_i + \sum_i \tau_i$, we have $\sum_i \kappa_i \leq 0$. We have,

$$\phi(\overline{\alpha}, \overline{\beta}) \quad = \quad T + a\left(\sum_i (2/k + \kappa_i)\tau_i\right)$$

where

$$\begin{aligned} T &= -\sum_i (2/k+\kappa_i)\ln(2/k+\kappa_i) - \left(\sum_i \tau_i \ln \tau_i\right) \\ &= -\left(\sum_i (2/k+\kappa_i)\ln(2/k)\right) - \left(\sum_i (2/k+\kappa_i)(k\kappa_i/2 - (k\kappa_i/2)^2/2 + \ldots)\right) - \left(\sum_i \tau_i \ln \tau_i\right) \\ &= 2\ln(k/2) - \left(\sum_i \kappa_i(1+\ln(2/k))\right) - \left(\sum_i k\kappa_i^2/4 - k^2\kappa_i^3/24 + \ldots\right) - \left(\sum_i \tau_i \ln \tau_i\right). \end{aligned}$$

Hence

$$\begin{aligned} \phi(\overline{\gamma}, \overline{\delta}) - \phi(\overline{\alpha}, \overline{\beta}) &= (1+\ln(2/k))\left(\sum_i \kappa_i\right) + \left(\sum_i \tau_i \left(\ln \tau_i - a(2/k+\kappa_i)\right)\right) \\ &\quad + \left(\sum_i k\kappa_i^2/4 - k^2\kappa_i^3/24 + \ldots\right). \end{aligned}$$

Let T_1, T_2, T_3 denote (respectively in that order) the three terms in the above expression. By our notation, $\sum_i \kappa_i \leq 0$. Also, since $k \geq 6 \geq 2e$, $1 + \ln(2/k) \leq 0$. Hence T_1 is non-negative.

$$\begin{aligned} T_3 &\geq \left(\sum_i k\kappa_i^2/4\right) - \left(\sum_i k^2|\kappa_i|^3/24(1-k|\kappa_i|)\right) \\ &\geq \left(\sum_i k\kappa_i^2/4\right) - \left(\sum_i k|\kappa_i|^2/24\right) \\ &= (5k/24)\left(\sum_i \kappa_i^2\right) \quad \geq \quad (5/24)\left(\sum_i |\kappa_i|\right)^2 \\ &\geq (5/96)||(\kappa,\tau)||_1^2 \quad \geq \quad 5/1536k^2. \end{aligned}$$

In the second term, by our choice of $(\overline{\alpha}, \overline{\beta})$, $|\kappa_i| \leq 1/2k$. Hence $\sum_i \tau_i(-a)(2/k + \kappa_i) \geq 3|a|/2k\,(\sum_i \tau_i)$. Divide the sum $\sum_i \tau_i \ln \tau_i$ into two sums S_1 (summing over $\tau_i \leq 1/k^9$) and S_2 (summing over the remaining). $S_1 \geq -k(9\ln k)/k^9 = -(9\ln k)/k^8$. Also, $S_2 \geq -(9\ln k)(\sum_i \tau_i)$. Hence

$$T_2 \quad \geq \quad \left(\frac{3|a|}{2k} - 9\ln k\right)\left(\sum_i \tau_i\right) - \frac{9\ln k}{k^8}.$$

Combining all these,

$$\begin{aligned}\phi(\overline{\gamma},\overline{\delta})-\phi(\overline{\alpha},\overline{\beta}) &\geq \frac{5}{1536k^2}-\frac{9\ln k}{k^8}+\left(\frac{3|a|}{2k}-9\ln k\right)\left(\sum_i \tau_i\right) \qquad (4)\\ &\geq \frac{1}{384k^2} = c_3(k) \text{ if we choose } |a|\geq 6k\ln k.\end{aligned}$$

Thus, for $k \geq 6$, this proves Theorem 3.2i].

Proof of Theorem 3.2 ii We only consider the case of even k. To prove (ii) Consider the point $(\overline{\gamma^\epsilon},\overline{\delta^\epsilon})$ defined for every small $\epsilon > 0$. It is defined as follows. For even k, $\gamma_i^\epsilon = \delta_{k/2+i}^\epsilon = 2/k - \epsilon/k$ and $\delta_i^\epsilon = \gamma_{k/2+i}^\epsilon = \epsilon/k$ for $1 \leq i \leq k/2$.

$$\begin{aligned}\phi(\overline{\gamma^\epsilon},\overline{\delta^\epsilon}) &= -k\left(\frac{2}{k}-\frac{\epsilon}{k}\right)\ln\left(\frac{2}{k}-\frac{\epsilon}{k}\right)-k\frac{\epsilon}{k}\ln(\epsilon/k)+ak\left(\frac{2}{k}-\frac{\epsilon}{k}\right)\frac{\epsilon}{k}\\ &= \phi(\overline{\gamma},\overline{\delta})+\epsilon\left(\ln\frac{2}{k}-\ln\frac{\epsilon}{k}+1\right)-\frac{\epsilon^2}{4}+\frac{a\epsilon}{k}(2-\epsilon)+O(\epsilon^3)\\ &> \phi(\overline{\gamma},\overline{\delta}) \quad \text{for all sufficiently small } \epsilon > 0.\end{aligned}$$

Hence $(\overline{\gamma},\overline{\delta})$ is not a local maximum of ϕ.

5 Consequences

The following corollary shows that Theorem 1.1 holds even if our random G is conditioned to satisfy some property which holds with "not a too small" probability. We also illustrate the immediate consequences of this corollary with some specific examples. Throughout, k is a value such that $6 \leq k \leq n^{1/3-\epsilon}$. Also, d,γ have the same meaning given in Theorem 1.1.

Definition : Let $\mathcal{P}$ denote an arbitrary property of bipartite graphs which is satisfied with positive probability by a random $G \in \mathcal{G}(n,n,p)$. Given $p = p(n)$, we define $s_p(\mathcal{P},n)$ as the probability that G satisfies $\mathcal{P}$. The random model $\mathcal{G}(n,n,p) \mid \mathcal{P}$ is defined to the model obtained by conditioning $\mathcal{G}(n,n,p)$ to the event "$G \in \mathcal{G}(n,n,p)$ satisfies $\mathcal{P}$".

Corollary 5.1 *Let $\mathcal{P}$ be any property such that there is a $p = p(n)$ with $np \geq 6k(\ln k)$ and $e^{-n/(k^2\ln n)} = o(s_p(\mathcal{P},n))$. Then, almost surely,*[(v)] *$G_0 \in \mathcal{G}(n,n,p) \mid \mathcal{P}$ satisfies the claim of Theorem 1.1.*

Corollary 5.2 *For any k, $6 \leq k = O(n^{1/3-\epsilon})$, for any dn-cautious ergodic markov chain $\mathcal{M}(K_{n,n})$ which is asymptotically uniform on $\mathcal{C}_k(K_{n,n})$, its mixing time is $\Omega(e^{\gamma(k)n})$.*

Proof: Let $\mathcal{P}$ denote the property that $G = K_{n,n}$ and choose $p = 1$ leading to $np = n$. Also, $s_p(\mathcal{P},n) = 1$ satisfying the requirements of Corollary 5.1. □

Corollary 5.3 *Let G' denote any one of the random bipartite graphs G_1 and G_2 on $n+n$ vertices defined below. Then, almost surely, G' satisfies : For any dn-cautious ergodic markov chain $\mathcal{M}(G')$ which is asymptotically uniform on $\mathcal{C}_k(G')$, its mixing time is $\Omega(e^{\gamma(k)n})$.*

(v) In other words, most of bipartite graphs satisfying $\mathcal{P}$ also satisfy the torpid mixing property.

(i) *G_1 is random with edge-density $dens(G_1) = |E(G_1)|/|V(G_1)| \leq s$, for any $s \geq 4k(\ln k)$.*

(ii) *G_2 is random with $0.9s \leq \Delta(G_2) \leq 1.1s$, for any $s \geq 850(\ln n)$ and any k where $6 \leq k \leq \min\{s/(6\ln s), n^{1/3-\epsilon}\}$.*

Proof: Let $G \in \mathcal{G}(n,n,p)$ be a random bipartite graph on $n+n$ vertices.

(i) Choose $p = 3s/2n$. Then, $np \geq 6k(\ln k)$. Let $\mathcal{P}$ denote the property that $dens(G) \leq s$ and let $G_1 \in \mathcal{G}(n,n,p) \mid \mathcal{P}$. Then $s_p(\mathcal{P}, n) = 1 - o(1)$ satisfying the requirements of Corollary 5.1.

(ii) Choose $p = s/n$. Then, $np \geq 6k(\ln k)$. Let $\mathcal{P}$ denote the property that $0.9s \leq \Delta(G) \leq 1.1s$ and let $G_2 \in \mathcal{G}(n,n,p) \mid \mathcal{P}$. For an arbitrary vertex u, by Chernoff-Hoeffding bounds (see [12]),

$$\mathbf{Pr}(\ \Delta(G) > 1.1s\ ) \ \leq\ 2n\,\mathbf{Pr}(\ deg_u(G) > 1.1(np)\ )\ \leq\ 2n\,e^{-np/400}\ =\ o(n^{-1}).$$

Also, $\mathbf{Pr}(\ \Delta(G) < 0.9s\ ) \ \leq\ \mathbf{Pr}(\ \forall u \in A \cup B,\ deg_u(G) < 0.9(np)\ )\ \leq\ e^{-np/200}\ =\ o(n^{-1})$. Hence $s_p(\mathcal{P}, n) = 1 - o(1)$ satisfying the requirements of Corollary 5.1. (vi)

□

6 Ergodicity and cautiousness

Theorem 1.1 assumes that there are some dn-cautious ergodic chains on $\mathcal{C}_k(G)$, otherwise there is no need to worry about mixing time. So it becomes important to study the conditions under which such ergodic chains exist. We answer this question partially by providing some positive answers and also some negative answers.

There are $O(1)$-cautious ergodic chains which are uniform on $\mathcal{C}_k(K_{n,n})$ for $k \geq 3$. Glauber dynamics (GD) which tries to recolor a uniformly chosen vertex with a uniformly chosen new color, is one such example. It is easy to verify that this chain is ergodic on the s-colorings of any complete l-partite graph G_l, for any $s > l$. In particular, it is ergodic on k-colorings (for $k \geq 3$) of any complete bipartite graph. The same is true of some variants of Glauber dynamics also (see [5]).

But, there exist some bipartite graphs for which GD is not ergodic (see the Fact 6.1 for a more general statement). As noted in [10], the chain associated with WSK algorithm is ergodic for any bipartite graph even though it is not (n/k)-cautious. However, if we restrict our sample space to those k-colorings of G in which some color is not used on any vertex of some partite set (A or B), then Glauber dynamics is ergodic for each $k \geq 3$ and for every bipartite graph. Also, Theorem 1.1 holds true for such colorings.

Also, a large value of b alone cannot guarantee that some ergodic b-cautious chain can be designed for sampling uniformly from $\mathcal{C}_k(G)$ for each bipartite G and k. This is shown in the following proposition.

Fact 6.1 *For every two positive integers $k, b \geq 2$, there exists a bipartite graph H on $n = kb$ vertices such that no b'-cautious ($b' < b$) chain is irreducible on $\mathcal{C}_k(H)$. Hence no such chain is asymptotically uniform on $\mathcal{C}_k(H)$.*

(vi) The lower bound on s in Corollary 5.3 is close to optimal within a multiplicative factor of $O(\ln k)$. This is because if $s \leq (6/11)k$ (or equivalently, $k \geq (11/6)s$), then for every G with $\Delta(G) \leq s$, the 1-cautious chain Glauber Dynamics is ergodic and mixes rapidly as shown by [16].

Proof: Let $b_1 = \lceil b/2 \rceil$ and $b_2 = \lfloor b/2 \rfloor$. H has two partite sets L and R defined by

$$L = \{u_{i,j} \ : \ 1 \leq i \leq k, \ 1 \leq j \leq b_1\}, \quad R = \{v_{i,j} \ : \ 1 \leq i \leq k, \ 1 \leq j \leq b_2\}$$

There is an edge between $u_{i,j}$ and $v_{i',j'}$ if and only if $i \neq i'$. H is bipartite and has n vertices. Consider a k-coloring C of H defined as $C(u_{i,j}) = C(v_{i,j'}) = i$ for any i, j, j'. For any $u_{i,j}$, if we try to change its color to some $i' \neq i$, then it forces each $v_{i',j'}$ to take a color $i'' \neq i'$. This further changes the colors of at least $\lceil b/2 \rceil$ vertices in L. Hence, at least b vertices are forced to change their colors. Hence $h(C, C') \geq b$ for every $C' \neq C$. Thus, there is no b'-cautious ($b' < b$) chain irreducible on $\mathcal{C}_k(H)$. □

Fact 6.1 does not weaken the statement of Theorem 1.1. It only shows that there are bipartite graphs for which less powerful chains cannot be ergodic. Theorem 3.1, on the other hand, looks at some graphs for which such less powerful but ergodic chains exist and show that they do not mix rapidly. However, it is very likely that there are $O(n/k)$-cautious chains (on k-colorings) which are ergodic on all or almost all bipartite graphs.

7 Examples of colorings admitting cautious ergodic chains

In this section, we provide some explicit examples of colorings which admit ergodic chains that are 1-cautious and which can be shown (as in Theorem 1.1) to have exponential mixing time.

7.1 Example

Let $\mathcal{C}_{1,k+1}(G)$ be the set of all $(k+1)$-colorings of $G = (A \cup B, E)$ such that

(i) color 1 is used only for vertices of A,

(ii) color $k+1$ is used only for vertices of B and

(iii) color 2 through k are used for vertices of both A and B.

Let H be the complete bipartite graph on $X = \{u_1, \ldots, u_k\} \cup Y = \{v_1, \ldots, v_k\}$ *minus* the near perfect matching $\{(u_2, v_2), \ldots, (u_k, v_k)\}$. It can be seen that sampling uniformly from $\mathcal{C}_{1,k+1}(G)$ is equivalent to sampling uniformly from $\mathcal{C}_H(G, X, Y)$ where

$$\mathcal{C}_H(G, X, Y) = \{c \in \mathcal{C}_H(G) \ \mid \ c(A) \subseteq X, \ c(B) \subseteq Y\}.$$

This is seen from the following bijection which maps a $c \in \mathcal{C}_H(G, X, Y)$ to the coloring

$$\langle c^{-1}(u_1), \ c^{-1}(\{u_2, v_2\}), \ c^{-1}(\{u_3, v_3\}), \ \ldots \ , c^{-1}(\{u_k, v_k\}), \ c^{-1}(v_1)) \rangle.$$

One can derive negative results (similar to Theorem 1.1) on sampling from $\mathcal{C}_H(G, X, Y)$. The proof is essentially the same as the proof of Theorem 1.1. For $i = 2, \ldots, k$, define α_i and β_i by $|A \cap c^{-1}(u_i)| = \alpha_i n$ and $|B \cap c^{-1}(v_i)| = \beta_i n$. Also, let $\alpha_1 = |A \cap c^{-1}(u_1)|/n$ and $\beta_1 = 0$ and let $\beta_{k+1} = |B \cap c^{-1}(v_1)|/n$ and $\alpha_{k+1} = 0$. Note that β_1 and α_{k+1} are forced to be zero always. Now, as in Theorem 1.1, we can define $\overline{\gamma}$ and $\overline{\delta}$ suitably and show that the conductance of the set S is exponentially small. We leave the details to the reader. This shows that Theorem 1.1 applies to the ergodic cautious chains for sampling from $\mathcal{C}_{1,k+1}(G)$.

The result of applying the procedure CLOSURE(H) (described in Section 3 of [2]) is a complete unlooped bipartite graph. Hence it follows from [2] (Theorem 3) that the 1-cautious chain Glauber dynamics is ergodic for all bipartite G and is uniform on the space $\mathcal{C}_{1,k+1}(G)$. The Glauber dynamics for a H-coloring is a 1-cautious chain which tries at each step to re-color at random a single randomly chosen vertex. By the afore-said arguments, GD has exponential mixing time for $G \in \mathcal{G}(n, n, p)$.

7.2 Example

In general, for $0 \leq a \leq k$, let $\mathcal{C}_{a,k}(G)$ denote the set of all k-colorings of $G = (A \cup B, E)$ such that

(i) colors 1 through a are used only for vertices of A,

(ii) colors $k - a + 1$ through k are used only for vertices of B and

(iii) colors $a + 1$ through $k - a$ are used for vertices of both A and B.

Let H be the complete bipartite graph on $X = \{u_1, \ldots, u_{k-a}\} \cup Y = \{v_{a+1}, \ldots, v_k\}$ *minus* the matching $\{(u_{a+1}, v_{a+1}), \ldots, (u_{k-a}, v_{k-a})\}$. It can be seen that sampling uniformly from $\mathcal{C}_{a,k}(G)$ is equivalent to sampling uniformly from $\mathcal{C}_H(G, X, Y)$ where

$$\mathcal{C}_H(G, X, Y) = \{c \in \mathcal{C}_H(G) \mid c(A) \subseteq X,\ c(B) \subseteq Y\}.$$

This is seen from the following bijection which maps a $c \in \mathcal{C}_H(G, X, Y)$ to the coloring

$$\langle c^{-1}(u_1),\ \ldots\ c^{-1}(u_a),\ c^{-1}(\{u_{a+1}, v_{a+1}\}),\ \ldots\ , c^{-1}(\{u_{k-a}, v_{k-a}\}),$$

$$c^{-1}(v_{k-a+1}),\ \ldots\ , c^{-1}(v_k))\rangle.$$

Just as in Example 7.1, one can derive negative results on sampling using cautious chains. The proof is again essentially the same as the proof of Theorem 1.1. Also, Glauber dynamics is ergodic for such H-colorings also.

7.3 Example

Suppose H is the complete graph plus just one loop. The preceding arguments also show that our torpid mixing results can be applied to H-colorings for this specific H. This is because if H is the complete graph on $V = \{1, 2 \ldots, k\}$ with a self-loop on vertex 1, then any H-coloring $\langle X_1, \ldots X_k\rangle$ of $G = (A \cup B, E)$ can be thought of as a standard $(k + 1)$-coloring $\langle Y_1, \ldots Y_{k+1}\rangle$ of G where

$$Y_1 = X_1 \cap A,\ \ Y_{k+1} = X_1 \cap B,\ \ Y_i = X_i\ \ \forall i = 2, \ldots, k.$$

This mapping establishes a bijection between $\mathcal{C}_H(G)$ and $\mathcal{C}_{1,k+1}(G)$ defined before. Hence, the negative results of Theorem 1.2 apply to this specific H. Also, the Glauber dynamics chain is ergodic in this case (follows from Theorem 2, [2]).

8 Conclusions

An interesting open problem is to design efficient samplers for k-colorings ($k \geq 3$) of bipartite graphs? An affirmative answer would lead to efficient randomized approximate counters for this case. The results of this paper indicate that such samplers may have to use tools other than Markov chains.

It seems very likely that the negative results obtained can be extended to k-colorings of l-colorable ($l \geq 3$) graphs where $l < k = O(n^\delta)$ for some positive δ which depends only on l. Finding a k-coloring is NP-hard for any fixed $k \geq 3$ except for some special classes like bipartite graphs, chordal graphs, etc. Even if we ignore the complexity of finding a starting state, the previous remark shows that markov chains are not likely to be useful as efficient samplers.

It is known that approximate counting and almost uniform sampling problems are polynomial time reducible to each other provided $k > \Delta(G)$ (see [8]). However, it would be interesting to see if this can be extended to $k \leq \Delta(G)$ for some special classes of graphs. For example, this reduction works also for bipartite G and for any $k \geq 3$. We can show this by a simple modification of the arguments of [8]. But, we skip the proof of this observation since we present only negative results on efficiently sampling colorings of bipartite graphs. A related question is : Are there other classes of graphs where the equivalence is true for $k \leq \Delta$?

Acknowledgements

We thank Eric Vigoda for the useful comments provided by him. We are also very grateful to the anonymous referee whose questions and comments prompted us to obtain the generalization to H-colorings (Subsection 1.2) and also to come up with explicit colorings admitting 1-cautious ergodic chains (Section 7).

References

[1] R. Bubley, M.E. Dyer, C. Greenhill, “Beating the 2Δ bound for approximately counting colorings: A computer-assisted proof of rapid mixing”, *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA’98).*

[2] C. Cooper, M.E. Dyer, A.M. Frieze, “On Markov chains for randomly H-colouring a graph”, *Journal of Algorithms*, **39**(1), April 2001, 117-134.

[3] M.E. Dyer and A.M. Frieze, “Randomly colouring graphs with lower bounds on girth and maximum degree”, *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science (FOCS’01)*, pages 579-587.

[4] M.E. Dyer, A.M. Frieze, M.R. Jerrum, “On counting independent sets in sparse graphs”, *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science (FOCS’99).*

[5] M.E. Dyer, C. Greenhill, “A more rapidly mixing Markov chain for graph colorings”, *Random Structures & Algorithms*, Vol. 13 (1998), pp. 285-317.

[6] Thomas P. Hayes, “Randomly Coloring Graphs of Girth at least Five”, *Proceedings of the 45th Annual ACM Symposium on Theory of Computing (STOC’03).*

[7] Thomas P. Hayes and E. Vigoda, “A Non-Markovian Coupling for Randomly Sampling Coloring”, *Proceedings of the 43rd IEEE Symposium on Foundations of Computer Science (FOCS’03)*.

[8] M.R. Jerrum, “A very simple algorithm for estimating the number of k-colorings of a low-degree graph”, *Random Structures and Algorithms*, 7(2), 1995.

[9] M.R. Jerrum and A. Sinclair, “The Markov chain Monte Carlo method : an approach to approximate counting and integration”, In *Approximation Algorithms for NP-hard Problems*, D. Hochbaum, ed., PWS, 1996, 482-520.

[10] T. Luczak and E. Vigoda, “Torpid mixing of the Wang-Swendsen-Kotecky algorithm for sampling colorings”, To appear in *J. of Discrete Algorithms*.

[11] M. Molloy, “The Glauber dynamics on colorings of a graph with high girth and maximum degree”, *Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC’02)*, pages 91-98.

[12] R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge University Press, 1995.

[13] J. Salas and A.D. Sokal, “Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem”, *Journal of Statistical Physics*, 86(3-4):551-579, 1997.

[14] A.J. Sinclair, Randomized Algorithms for Counting and Generating Combinatorial Structures. PhD thesis, Department of Computer Science, University of Edinburgh, 1988.

[15] A.J. Sinclair and M.R. Jerrum, “Approximate counting, uniform generation, and rapidly mixing Markov chains”, *Information and Computation* **82** (1989), 93-133.

[16] E. Vigoda, “Improved bounds for sampling colorings”, *Journal of Mathematical Physics*, 41 (2000), pp. 1555-1569. Also appeared in *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science (FOCS’99)*.