A survey on the use of Markov chains to randomly sample colorings

Alan Frieze* Eric Vigoda†

October 11, 2005

Dedicated to Dominic Welsh on the occasion of his 65th Birthday

Abstract

In recent years there has been considerable progress on the analysis of Markov chains for generating a random coloring of an input graph. These improvements have come in conjunction with refinements of the coupling technique, which is a classical tool in probability theory. We survey results on generating random colorings, and related technical improvements.

1 Introduction

Our focus is on Markov Chain Monte Carlo (MCMC) algorithms for approximately counting the number of k-colorings of a graph. For a graph $G = (V, E)$, a (proper) k-coloring is an assignment $\sigma : V \to [k]$ such that adjacent vertices receive different colors.

It is well known that it is NP-hard to compute the minimum number of colors in a proper k-coloring. On the other hand there are many known conditions that imply that G is k-colorable. For example, unless G is an odd cyle or a complete graph, then Brook's theorem states that G can be

*Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh PA 15213. Research supported in part by NSF grant CCF-0502793.

†College of Computing, Georgia Institute of Technology, Atlanta GA 30332. Research supported in part by NSF grant CCF-0455666

Δ-colored, where $\Delta = \Delta(G)$ is the maximum degree of G. In the two special cases we need $\Delta + 1$ colors. Thus if $k \geq \Delta + 1$ then we can k-color G. This raises the interesting computational challenge of finding the number of proper k-colorings for $k \geq \Delta + 1$.

Valiant [36] introduced the notion of #P-hardness, and proved that counting colorings is #P-complete. Hence, our focus is on whether or not there is an efficient approximation algorithm for counting proper colorings. An important paper by Karp, Luby and Madras [29] introduced the notion of a Fully Polynomial time Randomized Approximation Scheme (FPRAS). In the case of colorings, an FPRAS is a randomised algorithm that, for all $1 > \epsilon > 0$, can with high probability (i.e., probability at least $1 - 1/poly(n)$) estimate the number of k-colorings within a multiplicative factor $1 \pm \epsilon$ in time polynomial in $n = |V|$ and $1/\epsilon$.

Dominic Welsh has made major contributions to the computational theory of counting and has written a beautiful monograph on the subject [38]. The paper [25] by Jerrum, Valiant and Vazirani relates the estimation of the size of a finite set S to the uniform random choice of a member of S. In particular, it shows how given a (near) uniform sampler, we can construct an FPRAS for $|S|$. In this short survey, we discuss the success or otherwise in the quest to find an algorithm that can generate a (near) uniform k-coloring of a graph G.

Markov Chain Monte Carlo is an important tool in sampling from complex distributions such as the uniform distribution on k-colorings. It has been successfully applied in several areas of Computer Science, most notably volume computation [11], [28], [30] and estimating the permanent of a non-negative matrix [26].

The idea is to define an ergodic Markov chain whose state space is the same as the set from which we wish to sample and whose steady state distribution is the required distribution. We then run the chain until it is close to its steady state and use the current state as a sample point. The main issue in all of this is to determine how long it is required to run the chain before it is close enough to the steady state. To make this precise, let Ω denote the state space of the Markov chain $X_0, X_1, \ldots, X_t$, and let π denote the steady state distribution. Let $P^{(t)}(\sigma, \tau)$ denote the probability that $X_t = \tau$, given that the chain started with $X_0 = \sigma$. Define the mixing time as the time until is the chain is within variation distance $\leq 1/4$ from the stationary

distribution from the worst initial state, i.e., let

$$\tau_{\text{mix}} = \max_{\sigma \in \Omega} \min\{t : \left\| P^{(t)}(\sigma, \cdot) - \pi \right\|_{\text{TV}} \leq 1/4\}.$$

Recall the (total) variation distance for a pair of distributions μ and ν on a finite space Ω is defined as

$$\|\mu - \nu\|_{\text{TV}} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)| = \max_{A \subset \Omega} \mu(A) - \nu(A).$$

The constant $1/4$ in the definition of the mixing time is somewhat arbitrary and simply needs to be $< 1/2$. It then follows, for any $\delta > 0$, that after $\leq \tau_{\text{mix}} \log(1/\delta)$ iterations, the chain is within variation distance $\leq \delta$ of the stationary distribution (see, for example, [24, 3]).

1.1 Glauber Dynamics and Survey of Results

There is one particular Markov chain that has been the subject of much interest. For $G = (V, E)$, the state space of the Markov chain is $[k]^V = \{1, \ldots, k\}^V$. Let $N(v)$ denote the neighbors of vertex v. The following Markov chain is known as the Glauber dynamics. From $X_t \in [k]^V$,

1. Choose z uniformly at random from V, and c uniformly at random from $\{1, \ldots, k\}$.
2. For all $w \neq z$, set $X_{t+1}(z) = X_t(z)$.
3. If no neighbors of z have color c (i.e., $c \notin X_t(N(z))$), then set $X_{t+1}(z) = c$, otherwise set $X_{t+1}(z) = X_t(z)$.

The version above is called *Metropolis Dynamics*.

While the coloring problem came to the fore in 1995 with the publication of [23], already in 1983 Donnelly and Welsh [7] had considered a Markov chain for 2-coloring graphs, viz. the antivoter problem.

Note, while the chain is defined on all labelings of G, when $k \geq \Delta + 2$, for any initial state X_0, we eventually reach a proper k-coloring, and then the chain only walks on proper colorings. Moreover, it is straightforward to verify that the Glauber dynamics is irreducible on the set Ω of proper k-colorings of G when $k \geq \Delta + 2$. Thus, its stationary distribution π is uniform over the proper k-colorings of G.

A natural conjecture is that the mixing time of the Glauber dynamics is $O(n \log n)$ whenever $k \geq \Delta + 2$. Note, $O(n \log n)$ steps are needed so that every vertex is recolored at least once. Recently Hayes and Sinclair [18] proved the mixing time is in fact $\Omega(n \log n)$, at least for constant degree graphs.

There has been steady progress on the upper bound, and the following result now appears within reach: $O(n \log n)$ mixing time for $k \geq \Delta + o(\Delta)$. The first significant result on this problem (at least in the theoretical computer science community) was by Jerrum, who proved $O(n \log n)$ mixing time of the Glauber dynamics whenever $k > 2\Delta$. His proof is closely related to work in the statistical physics community, e.g., Salas and Sokal [35], on the Dobrushin uniqueness condition. In statistical physics terminology, counting colorings corresponds to computing the partition function of the zero-temperature anti-ferromagnetic Potts model.

Jerrum's proof was simplified using the path coupling approach introduced by Bubley and Dyer [5]. Vigoda then introduced an alternative Markov chain and proved $O(n \log n)$ mixing time of this new chain whenever $k > 11\Delta/6$. This implied $O(n^2)$ mixing time of the Glauber dynamics. Vigoda's $11\Delta/6$ result is still the best upper bound for general graphs.

Dyer and Frieze considered graphs with large girth ($g = \Omega(\log \Delta)$) and large degree ($\Delta = \Omega(\log n)$) in order to use properties of typical colorings. They proved $O(n \log n)$ mixing time of Glauber dynamics when $k > \alpha\Delta$ where $\alpha \approx 1.763...$ for this class of large girth/degree graphs. A series of results reduced the bound on k/Δ (Molloy [33], and Hayes and Vigoda [19]), lowered the girth requirement (Hayes [17], Hayes and Vigoda [20], Frieze and Vera [14]), and lowered the degree requirement (Dyer, Frieze, Hayes and Vigoda [10]). All of these results prove $O(n \log n)$ mixing time of the Glauber dynamics with various improvements in the parameters.

The two most notable results in this framework are Hayes and Vigoda [19] whose result holds for all $\epsilon > 0$, all $k > (1 + \epsilon)\Delta$ with girth $g \geq 11$ and $\Delta = \Omega(\log n)$; and Dyer et al [10] which holds for all $\epsilon > 0$, all $k > (1+\epsilon)\beta\Delta$ where $\beta \approx 1.489...$ with girth $g \geq 6$ and Δ a sufficiently large constant (which grows with $1/\epsilon$). When above the threshold $k/\Delta = \alpha$ we have rapid mixing for triangle free graphs [20] and more generally, graphs with sparse vertex neighbourhoods, e.g. planar graphs [14]. In this survey we focus on explaining the high-level ideas in this body of work. We introduce this approach in a non-historical, but in what we hope is a more intuitive manner.

This is a good place to mention the experimental work of Petford and Welsh [34] on the *Anti-Voter Algorithm* for 3-coloring a graph G. Their goal was to use a Markov chain to find a single proper 3-coloring. Starting with an arbitrary (non-proper) 3-coloring of G the algorithm chooses a vertex v which has the same color as one of its neighbors and randomly re-colors it according to a distribution proportional to $e^{-\lambda_c}$ where λ_c is the current number of neighbours of v with color c. The algorithm continues until G is 3-colored. They report good results up to average degree about 5.

1.2 Outline of Paper

In Section 2 we explain the coupling method, including the path coupling approach. We give a detailed proof of Jerrum's 2Δ result in Section 3. Then in Section 4 we explain how we can use typical properties of random colorings to improve upon the worst case coupling approach. These results require high probability events, and thus require $\Delta = \Omega(\log n)$. In Section 5 we explain how the approach extends to constant degree graphs. The constant k/Δ is further reduced in Section 6. In Section 7 we give a very rough idea of the $k > (1+\epsilon)\Delta$ result of Hayes and Vigoda.

In Section 8 we consider the alternative Markov chain used by Vigoda in his $k > 11\Delta/6$ result. Finally in Section 9 we mention related results, and present several interesting open problems in Section 10.

2 Coupling Method

The coupling method has been the primary tool for analyzing Markov chains in the context of random colorings. It was first introduced for the analysis of Markov chains by Doeblin [6]. For a pair of distributions μ, ν on a discrete space Ω, a coupling is any distribution ω on $\Omega \times \Omega$ where:

$$\sum_{y \in \Omega} \omega(x,y) = \mu(x), \qquad \text{for all } x \in \Omega,$$

$$\sum_{x \in \Omega} \omega(x,y) = \nu(y), \qquad \text{for all } y \in \Omega.$$

Thus, if we simply observe the first coordinate, it is distributed as μ, and similarly for the second coordinate with respect to ν. It is then straightforward

to show that

$$\|\mu - \nu\|_{\mathrm{TV}} \leq 1 - \sum_{z \in \Omega} \omega(z, z). \tag{1}$$

Therefore we can upper bound the variation distance by defining a coupling and bounding the probability that the coordinates differ. In fact, there always exists a coupling which achieves equality in (1).

2.1 Worst Case Coupling

We will consider couplings for Markov chains. Consider a pair of Markov chains (X_t) and (Y_t) on Ω with transition matrices P_X and P_Y, respectively. The Markov chain (X'_t, Y'_t) on $\Omega \times \Omega$ is a (Markovian) coupling if

$$\Pr(X'_{t+1} = c \mid X'_t = a, Y'_t = b) = P_X(a, c), \text{ for all } a, b, c \in \Omega; \text{ and}$$

$$\Pr(Y'_{t+1} = c \mid X'_t = a, Y'_t = b) = P_Y(b, c), \text{ for all } a, b, c \in \Omega.$$

In other words, if we simply observe the first coordinate, it behaves like P_X and similarly the second coordinate acts according to P_Y. This is a more restrictive form of coupling then is necessary. In general, the joint distribution of X_{t+1}, X'_{t+1} can depend on the whole previous history and as such is called a non-Markovian coupling. We will see an application of a non-Markovian coupling in Section 7.

For a coupling (X'_t, Y'_t) of (X_t) and (Y_t), by (1), we have for all X_0, Y_0,

$$\|X_t - Y_t\|_{\mathrm{TV}} \leq \Pr(X'_t \neq Y'_t \mid X'_0 = X_0, Y'_0 = Y_0), \tag{2}$$

where we are using the random variable X_t to denote its probability distribution, and similarly for Y_t.

Let d denote an arbitrary integer-valued metric on Ω, and let $d_{\max}$ denote the diameter of Ω under this metric. For $\epsilon > 0$, we say a pair $(x, y) \in \Omega^2$ is *ϵ distance decreasing* if there exists a coupling such that

$$\mathbf{E}\left(d(X_1, Y_1) \mid X_0 = x, Y_0 = y\right) < (1 - \epsilon) d(x, y).$$

The Coupling Theorem says that if *every pair* (x, y) is ϵ distance decreasing, then the Markov chain mixes rapidly:

Theorem 1 (cf. Aldous [2]). *Let $\epsilon > 0$ and suppose every $(x, y) \in \Omega^2$ is ϵ distance decreasing. Then,*

$$\tau_{\mathrm{mix}} \leq 2 \log(d_{\max}) \epsilon^{-1}.$$

Proof. By induction we have

$$\begin{aligned}
\mathbf{Pr}\left(X_t \neq Y_t \mid X_0, Y_0\right) &\leq \mathbf{E}\left(d(X_t, Y_t) \mid X_0, Y_0\right) \\
&\leq (1-\epsilon)\mathbf{E}\left(d(X_{t-1}, Y_{t-1}) \mid X_0, Y_0\right) \\
&\leq (1-\epsilon)^t d(X_0, Y_0) \\
&\leq \exp(-t\epsilon) d_{\max} \\
&\leq 1/4
\end{aligned}$$

for $t \geq 2\log(d_{\max})/\epsilon$. The theorem now follows from (2). □

2.2 Path Coupling

The path coupling lemma says it suffices to define and analyze a coupling for pairs in a subset of $\Omega \times \Omega$ assuming the subset "connects" the state space. Let $S \subseteq \Omega \times \Omega$ be such that (Ω, S) is connected. For $(X, Y) \in \Omega \times \Omega$, define $d(X, Y)$ as the shortest path length between X and Y in the graph (Ω, S). Thus, if $(X, Y) \in S$, then $d(X, Y) = 1$.

Lemma 2 (Bubley and Dyer [5]). *Let $\epsilon > 0$. Suppose (Ω, S) is connected and every $(x, y) \in S$ is ϵ distance decreasing. Then,*

$$\tau_{\text{mix}}(\delta) \leq 2\log(d_{\max})/\epsilon.$$

Proof. We construct a coupling for all $(X_t, Y_t) \in \Omega \times \Omega$ which is distance decreasing, and then we apply Theorem 1. We construct the coupling for an arbitrary pair X_t, Y_t by simply "composing" couplings along a shortest path between X_t and Y_t. Let $Z_t^0 = X_t, Z_t^1, \ldots, Z_t^j = Y_t, j = d(X_t, Y_t)$ denote an arbitrary such shortest path.

From the hypothesis of the lemma we know

$$\mathbf{E}\left(d(Z_{t+1}^i, Z_{t+1}^{i+1}) \mid Z_t^i, Z_t^{i+1}\right) \leq (1-\epsilon) d(Z_t^i, Z_t^{i+1}).$$

We then have

$$\begin{aligned}
\mathbf{E}\left(d(X_{t+1}, Y_{t+1}) \mid X_t, Y_t\right) &\leq \sum_{1 \leq i < j} \mathbf{E}\left(d(Z_{t+1}^i, Z_{t+1}^{i+1}) \mid Z_t^i, Z_t^{i+1}\right) \\
&\leq (1-\epsilon) \sum_{1 \leq i < j} d(Z_t^i, Z_t^{i+1}) \\
&= (1-\epsilon) d(X_t, Y_t)
\end{aligned}$$

The lemma now follows from Theorem 1. □

3 Worst Case Coupling for Glauber

We will prove fast convergence via path coupling. We begin with the following theorem to get started.

Theorem 3. *Let Δ denote the maximum degree of G. If $k > 3\Delta$, then*

$$\tau_{\text{mix}} \leq 2nk \log(n).$$

Proof. Let $\Omega = [k]^V$ denote the state space of the Glauber dynamics. For simplicity we will refer to the states of Ω as colorings. In this set up, the distance

$$d(X, Y) = |\{v \in V : \ X(v) \neq Y(v)\}|$$

corresponds to the Hamming distance between colorings X and Y. We apply the path coupling lemma with the subset S defined as pairs of colorings that differ at a single vertex, i.e.,

$$S = \{X, Y \in \Omega \times \Omega : d(X, Y) = 1\}$$

Note, since Ω is $[k]^V$ (not just proper colorings), for all $X, Y \in \Omega$ the length of the shortest path between X and Y in the graph (Ω, S) is of length $d(X, Y)$. Thus we can apply the path coupling lemma with S and distance d defined in this manner.

Now we need to define a coupling for pairs of colorings in S. Consider a pair of such colorings $(X_t, Y_t) \in S$ where $X_t(v) \neq Y_t(v)$, say $X_t(v) = c_X$ and $Y_t(v) = c_Y$. The coupling is simply the identity coupling. Both chains attempt to update the same vertex z to the same color c. Observe that only updates with $z \in N(v)$ and $c \in \{c_X, c_Y\}$ might succeed or fail in exactly one chain. All other updates succeed or fail in both chains. Thus, these are the only updates which might increase the distance. In summary, there are at most 2Δ updates which might increase the distance, and each occurs with probability $1/kn$.

The only updates which might decrease the distance are successful recolorings of v. Since there are at most Δ colors in the neighborhood of v, there are at least $k - \Delta$ available colors for v.

Combining these contributions, we have

$$\mathbf{E}\left(d(X_{t+1}, Y_{t+1}) - d(X_t, Y_t) \mid X_t, Y_t\right) \leq \frac{1}{kn}\left(2\Delta - (k - \Delta)\right) < -\frac{1}{kn},$$

for $k > 3\Delta$.

Plugging this bound into the path coupling lemma proves the theorem. □

Modifying the above coupling we can achieve the following improvement.

Theorem 4 (Jerrum [23]). *If $k > 2\Delta$, then*

$$\tau_{\text{mix}} \leq 2nk \log(n).$$

Proof. The set S remains the same, we simply modify the coupling slightly. If X_t attempts to recolor a neighbor w of v to color c_X, then Y_t attempts to recolor w to c_Y. Similarly, if X_t attempts to recolor $w \in N(v)$ to c_Y, then Y_t attempts to recolor w to c_X. In all other cases, X_t and Y_t attempt to modify the same vertex to the same color.

Observe that $w \in N(v)$ can not receive color c_X in X_t (since this is v's color), and w can not receive c_Y in Y_t. Thus, such a coupled update has no effect on the pair of colorings. Therefore, there are at most Δ coupled recolorings which might increase the distance – attempting to recolor $w \in N(v)$ to c_Y in X_t and to c_X in Y_t.

We now have

$$\mathbf{E}\left(d(X_{t+1}, Y_{t+1}) - d(X_t, Y_t) \mid X_t, Y_t\right) \leq \frac{1}{kn}\left(\Delta - (k - \Delta)\right) < -\frac{1}{kn},$$

for $k > 2\Delta$. □

Further Improvements

Notice that in the above proof for $k > 2\Delta$, we needed that the number of valid recolorings of v is greater than Δ. While in the worst case v has Δ distinct colors in its neighborhood, this might be unlikely in a typical coloring. In the next section we look at properties of random colorings, and how these properties can be used to improve upon the $k > 2\Delta$ result.

In Section 8 we look at a different Markov chain, and show how that improves upon the 2Δ result.

4 Using Properties of Typical Colorings

In an arbitrarily chosen coloring, a vertex v only has $\geq k - \Delta$ valid recolorings, which was used in the $k > 2\Delta$ bound. However, if G has few edges among

neighbors of v, then in a random coloring one expects that many neighbors of v have the same color, and thus v has more valid recolorings. In particular, we want to lower bound the number of available colors for v. Let

$$A(X, v) := [k] \setminus X(N(v)),$$

denote the set of available colors for a vertex v under a k-coloring X.

It is in fact easy to lower bound the expected number of available colors for any vertex in a triangle-free graph, as we show in the next lemma. By setting $\Delta = \Omega(\log n)$ we get the lower bound on the number of available colors with high probability for every vertex. It is then straightforward to combine this high probability event with the coupling framework to conclude rapid mixing of the Glauber dynamics for an improved range of k/Δ.

Note, if the neighbors of a vertex v were each receiving an independent random color, then the expected number of colors which do not appear in this neighborhood (i.e., the number of available colors) is $k(1 - 1/k)^{\Delta}$ if v had degree Δ. For large k this is roughly $k \exp(-\Delta/k)$, which is exactly the bound that we obtain in the following lemma.

Lemma 5 (Hayes and Vigoda [20]). *Let $G = (V, E)$ be a triangle-free graph with maximum degree Δ. For every $1 \geq \delta > 0$, there exists $C > 0$ such that, whenever $k \geq \max\{\Delta + 2/\delta, C \log n\}$, then for a random k-coloring X,*

$$\mathbf{Pr}\left(\exists v \in V : \ |A(X, v)| < k \exp(-\Delta/k)(1 - \delta)\right) \leq 1/n^{10}.$$

The above lemma is interesting in the case $\Delta \geq C \log n$, and then we only need $k = \Delta + \Omega(1)$ for the above property to hold with high probability. Notice the theorem is a statement about a random coloring. Thus, it is property of the stationary distribution of the Glauber dynamics, but has no connection to the transitions of the Markov chain.

Before proving the lemma let us explain how this improves upon Jerrum's 2Δ result. Recall the proof of Theorem 1 (Coupling Theorem). The proof couples an arbitrary pair of states. Our goal is to get to the stationary distribution, thus we could instead couple a chain starting at an arbitrary initial state X_0 with a chain starting in the stationary distribution $Y_0 \sim \pi$. By the definition of the stationary distribution, since $Y_0 \sim \pi$, then for all $t \geq 0$, we have $Y_t \sim \pi$. Hence, once the chains have coupled, i.e., $X_t = Y_t$, then the chain X_t has reached the stationary distribution.

The above Lemma 5 implies that Y_t locally looks random with high property. We call this a local uniformity property. For $T = O(n \log n)$, we will

prove that if Y_t, for all $0 \leq t \leq T$, has the local uniformity property, then we can couple $X_T = Y_T$ with probability $\geq 1 - 1/10n$. Using Lemma 5, with a union bound, we have that for all $0 \leq t \leq T$, Y_t has the local uniformity property, except with probability at most $T/n^{10} < n^{-8}$. In the case that some Y_t does not satisfy the local uniformity property we can bound the Hamming distance of X_T and Y_T as $d(X_T, Y_T) \leq n$. Hence, for $Y_0 \sim \pi$,

$$\mathbf{E}\left(d(X_T, Y_T) \mid X_0\right) \leq n(1/10n + 1/n^8) \leq 1/4.$$

One complication in the above approach is that it requires defining and analyzing a coupling for an arbitrary pair of states. In other words, we can no longer use the path coupling lemma. Jerrum's original proof of the $k > 2\Delta$ result did not use path coupling. Moreover, in his proof it turns out that it suffices for one of the coupled pairs to have the local uniformity property, in our case only Y_t has the property.

To get intuition for the effect of the local uniformity property on the coupling, recall the discussion after the proof of Theorem 4. Note, the proof required that the number of valid recolorings of the disagree vertex v was greater than Δ. The number of valid recolorings for v is the number of colors not appearing in its neighborhood, which is lower bounded in Lemma 5. Therefore, we need

$$k \exp(-\Delta/k) > \Delta$$

Let $\alpha = 1.763...$ denote the solution to $x = \exp(1/x)$. It then follows that for all $\epsilon > 0$, there exists $C > 0$ such that the Glauber dynamics has mixing time $O(n \log n)$ for any triangle-free graph with maximum degree Δ when $k \geq \max\{(1+\epsilon)\alpha\Delta, C \log n\}$.

We refer the reader to Hayes and Vigoda [20] for details, and to Frieze and Vera [14] for an improvement to locally-sparse graphs.

We now show how to prove Lemma 5. The main idea is to consider a random coloring X and any vertex v. Then we simply recolor the neighborhood of v, with all other vertices fixed to X. Since v is triangle-free we can simultaneously recolor all neighbors of v, i.e., their colors do not interact since X is fixed on $V \setminus N(v)$. Moreover, after this process we will still have a random coloring. Note, every $w \in N(v)$ will be rechoosing its new color from a reasonably large set of possible colors. Hence, many neighbors will choose the same color, and we will get the expected behaviour. Here is the formal proof.

Proof of Lemma 5. Let $v \in V$. By definition,

$$|A(X,v)| = \sum_{j\in[k]} \prod_{w\in N(v)} (1 - X_{j,w}),$$

where $X_{j,w}$ is the indicator variable for vertex w having color j, i.e., the event $\{X(w) = j\}$. Since $\delta \leq 1$, we always have $|A(X,v)| \geq 2$.

We will fix the coloring X on $V \setminus N(v)$, and consider the distribution of colors on $N(v)$. Let $\mathcal{F}$ denote the fixed coloring X on $V \setminus N(v)$.

Since G is triangle-free, given $\mathcal{F}$, each $X(w)$ is independent and uniformly distributed over its set $|A(X,v)|$. Note, for a color j,

$$\mathbf{E}\left(X_{j,w} \mid \mathcal{F}\right) = \begin{cases} \frac{1}{|A(X,w)|} & \text{if } j \in A(X,w) \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

Hence,

$$\begin{aligned} \mathbf{E}\left(|A(X,v)| \mid \mathcal{F}\right) &= \sum_{j\in[k]} \prod_{w\in N(v)} \left(1 - \mathbf{E}\left(X_{j,w} \mid \mathcal{F}\right)\right) \\ &\geq k \prod_{j\in[k]} \prod_{w\in N(v)} \left(1 - \mathbf{E}\left(X_{j,w} \mid \mathcal{F}\right)\right)^{1/k} \\ &= k \prod_{w\in N(v)} \left(1 - \frac{1}{|A(X,w)|}\right)^{|A(X,w)|/k} \\ &\geq k(\exp(-\Delta/k) - \delta/2), \end{aligned}$$

where the first inequality holds by the arithmetic-geometric means inequality, the subsequent identity follows from (3), and the final inequality uses $k \geq \Delta + 2/\delta$.

Since the colors $X(w), w \in N(v)$ are independent, and $|A(X,v)|$ is a Lipschitz function of these colors, with constant 1, it follows by Hoeffding's theorem [22] that

$$\mathbf{Pr}\left(|A(X,v)| < k\left(\mathrm{e}^{-\Delta/k} - \delta\right) \mid \mathcal{F}\right) \leq \mathrm{e}^{-\delta^2 k/2}.$$

Since this holds for every outcome of $\mathcal{F}$,

$$\mathbf{Pr}\left(|A(X,v)| < k\left(\mathrm{e}^{-\Delta/k} - \delta\right)\right) \leq \mathrm{e}^{-\delta^2 k/2} < 1/n^{11},$$

for sufficiently large C. Note C depends on δ. Also, by choosing C sufficiently large we get an arbitrarily small polynomial for the tail probability. Taking a union bound over $v \in V$ completes the proof. □

5 Constant Degree

The lower bound on the number of available colors holds with high probability because $\Delta = \Omega(\log n)$. In this section we describe how these results extend to constant degree graphs. The first observation is that in the proof of Lemma 5 for general Δ, the tail probability is roughly $\exp(-\Delta)$. Thus, in constant degree graphs, there will be a constant fraction of the vertices which do not achieve the desired lower bound as in Lemma 5. We will need a new approach.

The lower bound on the number of available colors held for a random coloring. In fact, it also holds for an arbitrary coloring after a sufficient number of steps (roughly $O(n \log \Delta)$) of the Glauber dynamics. This is much more difficult to prove, but crucial for the constant degree results in this section. Roughly speaking, we can mimic the recoloring of the neighborhood of a vertex v,

Recall, in the proof of Lemma 5, the key step was recoloring the neighborhood of a vertex v. For an arbitrary coloring, after $O(n \log \Delta)$ steps of the Glauber dynamics, most of the neighborhood of v will be recolored at least once. Thus, we might expect after this many steps, the coloring has similar behavior as in the proof of Lemma 5. This is the case, but the neighbors of v are not being recolored simultaneously. Thus, their new colors are not independently chosen. This independence was key to the simplicity of the proof of Lemma 5. Overcoming these dependencies is much more difficult.

Dyer and Frieze first assumed girth $\Omega(\log \log \Delta)$, and used a disagreement percolation argument to argue the neighbors are being recolored nearly independently. Hayes reduced the girth to 5, by coupling with the Glauber dynamics on a modified graph where the local neighborhood of v is disconnected from the rest of the graph. Here is the formal statement of Hayes' result.

Lemma 6 (Hayes [17]). *For every $\epsilon > 0$ there exists $C_{min} > 0$ such that for every graph $G = (V, E)$ with maximum degree $\Delta > C_{min}$ and girth at least 5, for $k > \alpha\Delta$, all $X_0 \in \Omega$, for every $t > C_{min} n \log \Delta$, for all $v \in V$,*

$$\mathbf{Pr}\left(A(X_t, v) < (1-\epsilon)k \exp(-\Delta/k)\right) \leq \exp(-\epsilon^2 \Delta/100).$$

Using this lemma, Dyer, Frieze, Hayes and Vigoda [10] were able to reduce the lower bound on maximum degree to *sufficiently large* as opposed to $\Omega(\log n)$ in previous papers.

Their approach uses path coupling for an arbitrary pair of initial colorings. Thus, consider an arbitrary pair of colorings X_0, Y_0 which differ at a single vertex v. They then couple this pair using the same coupling as used by Jerrum in his $k > 2\Delta$ result. However, they analyze the coupling over $T = O(n \log \Delta)$ steps (whereas Jerrum simply used $T = 1$).

They begin by running the pair for $O(n \log \Delta)$ steps. During this initial period the Hamming distance will increase in expectation at each step. This is a "burn-in" period. Looking at a constant radius ball of vertices around v, during the burn-in period, we expect most vertices in the ball will achieve the lower bound on the number of available colors. Moreover, even though the disagreement will likely spread away from v, with a large constant probability it is unlikely to spread outside of the local ball around v.

There are then two cases after the initial burn-in period. In the good case, all of the disagreements are contained in the local ball around v, and most vertices in the ball have the uniformity properties of Lemma 6. In this case the Hamming distance will then decrease in expectation. Also, the chains are coupled for a sufficient number of further steps so that the expected Hamming distance is small at the conclusion of the coupling. In the bad case, they simply obtain a reasonable upper bound on the Hamming distance, and prove this is offset by the small probability of actually reaching this bad case.

6 Further Properties of Typical Colorings

The above improvements to $k > 1.763\Delta$ used a lower bound on the number of available colors. If the neighbors of a vertex v were really receiving independent random colors, then in expectation the number of available colors for v would be $k(1 - /k)^{deg(v)} \sim k \exp(-k/deg(v))$. Thus, we have not yet proven that neighbors of v are getting random color choices. Improving beyond 1.763 requires further "local uniformity properties" which require that in a random coloring (or after a sufficient number of steps of Glauber dynamics) vertices are receiving (close to) uniform random color choices.

To get the matching upper bound on the number of available colors (or prove stronger local uniformity properties) requires a recursive argument. In particular, we need to successively improve the upper bound on available colors. The details of this recursion is complicated, and we refer the interested reader to Molloy [33]. Molloy first proved the desired upper bound on

available colors in the context of graphs with girth $\Omega(\log\log\Delta)$, and used it to improve the 1.763... threshold to 1.489....

A simpler set of recurrences were later used by Hayes (with only constant girth requirements). The recurrences used by Hayes are very similar to those used earlier by Jonasson to study uniqueness of Gibbs measures for colorings of the infinite tree [27]. The expanded girth requirement is a result of having to recolor a larger neighborhood in the proof of the local uniformity property.

6.1 1.489... Threshold

In Jerrum's 2Δ proof, we had a pair of colorings X, Y which differed only at v, say v had c_X =Red in X and c_Y =Blue in Y. Then, the (Hamming) distance decreased if we successfully recolored v in both chains. In the worst case there are only $k-\Delta$ valid recolorings of v. The improvements to 1.763Δ replaced this worst case bound with $\geq k\exp(-\Delta/k)$ valid recolorings.

On the other side, there are at most Δ recolorings which increase the distance. In particular, for every $w \in N(v)$, if w can be recolored to Red in Y and/or to Blue in X, then there's a coupled recoloring of w that increases the distance. If neighbors of w were colored independently at random then (assuming w is degree Δ) with probability

$$(1-(1-1/k)^{\Delta-1})^2 \sim (1-\exp(-\Delta/k))^2,$$

at least one neighbor of w (ignoring v) is colored Blue and at least one neighbor is Red. Hence, the attempted coupled update of w to Red in Y and Blue in X fails in both chains. Thus, in expectation (if neighbors of w are colored randomly) we only have $\Delta(1-(1-\exp(-\Delta/k))^2)$ recolorings that increase the distance. We might expect that the proof approach works when:

$$k\exp(-\Delta/k) > \Delta(1-(1-\exp(-\Delta/k))^2)$$

In other words, for $x = k/\Delta$, we want

$$x\exp(-1/x) + (1-\exp(-1/x))^2 > 1.$$

This holds when $x > 1.489...$.

Using the above approach, Molloy [33] proved $O(n\log n)$ mixing time of the Glauber dynamics when $k/\Delta > 1.489...\Delta$ for $\Delta = \Omega(\log n)$ and girth $g = \Omega(\log\log n)$. Hayes [17] improved the girth requirement to $g \geq 6$. Dyer, Frieze, Hayes and Vigoda [10] reduced the degree requirement to Δ sufficiently large constant.

7 Getting close to the Maximum Degree

Hayes and Vigoda [19] improved the threshold on k/Δ to 1. In particular, they proved $O(n \log n)$ mixing time of the Glauber dynamics for $k > (1+\epsilon)\Delta$ for all $\epsilon > 0$ assuming $\Delta = \Omega(\log n)$ and the girth $g \geq 11$.

Their result differs from all of the previously mentioned results in the use of a much more sophisticated coupling. The earlier results used the coupling studied by Jerrum in his $k > 2\Delta$ result with increasingly sophisticated analyses. Whereas Jerrum's coupling was a one-step coupling, Hayes and Vigoda's coupling pairs $T = O(n)$ steps of one chain with T steps of the other chain.

Notice that in Molloy's improvement to 1.489, he is accounting for $w \in N(v)$ where some $x \in N(w) \setminus \{v\}$ is colored Red and some $y \in N(w) \setminus \{v\}$ is colored Blue. In this case, the attempted update of w to Blue in X and to Red in Y, fails in both chains, and the distance stays the same. Hence, w is called a "doubly-blocked" neighbor. A "singly blocked" neighbor has Blue in its neighborhood or Red, but not both. In this case, the attempted update succeeds in one of the chains, and fails in the other, but the distance still increases by one.

Hayes and Vigoda construct a coupling so that attempted updates always succeed in both chains, or in neither chain. In particular, they couple singly blocked scenarios, so that if X is singly blocked for Blue (or Red respectively), then they couple the evolution of Y so that Y is singly blocked for Red (Blue). The improvement comes from the case when X is singly blocked for Blue, and now Y is singly blocked for Red, and then the attempted update for w fails in both chains, keeping the distance the same. In the earlier coupling, such an update in a singly blocked scenario would have increased the distance.

Coupling a Red singly blocked scenario for X with a Blue singly blocked scenario for Y implies that the neighborhood of w (specifically $N^*(w) = N(w) \setminus \{v\}$) differs between the two chains. Coloring X has Red, but not Blue in $N^*(w)$, whereas Y has Blue, but not Red in $N^*(w)$. Thus, they need to introduce disagreements on $N^*(w)$ to achieve this.

By running X for $T = O(n)$ steps before defining the coupled evolution for Y, they can introduce temporary disagreements on $N^*(w)$ so that these disagreements are guaranteed to disappear before they propagate. Suppose there is an attempted update in X of w to Blue, and w is singly blocked for Blue (thus, Blue appears in $N^*(w)$, but Red does not). Say $x \in N^*(w)$ is the only Red "blocker" of w in chain X.

Roughly speaking, we then find a $y \in N^*(w)$ whose current color can be

replaced by Red, without any effect on the rest of the coloring. In particular, suppose y is currently Yellow. When y was recolored to Yellow, if we instead colored it to Red, and this change would have "no effect" on the rest of the coloring, then y is a suitable "blocker" for Red in chain Y. More specifically, if chain X had recolored y to Red instead of Yellow, and kept all other attempted updates of the dynamics the same, then this change had no effect if the coloring at every other vertex did not change. Hence, at the last recoloring of y, whereas chain X will still use color Yellow, we will couple this with color Red in chain Y. Meanwhile, for the last recoloring of x, while X will still use color Blue, Y will now choose a suitable other color that will not effect the rest of the coloring. In this way, we have caused Y to be singly blocked for Red, while X is still singly blocked for Blue.

The difficulty is doing the above process in such a way that the coupling is still valid. This involves an inductive argument which relies on the above process, for creating temporary disagreements to handle another singly blocked scenario, is reversible (or invertible) in an appropriate sense.

8 Improved Markov Chain

In this section we present the Markov chain used by Vigoda [37] to prove his $11\Delta/6$ result. The chain 'flips' 2-colored clusters. The probability of flipping a cluster is a function of the size of the cluster, and only clusters with at most 6 vertices are flipped. The choice of probabilities for flipping clusters is essential in the details of the coupling analysis.

We need some notation before specifying the transitions of the chain. For a coloring σ, we will refer to a path $v = x_0, x_1, \ldots, x_l = w$ as an alternating path between vertices v and w using colors c and $\sigma(v)$ if, for all i, $(x_i, x_{i+1}) \in E$, $\sigma(x_i) \in \{c, \sigma(v)\}$, and $\sigma(x_i) \neq \sigma(x_{i+1})$. We let $S_\sigma(v, c)$ denote the following cluster of vertices.

$$S_\sigma(v,c) = \left\{ w \;\middle|\; \begin{array}{l} \text{there exists an alternating path between} \\ v \text{ and } w \text{ using colors } c \text{ and } \sigma(v) \end{array} \right\}$$

Let $S_\sigma(v, \sigma(v)) = \emptyset$. For every vertex x in the cluster $S_\sigma(v, c)$, notice that $S_\sigma(x, c) = S_\sigma(v, c)$ if $\sigma(x) = \sigma(v)$ and otherwise $S_\sigma(x, \sigma(v)) = S_\sigma(v, c)$.

For a coloring $\sigma \in \Omega$, the transitions $\sigma \mapsto \sigma'$ are defined as:

- Choose a vertex v and color c uniformly at random from the sets V, C respectively.

- Let $s = |S_\sigma(v, c)|$.
  With probability $\frac{p_s}{s}$, 'flip' cluster $S_\sigma(v, c)$ by interchanging colors c and $\sigma(v)$ on the cluster.

The reason for dividing the flip probability by s is that, as observed above, there are exactly s ways to pick the cluster (one for each of its elements). Thus, a cluster is actually flipped with weight p_s.

To complete the description of the chain, we specify the parameters p_s. They are $p_1 = 1, p_2 = \frac{13}{42}$ and for $s > 2$,

$$p_s = \max\left\{0, \frac{13}{42} - \frac{1}{7}\left[1 + \frac{1}{2} + \cdots + \frac{1}{s-2}\right]\right\}$$

Specifically, $p_3 = \frac{1}{6}, p_4 = \frac{2}{21}, p_5 = \frac{1}{21}, p_6 = \frac{1}{84}$,and $p_s = 0$ for $s \geq 7$.

With this choice of parameters, the chain can then be analyzed using a path coupling analysis for a one step coupling.

9 Other results

Here we briefly a few related results which we did not have space to discuss in detail.

Random Graphs: A random graph with n vertices and dn random edges, $d > 0$ constant has maximum degree $\sim \frac{\log n}{\log \log n}$ with high probability (**whp**), but there are only very few vertices of such large degree. It would therefore seem unnecessary to have k of order Δ in order to generate a random k-coloring. Dyer, Flaxman, Frieze and Vigoda [8] proved that one can construct a Markov chain algorithm that **whp** (over the space of graphs) mixes in $O(n \log n)$ time with k as small as $o(\log \log n)$. It is also shown that **whp** the Glauber dynamics mixes in polynomial time for $k = (\log n)^\alpha$, $\alpha < 1$.

Coupling with Stopping Times: There are several papers that have extended the path coupling approach so that the length of the coupling is a random stopping time. Various techniques and applications are presented in Dyer, Goldber, Greenhill, Jerrum and Mitzenmacher [12], Dyer, Greenhill and Molloy [13], Hayes and Vigoda [21], and Bordewich, Dyer and Karpinski [4].

Lattice graphs: For the 2-dimensional integer lattice (and the torus) there are improved results for $k = 3$ colors, Goldberg, Martin and Paterson

[15], and $k \geq 6$ using computer-assisted proofs, Achlioptas, Molloy, Moore and Van Bussell [1].

Strong Spatial Mixing: For amenable graphs, such as lattice graphs, the notion of strong spatial mixing from Statistical Physics is closely related to $O(n \log n)$ mixing time of the Glauber dynamics. At a high-level, strong spatial mixing says that for uniform random colorings, correlations decay exponentially fast with distance. Goldberg Martin and Paterson [16] obtained interesting results for strong spatial mixing when $k > \alpha\Delta$ and G is triangle-free (without restriction on Δ).

Trees: Martinelli, Sinclair and Weitz [32] studied the mixing rate of the Glauber dynamics on trees. They show $O(n \log n)$ mixing time on complete trees when $k > \Delta + 2$ and Δ is constant.

10 Open Problems

We point out several interesting open questions:

- Combining the proofs of Dyer et al [10] and Hayes and Vigoda [19], can one prove $O(n \log n)$ mixing time of the Glauber dynamics for girth $g \geq 11$ graphs when $k \geq (1 + o(1))\Delta$ or even $k \geq (1 + \varepsilon)\Delta$ for all $\epsilon > 0$ with $\Delta \geq \Delta_0$ where Δ_0 is a constant (that grows with $1/\epsilon$)?
- Can one remove the girth restrictions at least for $k > 1.763\Delta$?
- Can the approach of Vigoda [37] be pushed below $11\Delta/6$?
- For bipartite graphs can we get below Δ? See [31] for a Markov chain that is ergodic on bipartite graphs for all $k \geq 2$ and related negative results on its mixing time.
- Is $k = O(1)$ colors enough for planar graphs?
- Can the results outlined in this survey be extended to *list colorings*?
- Does $k = O(d)$ colors suffice in the case of a random graph of density d/n?

Acknowledgement: We thank Nayantara Bhatnagar for her helpful comments.

References

[1] D. Achlioptas, M. Molloy, C. Moore and F. van Bussell, Sampling grid colorings with fewer colors. In *LATIN 2004: Theoretical Informatics, 6th Latin American Symposium*, 2004, 80-89.

[2] D. J. Aldous. Random walks on finite groups and rapidly mixing Markov chains. In *Séminaire de Probabilities XVII*, 243–297. Springer-Verlag, 1983. Lecture Notes in Mathematics 986.

[3] D. J. Aldous and J. Fill, Reversible Markov Chains and Random Walks on Graphs. Available from: http://stat-www.berkeley.edu/pub/users/aldous/RWG/book.html.

[4] M. Bordewich, M.E. Dyer and M. Karpinski, Using stopping times and counting independent sets and colourings in hypergraphs, Preprint, 2005.

[5] R. Bubley and M.E. Dyer, Path coupling: a technique for proving rapid mixing in Markov chains, *Proceedings of the 38th Annual IEEE Symposium on Foundations of Computer Science* (FOCS), 223–231, 1997.

[6] W. Doeblin, Exposé de la théorie des chaînes simples constantes de Markov à un nombre fini d'états, *Revue Mathématique de l'Union Interbalkanique*, 2:77–105, 1938.

[7] P. Donnelly and D.J.A. Welsh, The antivoter problem: Random 2-colorings of graphs, in *Graph Teory and Combinatorics, B. Bollobás Ed.*, Cambridge University Press (1983) 133-144.

[8] M. Dyer, A. Flaxman, A. Frieze, and E. Vigoda. Randomly coloring sparse random graphs with fewer colors than the maximum degree, to appear in *Random Structures and Algorithms*.

[9] M. Dyer, A. Frieze. Randomly colouring graphs with lower bounds on girth and maximum degree. *Random Structures and Algorithms*, 23 (2):167-179, 2003.

[10] M. Dyer, A. Frieze, T. Hayes and E. Vigoda. Randomly coloring constant degree graphs, in *Proceedings of the 45th Annual Symposium on Foundations of Computer Science* (FOCS), 582-589, 2004.

[11] M.E. Dyer, A.M. Frieze and R. Kannan, A random polynomial time algorithm for approximating the volume of convex bodies. *Journal of the ACM*, 38: 1991, 1-17.

[12] M.E. Dyer, L. Goldberg, C. Greenhill, M.R. Jerrum and M. Mitzenmacher, An extension of path coupling and its application to the Glauber dynamics for graph colourings. In *Proceedings of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms* (SODA), 616–624, 2000.

[13] M.E. Dyer, C. Greenhill and M. Molloy, Very rapid mixing of the Glauber dynamics for proper colourings on bounded-degree graphs, *Random Structures and Algorithms*, 20 (2002), 98–114.

[14] A.M. Frieze and J. Vera, On randomly colouring locally sparse graphs, Preprint, 2004.

[15] L. A. Goldberg, R. Martin and M. Paterson. Random sampling of 3-colourings in Z^2, *Random Structures and Algorithms*, 24(3):2004, 279-302.

[16] L. A. Goldberg, R. Martin and M. Paterson. Strong Spatial Mixing for Graphs with Fewer Colours, in *Proceedings of the 45th Annual Symposium on Foundations of Computer Science* (FOCS), 562-571, 2004.

[17] T. P. Hayes. Randomly coloring graphs of girth at least five. Preprint. A preliminary version appears in *Proceedings of the 35th Annual ACM Symposium on Theory of Computing* (STOC), 269–278, 2003.

[18] T. P. Hayes and A. Sinclair, A general lower bound for mixing of single site dynamics on graphs, in *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science* (FOCS), 2005.

[19] T. P. Hayes and E. Vigoda. A Non-Markovian Coupling for Randomly Sampling Colorings. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science* (FOCS), 618–627, 2003.

[20] T.P. Hayes and E. Vigoda, Coupling with the Stationary Distribution and Improved Sampling for Colorings and Independent Sets. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (SODA), 2005, 971-979.

[21] T. P. Hayes and E. Vigoda. Variable length path coupling. In *Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms* (SODA), 96–103, 2004.

[22] W. Hoeffding, Inequalities for sums of bounded random variables, *Journal of the American Statistical Association* 27, (1963) 13-30.

[23] M.R. Jerrum, A very simple algorithm for estimating the number of k-colourings of a low-degree graph, *Random Structures and Algorithms* 7(2):157–165, 1995.

[24] M.R. Jerrum, *Counting, sampling and integrating: algorithms and complexity*, Lectures in Mathematics–ETH Zürich, Birkhäuser, 2003.

[25] M.R. Jerrrum, L.G. Valiant and V.V. Vazirani, Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43: 1986, 169–188.

[26] M.R. Jerrum, A. Sinclair and E. Vigoda, A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries, *Journal of the ACM*, 51: 2004, 671–697.

[27] J. Jonasson, Uniqueness of uniform random colorings of regular trees. *Statistics & Probability Letters*, 57: 2002, 243-248.

[28] R. Kannan, L. Lovász and M. Simonovits, Random walks and an $O^*(n^5)$ volume algorithm for convex bodies, *Random Structures and Algorithms*, 11: 1987, 1–50.

[29] R.M. Karp, M. Luby and N. Madras, Monte-Carlo Approximation Algorithms for Enumeration Problems. *Journal of Algorithms*, 10: 1989, 429-448.

[30] L. Lovász and S. Vempala, Simulated Annealing in Convex Bodies and an $0^*(n^4)$ Volume Algorithm. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science* (FOCS), 2003, 650-659.

[31] T. Łuczak and E. Vigoda. Torpid mixing of the WangSwendsenKotecký algorithm for sampling colorings. *Journal of Discrete Algorithms*, 3(1):2005, 92-100.

[32] F. Martinelli, A. Sinclair, and D. Weitz, Fast mixing for independent sets, colorings and other emodels on trees, in *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (SODA), 2004, 456-465.

[33] M. Molloy, The Glauber dynamics on colorings of a graph with high girth and maximum degree. *SIAM Journal on Computing*, 33(3):712-737, 2004.

[34] A.D. Petford and D.J.A. Welsh, A randomised 3-colouring algorithm, *Discrete Mathematics* 74 (1989) 253-261.

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

[36] L.G. Valiant, The Complexity of Enumeration and Reliability Problems. *SIAM Journal on Computing*, 8: 1979, 410-421.

[37] E. Vigoda, Improved bounds for sampling colorings, *Journal of Mathematical Physics* 41(3): 1555-1569, 2000.

[38] D.J.A. Welsh, *Complexity: Knots, Colourings and Countings*, London Mathematical Society Lecture Note Series, 1993.