Vigoda's Improvement On Sampling Colorings

Shahrzad Haddadan

Winter 2011

Abstract

- In 1995, Jerrum proved a rapid mixing property on a certain Markov chain to solve the problem of sampling uniformly at random from the set of all proper k-colorings on a graph with maximum degree Δ. Jerrum's proof was for the colorings for which $k > 2\Delta$. In 1999, Vigoda improved this bound to $k > \frac{11}{6}\Delta$. There has been no improvement for this problem after that for general graphs. Here, we introduce the problem and Jerrum's solution very briefly and then we will show Vigoda's approach to the problem.

1 Introduction

Consider a graph $G = (V, E)$ and a labeling σ of the vertices with colors from the set $C = \{1, \ldots, k\}$. We call such a labeling, a *k-coloring* on graph G. If neighboring vertices have different colors then the coloring is called a *proper k-coloring*. The problem we are here interested in, is to sample uniformly from the set of all proper k-colorings on a graph with maximum degree Δ.

This problem, other than being interesting from the combinatorial point of view has some applications in statistical physics and as a result there has been some work by statistical physicists on the subject[6].

We can use Markov chains to solve the sampling coloring problem. Consider a rapidly mixing Markov chain which has a state for each proper coloring of the graph so that the stationary distribution is uniform over all states. We can run that Markov chain until it is closed enough to the stationary distribution and then sample from it. This gives us a polynomial randomized approximation scheme for solving the problem.

The simplest Markov chain one can think of is called *Glauber dynamics.* The Markov chain is defined as,

- Let the state space be all the proper k-colorings on a graph $G = (V, E)$.
- At each step move from a coloring σ to a coloring σ' like this,
  1. Choose a vertex $v \in V$ and color $c \in C$ uniformly at random.
  2. Let $\sigma'(x) = \sigma(x)$ for all $x \in V, x \neq v$.
  3. If setting $\sigma'(v) = c$ results in a proper coloring then $\sigma'(v) = c$, else $\sigma'(v) = \sigma(v)$.

Jerrum proved the following theorem, for mixing time of Glauber dynamics.

Theorem. *(Jerrum[2]) Let G be a graph of maximum degree Δ and n vertices. Assuming $k > 2\Delta$, the mixing time τ of the Glauber dynamics for G and k is bounded above by*

$$\tau_x(\epsilon) \leq \frac{k}{k-2\Delta} n \ln(\frac{n}{\epsilon})$$

regardless of the initial state.

The condition $k > 2\Delta$ in Jerrum's theorem was improved by Bubley, Greenhill and Dyer[3] for Glauber dynamics of 5-colorings when Δ is at most three and of 7-colorings on triangle free four-regular graphs. Their proof relies on computer aid for solving linear programs.

Vigoda, in his paper[1], introduces a new Markov chain which he calls *flip dynamics*. He proves rapid mixing property for this Markov chain where the number of colors k, and the maximum degree Δ, of the underlying graph satisfy $\Delta > \frac{11}{6}$. His paper is remarkable since it has the advantages of improving the 2Δ barrier for arbitrary general graphs and also it is not computer assisted. He also uses flip dynamics rapidly mixing property to show rapidly mixing property for Glauber dynamics for k-colorings on the graphs with maximum degree Δ when $k > \frac{11}{6}\Delta$.

In the following sections, we briefly go through Vigoda's paper.

2 Vigoda's Bound

As we said previously, Vigoda introduced a new Markov chain, named flip dynamics. For this Markov chain he proves the following theorem.

Theorem. *(Vigoda[1]) Let G be a graph of maximum degree Δ on n vertices. Assuming $k > \frac{11}{6}\Delta$, the mixing time τ of the flip dynamics for G and k is polynomial in n[1].*

In order to prove this theorem Vigoda uses path coupling.

In this section we first define Vigoda's chain. Then, we show how he does the coupling which gives him the desired polynomial mixing time. Then, we state the theorem for Glauber dynamics.

2.1 Vigoda's markov chain

In order to define the flip dynamics we first need some notations.[2]

- For a coloring σ, we will refer to a path $v = x_0, x_1, ..., x_l = w$ 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.

$$S_\sigma(v,c) = \left\{ w| \begin{array}{c} \text{there exists an alternating path between} \\ v \text{ and } w \text{ using colors c and } \sigma(v) \end{array} \right\} \tag{1}$$

Let $S_\sigma(v, \sigma(v)) = \emptyset$.

[1]In [1], the theorem is stated more precisely. i.e. the mixing time is bounded by $O(nk\log(n))$
[2]All the definitions in this section are from Vigoda's paper[1]

- In *flip dynamics* the transitions between states of the Markov chain are defined as follows:
  - Choose a vertex $v \in V$ and color $c \in C$ uniformly at random.
  - Let $\alpha = |S_\sigma(v,c)|$. With probability $\frac{p_\alpha}{\alpha}$, 'flip' cluster $S_\sigma(v,c)$ by interchanging colors c and $\sigma(v)$ on the cluster. The parameters p_α are chosen to satisfy certain conditions that will be helpful in proving the rapidly mixing property.[3]

  In fact, we wish to flip each cluster with probability p_α. We divide p_α by factor α, since there are α ways to pick the cluster. This is because, for every vertex x in the cluster $S_\sigma(v,c)$, we have, $S_\sigma(x,c) = S_\sigma(v,c)$ or $S_\sigma(x,\sigma(v)) = S_\sigma(v,c)$.

  Suppose for now, that the state space of the Markov chain is the set of all proper colorings.[4]

All the definitions above give us our Markov chain. It is not very hard to check aperiodicity and irreducibility of this Markov chain. As a result, we can conclude that this Markov chain has an stationary distribution. Moreover, the chain is symmetric and therefore, it has a uniform stationary distribution.

2.2 Vigoda's Coupling

Before going through Vigoda's path coupling, we first state the *path coupling theorem.*

Theorem. *(Path Coupling[5])*
Let Φ be an integer-valued metric defined on $\Omega \times \Omega$ such that, for all $\sigma, \xi \in \Omega$, there exists a path $\eta \in \rho(\sigma, \xi)$ with

$$\Phi(\sigma,\xi) = \sum_i \Phi(\eta_i, \eta_{i+1}). \tag{2}$$

Suppose there exists a constant $\beta < 1$ and a coupling (σ_t, ξ_t) of the Markov chain such that, for all $\sigma_t \sim \tau_t$,

$$\mathbf{E}[\Phi(\sigma_{t+1}, \tau_{t+1})] \leq \beta \Phi(\sigma_t, \tau_t).$$

Then the chain is rapidly mixing.

To use the path coupling, we let Φ to be the Hamming distance which is the number of vertices that are colored differently in the two states. If Ω is equal to the set of all proper colorings, then equation(2) in path coupling theorem can not be satisfied. Consider a pair of adjacent vertices v and w. Moreover, suppose that $\sigma(v) = \eta(w), \sigma(w) = \eta(v)$. Thus, $\Phi(\sigma,\tau) = 2$ but the shortest path in Ω between these states is of length three. To overcome this problem, we extend Ω to the set of all colorings-proper or improper. It is not very hard to see that the improper colorings are transient states in the Markov chain and therefore they do not appear in the stationary distribution and as a result the stationary distribution is still the uniform distribution over all proper colorings.

Now consider two neighboring states σ and τ such that they differ only in vertex v. We aim to introduce a coupling for σ, τ so that the expected change of $\Phi(\sigma,\tau)$ under this coupling

[3] For more detail see [1] page. 6.

[4] For a reason that we will explain in the next subsection, we will extend the state space of this Markov chain to the set of all colorings of a graph- proper and improper colorings.

is negative. It is not very hard to see that the set D of clusters which might be different in σ and τ are

- $S_\sigma(w, \tau(v)), S_\tau(w, \sigma(v))$ for any neighbor w of v,
- $S_\sigma(v, c), S_\tau(v, c)$ for any color c.

We partition D to sets D_c. Each set D_c consists of colors c, $\sigma(v)$ and $\tau(v)$. Let,

$\Gamma_c = \{w | \sigma(w) = c, w \text{ is a neighbor of } v\}$,

$D_c = \{S_\sigma(v, c), S_\tau(v, c)\} \cup \{S_\sigma(w, \tau(v)), S_\tau(w, \sigma(v))\}_{w \in \Gamma_c}$

We should note here that for $c = \sigma(v), \tau(v)$ the sets may intersect. That issue is not very important and is handled as a special case in Vigoda's paper.

The idea for coupling is to couple the clusters in each partition together. Suppose f is the function that specifies the coupling such that if we choose a move in σ that flips cluster S, then f defines the coupled move in τ that moves cluster $f(S)$ in τ.

- For $S \notin D, f(S) = S$.
- For $S \in D_c, f(S) \in D_c$. We try to set $f(S)$ so that size of $f(S)$ is close to size of S.

In order to find a good way to couple the clusters in D_c, consider the following equations that can be easily proved. Let $\Gamma_c = \{w_1, ..., w_{\delta_c}\}$,

For $c \neq \sigma(v)$, $\quad S_\sigma(v, c) = \{\cup_i S_\tau(w_i, \sigma(v))\} \cup \{v\}$.

For $c \neq \tau(v)$, $\quad S_\tau(v, c) = \{\cup_i S_\sigma(w_i, \tau(v))\} \cup \{v\}$.

Let $a_i = a_i(c) = |S_\tau(w_i, \sigma(v))|$, $A = A(c) = |S_\sigma(v, c)| \leq 1 + \sum_i a_i$. Similarly, let $b_j = b_j(c) = |S_\sigma(w_j, \tau(v))|, B = B(c) = |S_\tau(v, c)| \leq 1 + \sum_j b_j$. Let $a_{\max} = \max_i a_i$ and $i_{\max}$ is the corresponding index for $a_{\max}$ (similarly for $b_{\max}$ and $j_{\max}$). Now, we do the coupling.[5]

- We couple the big flips, $S_\sigma(v, c)$ and $S_\tau(v, c)$, with the largest of the other flips, $S_\tau(w_{i_{\max}}, \sigma(v)), S_\sigma(w_{j_{\max}}, \tau(v))$.

Since each cluster should be flipped by probability p_α where α is its size, we need to perform some other coupling on the same clusters so that the weight by which each cluster is flipped is equal to p_α.

- For each w_i, couple together (as much as possible) the remaining weights of flips, $S_\sigma(w_i, \tau(v)), S_\tau(w_i, \sigma(v))$.

Having this coupling, Vigoda proves $\mathbf{E}[\Delta\Phi] \leq \frac{-k + \frac{11}{6}\Delta}{nk}$. Therefore, using path coupling theorem, we have the mixing property for $k > \frac{11}{6}\Delta$.

2.3 Comparison with Galuber dynamics

By having a bound on mixing time of flip dynamics, Vigoda uses the comparison theorem of Diaconis and Saloff-Coste [4] to prove the same bound for Glauber dynamics.

Theorem. *The Glauber dynamics is rapidly mixing, with mixing time $O(n^2 k \log n \log k)$, provided $k > \frac{11}{6}\Delta$.*[6]

[5] Here we just give the idea of Vigoda's coupling for my detail you can see [1], Page. 9.

[6] [1]

3 Conclusion and further work

For the Glauber dynamics, it seems that we can not improve the coupling better than what Jerrum did. The main idea of Vigoda's paper was changing the chain and flipping clusters instead of vertices.

After Vigoda's paper, there have been other efforts to improve the bound. Although there has been some improvement on some special graphs[7], there has not been any improvement for bound on general graphs.

There is a paper by Dyer et all [7] which proves that, for every $\epsilon > 0$, the Glauber dynamics converges to a random coloring within $O(n \log n)$ steps assuming $k = k_0(\epsilon)$ and either: (i) $\frac{k}{\Delta} > \alpha + \epsilon$ where $\alpha \approx 1.763$ and the girth $g = 5$, or (ii) $\frac{k}{\Delta} > \beta + \epsilon$ where $\beta \approx 1.489$ and the girth $g = 7$.

In another paper, Hayes and Vigoda [8] 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 $g > 11$.

We end the discussion by pointing out several open questions:

- Combining the proofs of Dyer et al [7] and Hayes and Vigoda [8], can one prove $O(n \log n)$ mixing time of the Glauber dynamics for girth $g > 11$ graphs when $k > (1 + o(1))\Delta$ or even $k > (1+\epsilon)\Delta$ for all $\epsilon > 0$ with $\Delta > \Delta_0$ 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 [1] be pushed below $\frac{11}{6}\Delta$?

References

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

[2] 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):157165, 1995.

[3] R. Bubley and M. Dyer and C. Greenhill. Beating the 2Δ bound for approximating colorings: A computer assisted proof of rapidly mixing. In Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, pages 335-363, San Fransisco, CA, January 1998.

[4] P.Diaconis and L.Saloff-Coste , Comparison theorems for reversible Markov chains. The Annals of applied probability, 3(3):696-730, 1993.

[5] R.Bubley and M. Dyer. Path coupling, Dobrushin uniqueness, and approximate countings. In 38th Annual Symposium on Foundation of Computer Science, pages 223-231, Miami Beach, FL, October 1997. IEEE.

[6] R.Potts. Some generalized order-disorder transformations. Proseeding of the Cambridge Philosophical Society, 48:106-109, 1952.

[7] For a survey on the related results see [9].

[7] 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.

[8] 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), 618627, 2003.

[9] Alan Frieze and Eric Vigoda, A survey on the use of Markov chains to randomly sample colourings. In Combinatorics, complexity, and chance. Oxford Lecture Ser. Math. Appl., vol. 34. Oxford Univ. Press, Oxford, UK, 5371.