A New Approach Towards Graph Coloring

Anand Louis

June 1, 2012

Mentors : Santosh Vempala, Prasad Tetali

1 Introduction

Consider the problem of deciding whether a graph is 3-colorable. This problem is NP-complete. The optimization version of this problem is, given a 3-colorable graph G, color it using as few colors as possible. In the regime of polynomial time approximation algorithms, this problem has been studied extensively [Wig82, Blu94, KMS98, ACC06, Chl07]. The current best known approximation factor is achieved by an algorithm due to Kawarabayashi and Thorup [iKT12] which uses at most $n^{0.2038}$ colors to color a 3-colorable graph. Surprisingly, very few lower bounds are known for the approximability of the chromatic number of a graph. Dinur, Mossel and Regev [DMR06] showed that it's hard to color a 3-colorable graph with $O(1)$ colors based on a variant of Unique Games Conjecture. Feige, Langberg and Schechtman [FLS02] showed that the simple SDP relaxation has an integrality gap of at least $n^{0.157}$.

In a recent work, Barak, Raghavendra and Steurer [BRS11] and Guruswami and Sinop [GS11] showed a new way to round vector solutions of semidefinite programming (SDP) hierarchies into integral solutions, based on a connection between these hierarchies and the spectrum of the input graph. In another recent work [BBH+12], Barak et. al. gave a natural polynomial time algorithm based on the 'Sum of Squares' SDP hierarchy that solves the previously proposed hard instances for Unique Games [RS09, KS09, KPS10, BGH+11]. Along these line, we propose to study a new approach towards sub-exponential time algorithms with improved approximation guarantees for combinatorial problems like graph coloring, vertex cover, etc.

Barak et. al.'s algorithm [BRS11] can be viewed as a divide-and-conquer approach. Consider an instance $\mathcal{I}$ of *MAX* 2-CSP. The constraint graph G of $\mathcal{I}$ is first partitioned into sets with small ε-threshold rank [1] such that only $O(\varepsilon)$ fraction of the edges are cut. This step is same as the preprocession step used in Arora, Barak and Steurer's sub-exponential time algorithm for unique games [ABS10]. Barak et. al. show that graphs with low threshold rank are 'easy'. More formally, they show that $O(\mathsf{rank}_{1-\varepsilon}(G)$ levels of the Lasserre hierarchy is sufficient to recover an integral solution from the SDP's vector solution whose cost[2] is at most $O(\varepsilon)$ lesser than that of the SDP solution.

2 Our Approach

The graph coloring problem can also be viewed as a constraint satisfaction problem. The problem asks to assign labels to every vertex such that every pair of adjacent vertices has a different label. If the graph has low threshold rank, then we can use the [BRS11] approach to find a coloring using a small number of colors. However, if the threshold rank is high, then partitioning the graph into pieces with low threshold rank will not be helpful as there doesn't seem to be any coherent way of ensuring that adjacent vertices which are separated by the partitioning get different colors. One difference between graph coloring CSP and the usual *MAX* − CSP regime is that in the graph coloring CSP the constraints are hard, i.e. every pair of adjacent vertices must have a different labels. Though this might seem as an

[1] ε-threshold rank of a graph G, denoted by $\mathsf{rank}_{1-\varepsilon}(G)$, is defined as the number of eigenvalues of the normalized adjacency matrix of G which are larger than $1 - \varepsilon$.

[2] The *cost* of a solution is the fraction of constraints satisfied by that assignment.

additional constraint, this is actually usefull for our approach as it allows us to re-weigh the edges of the graph (without changing the weighted degrees of the vertices) in an attempt to improve the threshold rank of the graph.

Closely related to threshold rank is the notion of collision probability. The t-step collision probability of a random walk is the probability that two independent random walks of t steps starting from the same vertex [3] also finish at the same vertex. Let $CP_t(G)$ denote the t-step collision probability of the random walk on the graph G. It is easy to see that $CP_t(G) \leqslant \mathsf{rank}_{1-\varepsilon}(G)(1-\varepsilon)^{2t}$. Computing the re-weighting of the graph which has the least t-step collision probability is captured by a simple convex program and hence is tractable. We conjecture that if the graph has high collision probability under every re-weighting, then the graph must have a small set with small vertex expansion.

Conjecture 2.1. *Let G be a on n vertices such that $\mathsf{rank}_{1-\varepsilon}(G') \geqslant n^{\Theta(\eta/\gamma)}$ for every degree preserving re-weighting G' of G. Then there exists a vertex set S of size at most $n^{1-\eta/\gamma}$ that satisfies $\phi^V(S) \leqslant \gamma^c$ for some absolute constant c.*

If this conjecture is true, it would yield a sub-exponential time algorithm for coloring a 3-colorable graph using $\mathsf{polylog}\, n$ colors as follows. We fix γ to be some small constant, say 0.01. If the graph has high threshold rank under every re-weighting, then by repeatedly applying the statement of this conjecture, we get a set S of size at most $n/2$ vertices such that every connected component in $G \backslash S$ has low threshold rank. Using the approach of [BRS11], we can color the vertices in $G \backslash S$ with a small number of colors. We recurse on the vertices in S, but we make use of a different set of colors for each recursion. Since the number of uncolored vertices after each round decreases by a factor of 2, the algorithm terminates after at most $\log n$ rounds of recursion and having used $O(\log n)$ colors in all.

3 Global Correlation in Vector Solutions

As an attack on Conjecture 2.1, we propose a new random process on graphs, and show that if this process has high collision probability, then the graph must have a small set with small vertex expansion. We conjecture that if the graph has many eigenvalues close to 1 under every re-weighting of its edges, then this process must have a large collision probability for most starting distributions.

The process starts with a probability mass of 1 at vertex some vertex, say i, i.e., $\mu_{i,0}(j) = 1$ if $j = i$ and 0 otherwise. At time t, for each vertex j, the probability mass flows towards the vertex k with the maximum difference in probability with rate proportional to the difference. More formally, let $a_j \stackrel{\text{def}}{=} \mathsf{argmax}_{a\sim j}\mu_{i,t}(a) - \mu_{i,t}(j)$, then the probability mass flows from j to a_j if $\mu_{i,t}(a_j) - \mu_{i,t}(j) > 0$, and at a rate proportional to $\mu_{i,t}(a_j) - \mu_{i,t}(j)$. We show that the collision probability of this process decreases exponentially with time.

Lemma 3.1. *For the process outlined above, the collision probability is bounded by* [4]

$$CP_t \leqslant e^{-\lambda_\infty t}$$

For our purpose, it would be interesting to view this process in connection with the SDP vector solution $\{v_i | i \in [n]\}$. We define the local correlation to be $\mathbb{E}_{j,k\sim\mu_{i,t}}\langle v_j, v_k\rangle^2$. We conjecture that this process has high local correlation.

Conjecture 3.2. *For any set of vectors $\{v_i | i \in [n]\}$ which form a feasible solution to the SDP relaxation of Graph 3-coloring*

$$\mathop{\mathbb{E}}_{j,k\sim\mu_{i,t}}\langle v_j, v_k\rangle^2 \geqslant \left(\mathop{\mathbb{E}}_{j\sim k}\langle v_j, v_k\rangle^2\right)^t$$

We define the global correlation to be $\mathbb{E}_{i,j}\langle v_i, v_j\rangle^2$. [BRS11] show that if the SDP solution has high global correlation, then rounding it to an integral solution is 'easy'. It can be shown that if Conjecture 3.2 is true, then the vector solution either has high global correlation, or high collision probability. In the former case we can use the approach of [BRS11] to make progress. This case is akin to the case when the graph has low threshold rank. In the latter case, using Lemma 3.1 and some additional machinery, we can show that graph has a small set with small vertex expansion. Infact, Conjectures 2.1 and 3.2 can be shown to be equivalent.

[3] The starting vertex is chosen according to the stationary distribution of the random walk.

[4] λ_∞ of a graph G is defined as $\lambda_\infty \stackrel{\text{def}}{=} \min_{x\in\mathbb{R}^n} \frac{\sum_i \max_{j\sim i}(x_i - x_j)^2}{\sum_i x_i^2 - (\sum_i x_i)^2/n}$

References

[ABS10] Sanjeev Arora, Boaz Barak, and David Steurer, *Subexponential algorithms for unique games and related problems*, FOCS, 2010. 1

[ACC06] Sanjeev Arora, Moses Charikar, and Eden Chlamtac, *New approximation guarantee for chromatic number*, in Kleinberg [Kle06], pp. 215–224. 1

[BBH+12] Boaz Barak, Fernando G. S. L. Brandão, Aram Wettroth Harrow, Jonathan A. Kelner, David Steurer, and Yuan Zhou, *Hypercontractivity, sum-of-squares proofs, and their applications*, STOC (Howard J. Karloff and Toniann Pitassi, eds.), ACM, 2012, pp. 307–326. 1

[BGH+11] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer, *Making the long code shorter, with applications to the unique games conjecture*, Electronic Colloquium on Computational Complexity (ECCC) **18** (2011), 142. 1

[Blu94] Avrim Blum, *New approximation algorithms for graph coloring*, J. ACM **41** (1994), no. 3, 470–516. 1

[BRS11] Boaz Barak, Prasad Raghavendra, and David Steurer, *Rounding semidefinite programming hierarchies via global correlation*, FOCS, 2011, pp. 472–481. 1, 2

[Chl07] Eden Chlamtac, *Approximation algorithms using hierarchies of semidefinite programming relaxations*, FOCS, IEEE Computer Society, 2007, pp. 691–701. 1

[DMR06] Irit Dinur, Elchanan Mossel, and Oded Regev, *Conditional hardness for approximate coloring*, in Kleinberg [Kle06], pp. 344–353. 1

[FLS02] Uriel Feige, Michael Langberg, and Gideon Schechtman, *Graphs with tiny vector chromatic numbers and huge chromatic numbers*, FOCS, IEEE Computer Society, 2002, pp. 283–292. 1

[GS11] Venkatesan Guruswami and Ali Kemal Sinop, *Lasserre hierarchy, higher eigenvalues, and approximation schemes for graph partitioning and quadratic integer programming with psd objectives*, FOCS, 2011, pp. 482–491. 1

[iKT12] Ken ichi Kawarabayashi and Mikkel Thorup, *Combinatorial coloring of 3-colorable graphs*, CoRR **abs/1205.1254** (2012). 1

[Kle06] Jon M. Kleinberg (ed.), *Proceedings of the 38th annual acm symposium on theory of computing, seattle, wa, usa, may 21-23, 2006*, ACM, 2006. 3

[KMS98] David R. Karger, Rajeev Motwani, and Madhu Sudan, *Approximate graph coloring by semidefinite programming*, J. ACM **45** (1998), no. 2, 246–265. 1

[KPS10] Subhash Khot, Preyas Popat, and Rishi Saket, *Approximate lasserre integrality gap for unique games*, APPROX-RANDOM (Maria J. Serna, Ronen Shaltiel, Klaus Jansen, and José D. P. Rolim, eds.), Lecture Notes in Computer Science, vol. 6302, Springer, 2010, pp. 298–311. 1

[KS09] Subhash Khot and Rishi Saket, *Sdp integrality gaps with local ℓ_1-embeddability*, FOCS, IEEE Computer Society, 2009, pp. 565–574. 1

[RS09] Prasad Raghavendra and David Steurer, *Integrality gaps for strong sdp relaxations of unique games*, FOCS, 2009, pp. 575–585. 1

[Wig82] Avi Wigderson, *A new approximate graph coloring algorithm*, STOC (Harry R. Lewis, Barbara B. Simons, Walter A. Burkhard, and Lawrence H. Landweber, eds.), ACM, 1982, pp. 325–329. 1