New Tools for Graph Coloring*

Sanjeev Arora and Rong Ge

Department of Computer Science, Princeton University
and Center for Computational Intractability
arora/rongge@cs.princeton.edu

Abstract. How to color 3 colorable graphs with few colors is a problem of longstanding interest. The best polynomial-time algorithm uses $n^{0.2072}$ colors. There are no indications that coloring using say $O(\log n)$ colors is hard. It has been suggested that SDP hierarchies could be used to design algorithms that use n^{ϵ} colors for arbitrarily small $\epsilon > 0$.

We explore this possibility in this paper and find some cause for optimism. While the case of general graphs is till open, we can analyse the Lasserre relaxation for two interesting families of graphs.

For graphs with low *threshold rank* (a class of graphs identified in the recent paper of Arora, Barak and Steurer on the unique games problem), Lasserre relaxations can be used to find an independent set of size $\Omega(n)$ (i.e., progress towards a coloring with $O(\log n)$ colors) in $n^{O(D)}$ time, where D is the threshold rank of the graph. This algorithm is inspired by recent work of Barak, Raghavendra, and Steurer on using Lasserre Hierarchy for unique games. The algorithm can also be used to show that known integrality gap instances for SDP relaxations like *strict vector chromatic number* cannot survive a few rounds of Lasserre lifting, which also seems reason for optimism.

For *distance transitive* graphs of diameter Δ, we can show how to color them using $O(\log n)$ colors in $n^{2^{O(\Delta)}}$ time. This family is interesting because the family of graphs of diameter $O(1/\epsilon)$ is easily seen to be *complete* for coloring with n^{ϵ} colors. The distance-transitive property implies that the graph "looks" the same in all neighborhoods.

The full version of this paper can be found at:
http://www.cs.princeton.edu/~rongge/LasserreColoring.pdf .

1 Introduction

In the graph coloring problem we are given a graph $G = (V, E)$. A *coloring with t colors* is a function $f : V \to [t]$, such that for any $(p, q) \in E$, $f(p) \neq f(q)$. The smallest t such that a coloring exists is called the *chromatic number* of the graph, and the graph is said to be *t-colorable.*

Despite much research we still have no good coloring algorithms even in very restricted cases. This is explained to some extent because it is NP-hard to approximate the chromatic number of a graph up to a factor of $n^{1-\epsilon}$ for any $\epsilon > 0$ ([20], following a long line of work in PCPs). Therefore attention has shifted to the case where the graph is 3-colorable. In this restricted case known algorithms can color the graph using $\tilde{O}(n^c)$[1] colors for some constants c. Wigderson's

* Research supported by NSF Grants CCF-0832797, 0830673, and 0528414.

[1] Here and throughout the paper $\tilde{O}$ hides logarithmic factors.

L.A. Goldberg et al. (Eds.): APPROX/RANDOM 2011, LNCS 6845, pp. 1–12, 2011.

purely combinatorial algorithm [19] works for $c = 1/2$. Using more combinatorial tools Blum achieved $c = 3/8$ [5]. Karger, Motwani, and Sudan [13] used SDP relaxations to achieve $c = 1/4$, which was combined with combinatorial tools to achieve $c = 3/14$ by Blum and Karger [6]. Arora, Charikar and Chlamtac [2] then carefully analyzed the SDP relaxation to reduce c to 0.2111. Chlamtac [8] further reduced c to 0.2072 using $O(1)$ levels of Lasserre lifting of the basic SDP relaxation (Lasserre lifting is defined in Section 2).

The seeming difficulty in getting even small improvements in c suggests that substantial improvement to c (achieving $c = o(1)$ for example) is intractable, but few lowerbounds are known. Dinur, Mossel and Regev[10] showed that it's hard to color with any constant number of colors (i.e., $O(1)$ colors) based on a variant of Unique Games Conjecture. Some integrality gap results[11,13,18] show that the simple SDP relaxation has an integrality gap at least $n^{0.157}$.

Arora et al. [2] suggested that using $O(1)$ or $O(\log n)$ levels of Lasserre lifting on the standard SDP relaxation should allow us to find an n^{ϵ}-coloring (running time would be $n^{O(k)}$ where k is the number of levels). In general researchers have hoped that a few rounds of lift-and-project strengthening of SDP relaxations (via Lasserre or other methods) should allow better algorithms for many other problems, though few successes have resulted in this endeavor.

The current paper is related to recent developments about the unique games problem. A surprising recent result of Arora et.al. [1] showed that unique games can be solved in subexponential time using the idea of *threshold rank*. More recently, Barak, Raghavendra and Steurer [3] showed that the surprising subexponential algorithm for unique games can be rederived using Lasserre lifting. Their rounding algorithm involves a new convex programming relaxation for threshold rank which we also use in a key way. It gives a way to round the SDP solution by showing that the solution vectors exhibit "global correlation."

We extend the techniques of Barak et.al. to show that low threshold rank also helps in coloring 3-colorable graphs with fewer colors. Our algorithm is also derived using Lasserre liftings. In general we think our approach may lead to n^{ϵ}-coloring in subexponential or even quasi-polynomial time.

1.1 Our Results

The difficulty in using Lasserre liftings for colorings as well as any other problem is the lack of an obvious rounding algorithm. The paper [3] gives such a rounding algorithm for the unique games problem for graphs of low *threshold rank*. Our first result is a similar algorithm for graph coloring. We state the theorem here and will prove it in Section 4. The hypothesis uses a somewhat different notion of threshold rank than [3].

Theorem 1. *There is a constant $c > 1$ and a randomized rounding algorithm such that the following is true. If a regular 3-colorable graph G has threshold rank $Rank_{-1/16}(G)$ (i.e., the number of eigenvalues less than $-1/16$, where eigenvalues are scaled to lie in $[-1, 1]$) at most D, then the algorithm can find an independent set of size at least $n/12$ in time $n^{O(D)}$ with constant probability.*

Moreover, if the graph is vertex-transitive, there is a randomized algorithm that finds a coloring with $O(\log n)$ colors in $n^{O(D)}$ time.

As a corollary of the above result we can show that existing "counterexamples" for graph coloring algorithms (eg integrality gap examples [12]) are easyfor high level Lasserre liftings since they all have low threshold rank.

When we try to apply similar ideas to general graphs, we quickly realize that the problematic cases (if they exist at all) must be such that different neighborhoods look "different." Of course, this flies against the usual intuition about SDP relaxations: the usual reason for high integrality gaps (at least in explicit examples) is precisely that all neighborhoods look the same and the SDP gives no meaningful clues.

To quantify the notion of all neighborhoods "looking the same," we focus on a specific kind of symmetric graph, the *distance transitive graphs*, which have been well-studied in graph theory (see the book [7]). In fact we restrict attention to such graphs that in addition have low diameter. The reason is that using simple combinatorial arguments one can show that in order to color the graph with n^ϵ colors, it suffices to restrict attention to graphs of diameter $O(1/\epsilon)$. If a 3-colorable distance transitive graph has diameter Δ we show how to find a $O(\log n)$ coloring in $O(n^{2^{O(\Delta)}})$ time. See Section 5.

How can our ideas be generalized to all graphs? In Section 6 we formulate a conjecture which if true would yield subexponential time coloring algorithms that find an n^ϵ-coloring.

2 The SDP and Lasserre Hierarchy

The standard SDP relaxation for graph 3-coloring uses *vector chromatic number*, but it is not amenable to Lasserre lifting. So we start with an equivalent (see [8]) relaxation based upon 0/1 variables. For each vertex p of the graph $G = (V, E)$, we have three variables $x_{p,R}, x_{p,Y}, x_{p,B}$ where $x_{p,C} = 1$ for $C \in \{R, Y, B\}$ "means" the vertex p is colored with color C. Thus exactly one of the three variables will be 1. The integer program makes sure $x_{p,R} + x_{p,B} + x_{p,Y} = 1$, and $x_{p,C}x_{q,C} = 0$ if p and q are adjacent.

Now we relax the integer program by replacing each $x_{p,C}$ with a vector $v_{p,C}$. The result is an SDP relaxation. Then we lift this SDP using k levels of Lasserre lifting. (For Lasserre lifting see the surveys[9,15]).The lifted SDP contains vector variables v_S, where S is a subset of the set $V \times \{R, Y, B\}$ (later denoted by Ω) and has size at most k. The resulting SDP is

$$\forall p \in V \quad v_{p,R} + v_{p,B} + v_{p,Y} = v_\emptyset \tag{1}$$

$$\forall p \in V, C_1 \neq C_2 \quad \langle v_{p,C_1}, v_{p,C_2} \rangle = 0 \tag{2}$$

$$\forall (p, q) \in E, C \in \{R, Y, B\} \quad \langle v_{p,C}, v_{q,C} \rangle = 0 \tag{3}$$

$$\forall P, Q, S, T \subseteq \Omega, P \bigcup Q = S \bigcup T, |P \bigcup Q| \leq k \quad \langle v_P, v_Q \rangle = \langle v_S, v_T \rangle \tag{4}$$

In this SDP, Equations (1) to (3) are constraints obtained from the integer program; Equations (4) are the consistency constraints imposed by Lasserre lifting; we also require $\langle v_\emptyset, v_\emptyset \rangle = 1$ for normalization. Notice that here we are abusing notation a bit: if the set S contains only one event (p, C), we use both $v_{p,C}$ and v_S for the same vector. We call this SDP Las^k and its solution SDP^k.

2.1 Understanding the SDP Solution

Here we discuss how we should interpret the solution of coloring SDP. Throughout the discussion below, an "atomic event" (abbreviated to just "event" when this causes no confusion) consists of assigning a vertex p some color C. We denote by $\Omega = V \times \{R, Y, B\}$ the set of all atomic events. Our rounding algorithm will iteratively assign colors to vertices. Each step may assign a color C to p, or declare that color C will *never* be assigned to p. In the former case the atomic event (p, C) has happened; in the latter case the complement event happened. It is common to interpret the SDP solution as giving a distribution over these events whose probabilities are equal to the innerproducts of the vectors. We formulate this by the following theorem:

Theorem 2 ([14,8]). *A solution to k-level Lasserre lifting SDP (Las^k) encodes a locally consistent coloring for any set of k vertices. Locally consistent means all colorings with positive probability are valid colorings. If W is a set of atomic events then the probability that they happen is equal to the inner-product of v_S and v_T, where $S \bigcup T = W$. In particular, each vector can be decomposed as $v_W = r_W v_\emptyset + u_W$ where r_W is the probability that all events in W happen and u_W is perpendicular to $v_\emptyset$.*

If $w = (p, C)$ is an atomic event, properties of Lasserre lifting allow us to construct a subsolution in which event w happens (ie vertex p is assigned colored C), and a subsolution in which event w does not happen (ie color C is forbidden for p from now on). We randomly choose one of the subsolutions to preserve the probability of w. That is, if r_w is the probability of event w, we pick the subsolution in which w happens with probability r_w, and pick the subsolution in which w does not happen with probability $1 - r_w$. We call this *"conditioning the solution on event w"*. The result of such an operation will be a solution for $k - 1$ level of Lasserre lifting, which we call SDP^{k-1}.

The computation needed to compute the new vectors in SDP^{k-1} is simple and follows from the above theorem: the probabilities of the new $k-1$-level solution must be the appropriate conditional probabilities in the locally consistent distributions in the k-level solution. For details see the full version or [3,9,8].

Note that we must use Lasserre instead of weaker relaxations: Sherali-Adams [17] and Lovász-Schrijver[16], because we consider solutions as locally consistent solutions (which rules out Lovász-Schrijver) and we use critically that probabilities correspond to inner-products of vectors (which rules out Sherali-Adams). Detailed comparison between the hierarchies are given in the surveys [9,15].

3 Global Correlation, Local Correlation and Rounding

Given a solution to the k-level lifting Las^k of a graph G, we shall define the global correlation of this solution and show how global correlation of order $\Omega(1/k)$ can help to round the solution. Intuitively global correlation measures the average correlation between the colors of two vertices chosen uniformly at random. In general, this correlation may be close to 0: knowing the color of one doesn't give much information about the color of the other. If the global correlation is bounded away from 0 however, then intuitively speaking, fixing the color for a randomly chosen vertex should bias the average remaining vertex a bit towards a particular color. Thus after fixing the colors for a sufficiently large set of vertices, the colors for most of the remaining vertices must get more or less fixed. This the main idea of Barak et.al.[3] in the context of unique games, and Lemma 1 is adapted from there. The amount of variability in the color of the average vertex is quantified using *variance*.

We first examine how conditioning on one atomic event reduces the variance of another. Let w_1, w_2 be two atomic events, r_1, r_2 be their probabilities respectively, and r_{12} be the probability that both of them happen. The variance of the conditional random variable $w_2|w_1$ is given by:

$$\mathrm{Var}[w_2|w_1] = \mathrm{Var}[w_2] - \frac{(r_1 r_2 - r_{12})^2}{\mathrm{Var}[w_1]}. \tag{5}$$

By the equation we see that the expected variance always drops, and the drop is proportional to $(r_1 r_2 - r_{12})^2$. Below we call this quantity the *correlation* between the two events.

Correlation has a geometric meaning in Lasserre solutions. Notice that $r_1 = \langle v_{w_1}, v_\emptyset \rangle$, $r_2 = \langle v_{w_2}, v_\emptyset \rangle$, and $r_{12} = \langle v_{w_1} v_{w_2} \rangle$ (by Theorem 2). As in Theorem 2 we express $v_{w_i} = r_i v_\emptyset + u_{w_i}$, then $\langle u_{w_1}, u_{w_2} \rangle = \langle v_{w_1}, v_{w_2} \rangle - r_1 r_2 = r_{12} - r_1 r_2$. Therefore we have $(r_1 r_2 - r_{12})^2 = \langle u_{w_1}, u_{w_2} \rangle^2$.

Definition 1 (Correlation, Global Correlation, Variance). *Given a solution SDP^k and two events w_1, w_2, The* correlation *between w_1 and w_2 is defined as (where probabilites r and vectors u are as in Theorem 2):*

$$Cor[w_1, w_2] = (r_{w_1} r_{w_2} - r_{\{w_1, w_2\}})^2 = \langle u_{w_1}, u_{w_2} \rangle^2.$$

The global correlation *of a set of vectors $\{z_p\}$ ($p \in U$) is just the expected correlation between two randomly picked vectors: $GC(\{z_p\}) = \mathrm{E}_{p,q \in U} \langle z_p, z_q \rangle^2$.*

The global correlation *of the SDP solution is the global correlation of all the u vectors for the set of atomic events (Ω). Intuitively it is the average correlation between a pair of atomic events.*

$$GC^k = \mathop{\mathrm{E}}_{w_1, w_2 \in \Omega} \langle u_{w_1}, u_{w_2} \rangle^2. \tag{6}$$

The variance *of the solution is $VAR^k = \mathrm{E}_{w \in \Omega}\, r_w (1 - r_w)$.*

Now we are ready to state the following Lemma for one step of rounding.

Lemma 1. *Suppose SDP solution SDP^k has global correlation GC^k and variance VAR^k. Upon picking a random event $w \in \Omega$ and conditioning on that event, the new solution SDP^{k-1} has expected variance at most $VAR^k - 4GC^k$.*

Proof. Due to space limit please see the full version.

Lemma 1 corresponds to a single step in our iterative rounding. So long as the global correlation is substantial —say, at least $10/k$– we can repeat this step up to k times and drive the variance of the solution towards zero. Intuitively, once the variance is small enough, the solution should be almost integral and thus easy to round. Indeed we show the following:

Lemma 2. *Given a vector solution SDP^k ($k \geq 2$) for k-level Lasserre lifting Las^k , we say a vertex p is* determined *if there's a color C such that event $\{p, C\}$ that happens with probability more than $1/2$. Otherwise the vertex is* undetermined. *If SDP^k has variance $VAR^k < 1/8$ then at least $1/4$ of the vertices are determined. Moreover, if we color the determined vertices with the color that makes them determined, then this is a valid partial coloring (i.e., no two adjacent vertices will have the same color).*

Proof. First rewrite the definition of variance as $VAR^k = \mathrm{E}_{w\in\Omega}\, r_w(1 - r_w) = \mathrm{E}_{p\in V}\, \mathrm{E}_{C\in\{R,Y,B\}}\, r_{(p,C)}(1 - r_{(p,C)})$.

From this formula we know for any vertex p and its 3 events w_1, w_2, w_3, their contribution to VAR^k is proportional to $(Var[w_1]+Var[w_2]+Var[w_3])/3$ (the second expectation in the right hand side). For undetermined vertices, the probabilities for w_1, w_2, w_3 can be more than $1/2$ and they sum up to 1, thus the minimum possible value of the contribution of this vertex p is $(1/4 + 1/4 + 0)/3 = 1/6$. If more than $3/4$ of the vertices are undetermined, we would have $VAR^k > 3/4 \cdot 1/6 = 1/8$, which contradicts our assumption.

For the moreover part, notice that the solution SDP^k is valid for the second level of Lasserre, which means it induces locally consistent distributions for any two vertices. For any edge (p, q) in the graph, if both p and q have events $\{p, C_1\}$, $\{q, C_2\}$ that happen with probability more than $1/2$, then we show $C_1 \neq C_2$. Suppose for contradition that $C_1 = C_2 = C$. If we look at the distribution that the Lasserre solution induces on these two vertices, with positive probabililty both of them will be colored with color C. This contradicts with the validity of the Lasserre solution. Therefore we must have $C_1 \neq C_2$.

Local correlation. For an SDP solution, we would want to argue either Lemma 2 can be applied or the solution has large global correlation. To show this, we introduce local correlation as an intermediate step. We first show that if we cannot apply Lemma 2, the solution SDP^k always has *local correlation*, then we analyze the relationship between local correlation and global correlation in the next section and show high local correlation implies high global correlation.

For a vertex p, we construct a new vector $z_p = (u_{p,R}, u_{p,B}, u_{p,Y})$ (which means z_p is the concatenation of the 3 vectors, the vector u comes from Theorem 2). It's easy to see that $\langle z_p, z_q\rangle = \sum_{C\in\{R,Y,B\}}\langle u_{p,C}, u_{q,C}\rangle$. Since $\langle z_p, z_q\rangle^2 \leq$

$27\,\mathrm{E}_{C_1,C_2\in\{R,Y,B\}}\langle u_{p,C_1},u_{q,C_2}\rangle^2$, we know $GC \ge 1/27\cdot \mathrm{E}_{p,q\in V}\langle z_p,z_q\rangle^2$. Hence the global correlation of the $\{z_p\}$ vectors $\mathrm{E}_{p,q\in V}\langle z_p,z_q\rangle^2$ can be used to lowerbound the global correlation of the solution SDP^k.

Local correlation is the expected correlation between endpoints of edges.

Definition 2 (Local Correlation). *Given a graph G and an SDP solution SDP^k, first construct vectors $z_p = (u_{p,R}, u_{p,B}, u_{p,Y})$. Then local correlation for this solution is defined to be $LC = \mathrm{E}_{(p,q)\in E}\langle z_p, z_q\rangle$.*

Local correlation depends on both the solution (SDP^k) and the graph G, unlike global correlation which only depends on the solution. Also, local correlation can be negative because we are not taking the squares of inner-products.

We shall prove the following Lemma which ensures high local correlation until we can find a large independent set.

Lemma 3. *If G is a regular 3-colorable graph, and in an SDP solution SDP^k at most $n/4$ vertices are determined in the sense of Lemma 2 then the local correlation $\mathrm{E}_{(p,q)\in E}\langle z_p, z_q\rangle \le -1/8$.*

Proof. If (p,q) is an edge, both p and q are undetermined (as in Lemma 2), we shall prove $\langle z_p, z_q\rangle \le -1/4$. Indeed, since (p,q) is an edge by (2) we have $\langle v_{p,R}, v_{q,R}\rangle = r_{p,R}r_{q,R} + \langle u_{p,R}, u_{q,R}\rangle = 0$. Which means $\langle z_p, z_q\rangle = -r_{p,R}r_{q,R} - r_{p,Y}r_{q,Y} - r_{p,B}r_{q,B} \le -1/4$. The inequality holds because the r values are all in [0,1/2], the worst case for the r values are $(1/2, 1/2, 0)$ and $(0, 1/2, 1/2)$.

Since only 1/4 of the vertices are determined (in the sense of Lemma 2), we consider the set S of undetermined vertices. At least 1/2 of edges of G have both endpoints in S. Therefore $\mathrm{E}_{(p,q)\in E}\langle z_p, z_q\rangle \le -1/4 * 1/2 = -1/8$.

4 Threshold Rank and Global Correlation

In this section we show how local correlation and global correlation are connected through *threshold rank*. Threshold rank of a graph $Rank_C(G)$ is defined by Arora et.al. in [1] as the number of eigenvalues larger than C. As they observed in [1], many problems have subexponential time algorithms when the underlying graph has low (i.e. sublinear) threshold rank. We show that 3-*Coloring* also lies in this category. If the underlying graph has low threshold rank, then an SDP solution will have high global correlation as long as it has local correlation.

Our definition for threshold rank is different from [1]. We are interested in eigenvalues that are smaller than a certain negative constant $-C$. For a graph G, we use $Rank_{-C}(G)$ to denote the number of eigenvalues of G's normalized ajacency matrix whose value is at most $-C$. In all discussions C should be viewed as a positive constant, and we use negative sign to indicate that we are interested in eigenvalues smaller than $-C$.

Consider a convex relaxation of threshold rank given by Barak et.al.[3]. In this relaxation each vertex in the graph has a vector z_p (later we will see that they are indeed related to the vectors $\{z_p\}$ in Lemma 3). We try to maximize D, the reciprocal of global correlation subject to the following constraints

$$\mathop{\mathrm{E}}_{p\in V} \|z_p\|_2^2 = 1 \tag{7}$$

$$\mathop{\mathrm{E}}_{(p,q)\in E} \langle z_p, z_q\rangle \le -C \tag{8}$$

$$\mathop{\mathrm{E}}_{p,q\in V} (\langle z_p, z_q\rangle)^2 \le 1/D. \tag{9}$$

Barak et.al.[3] proved the following Lemma explaining why this is a relaxation to threshold rank. Due to space limit the proofs are omitted here.

Lemma 4. *If the $Rank_{-C/2}(G) = D$, then the optimal value D^* of the convex relaxation is at most $4D/C^2 = Rank_{-C/2}(G)/(C/2)^2$.*

Lemma 4 is important for our analysis because it implies if the local correlation (left-hand-side of Equation (8)) is smaller than a negative constant, and threshold rank is low, the global correlation (left-hand-side of Equation (9)) must be of order $\Omega(1/D)$. Now we are ready to prove Theorem 1:

Proof. Write the SDP in Section 2 with $c \cdot D$ levels of Lasserre lifting ($Las^{c\cdot D}$), and solve it in time $n^{O(D)}$. We apply the following rounding algorithm inspired by Lemma 1.

1. Initialize SOL to be $SDP^{c\cdot D}$
2. Repeat
3. If at least $n/4$ of the vertices are "determined" in SOL
4. Then apply Lemma 2 to get a partial coloring .
5. Pick a random event w, condition the solution SOL on this event
6. Until SOL is only valid for the first Level of Lasserre

Clearly, if the condition in Step 3 is satisfied and we proceed to Step 4, by Lemma 2 we get a partial coloring for $n/4$ vertices. In particular, one of the colors will have more than $n/12$ vertices, and they form an independent set. Therefore we only need to prove the probability that we reach Step 4 is large.

Let r_i be the probability that Step 4 is reached before iteration i. We would like to prove $r_{c\cdot D} \ge 1/2$. Assume we continue to run the algorithm even if Step 4 is reached (and we have already found an independent set). Let SOL_i be the solution at step i, GC_i be its global correlation and VAR_i be its variance.

We first prove the following Claim:

CLAIM: If the number of undetermined vertices in SOL_i is smaller than $n/4$, the global correlation GC_i is at least $\Omega(1/D)$.

Proof. Given the assumption, we can apply Lemma 3. From the solution SOL_i, Lemma 3 constructs vectors $\{z_p\}(p \in V)$, and $\mathrm{E}_{(p,q)\in E}\langle z_p, z_q\rangle \le -1/8$.

We shall normalize these vectors so that they satisfy Equations (7) and (8). The norm of z_p is $\|z_p\|_2^2 = \|u_{p,R}\|_2^2 + \|u_{p,Y}\|_2^2 + \|u_{p,B}\|_2^2 = 1 - r_{p,R}^2 - r_{p,Y}^2 - r_{p,B}^2$. Here $r_{p,X}$ is the probability that p is colored with color X, and the equation follows from Theorem 2. If for vertex p no event has probability more than $1/2$, then $\|z_p\|_2^2$ is a value between 1/4 and 1. As assumed the number of such vertices is at least $3n/4$ (otherwise Step 4 has already been performed), thus $\mathrm{E}_{p\in V} \|z_p\|_2^2$ is between 3/16 and 1. We can normalize these vectors by multiplying with $c' = \sqrt{1/\mathrm{E}_{p\in V} \|z_p\|_2^2}$. For the normalized vectors $\{\overline{z_p}\}$, we have $\mathrm{E}_{p\in V} \|\overline{z_p}\|_2^2 = 1$. And since $c' \geq 1$ we still have $\mathrm{E}_{(p,q)\in E}\langle\overline{z_p}, \overline{z_q}\rangle \leq -1/8$.

The vectors $\{\overline{z_p}\}$ satisfy Equation (7) and (8) for $C = -1/8$. Since we know $Rank_{-1/16}(G) = D$, Lemma 4 shows that the left-hand-side of Equation (9) must be at least $(1/16)^2/D = \Omega(1/D)$. That is, the global correlation between vectors $\{\overline{z_p}\}$ is at least $\Omega(1/D)$.

By analysis in Section 3, we know GC_i is within a constant factor of $\mathrm{E}_{p,q}\langle z_p, z_q\rangle^2$. Since the normalization factor c' between z_p and $\overline{z_p}$ is also bounded by a constant, $\mathrm{E}_{p,q}\langle z_p, z_q\rangle^2$ and $\mathrm{E}_{p,q}\langle\overline{z_p}, \overline{z_q}\rangle^2$ are also within a constant factor. Thus $GC_i \geq \Omega(1/D)$.

The proof proceeds as follows: when r_i, the probability that the solution has more than 3/4 "determined" vertices, is large we can already get a good solution by applying the moreover part of Lemma 2. Otherwise we can apply the claim and Lemma 4 to conclude that the expected global correlation must be high at step i; then Lemma 1 reduces r_i significantly.

In step i, with probability $1 - r_i$ the number of "determined" vertices (in the sense of Lemma 2) is smaller than $n/4$. When this happens (number of determined vertices small), Lemma 4 shows the global correlation is at least $\Omega(1/D)$. Therefore the expected global correlation at step i is at least $\mathrm{E}[GC_i] \geq \Omega(1/D) * 1/2 = \Omega(1/D)$ (the expectation is over random choices of the algorithm) just by considering the situations when number of determined vertices is small. By Lemma 1 we know every time Step 5 is applied, the variance is expected to reduce by GC_i. That is, $\mathrm{E}[VAR_{i+1}] \leq \mathrm{E}[VAR_i] - 4\,\mathrm{E}[GC_i] \leq \mathrm{E}[VAR_i] - \Omega(1/D)$. If r_i remains smaller than 1/2 for all the $c \cdot D$ rounds (where c is a large enough constant), we must have $\mathrm{E}[VAR_{c \cdot D}] < 1/16$. By Markov's Inequality with probability at least 1/2 the variance is at most 1/8, in which case Lemma 2 can be applied. That is, $r_{c\cdot D} \geq 1/2$. This is a contradiction and we must have $r_i \geq 1/2$ for some $i \leq c \cdot D$.

Therefore with probability at least 1/2 the rounding algorithm will reach Step 4 and find a large independent set.

For the moreover part, we apply a random permutation π over the vertices before running the whole algorithm. In the permuted graph $n/12$ of the vertices are in the independent set S found by the algorithm above. If we apply the inverse permutation π^{-1} to the independent set found, we claim that any vertex q of the original graph is inside the independent set $\pi^{-1}(S)$ with probability at least 1/12. This is because the graph is vertex transitive and essentially the algorithm cannot distinguish between vertices. More rigorous argument can be found in full version.

Repeat this procedure $100 \log n$ times, each vertex is in one of the $100 \log n$ independent sets with probability at least $1 - n^{-2}$. Union bound shows with high probability the union of these independent sets is the vertex set. We use one color for each independent set (if a vertex belongs to multiple sets then choose an arbitrary one among them), which gives a valid $O(\log n)$ coloring.

5 Threshold Rank Bound for Distance Transitive Graphs

As we explained in the Introduction, symmetric graphs are a natural class of hard instances for graph coloring problem. Also, by Blum Coloring Tools [5], it is enough to consider graphs with low diameter for 3-Coloring.

In this section we focus on a class of symmetric graphs: distance transitive graphs, and we prove for a distance transitive graph with diameter Δ, the threshold rank $Rank_{-C}(G)$ is at most $(O(1/C^2))^{\Delta}$. We begin by defining distance transitive graphs:

Definition 3 (Distance Transitive Graph). *A graph $G = (V, E)$ is distance-transitive if for any pairs of vertices (p, q) and (s, t), where the shortest-path distance between p, q and s, t are the same, there is always an automorphism that maps p to s and q to t.*

Distance transitive graphs have many nice properties, especially when we look at the neighborhoods of vertices. Define $\Gamma^k(p)$ to be the k-th neighborhood of p (which is the set of vertices at distance k of p), by the distance transitive condition, we know if a pair of vertices p, q have distance k, then $|\Gamma^{k-1}(p) \bigcap \Gamma(q)|$, $|\Gamma^k(p) \bigcap \Gamma(q)|$, $|\Gamma^{k+1}(p) \bigcap \Gamma(q)|$ are three numbers that depend only on k. As a convention, we call these numbers c_k, a_k and b_k respectively. The size of k-th neighborhood ($|\Gamma^k(p)|$) is represented by n_k. The following is known about spectral properties of distance transitive graphs[4]:

Lemma 5. *A distance transitive graph G has $\Delta + 1$ distinct eigenvalues, which are the eigenvalues of the matrix*

$$B = \begin{pmatrix} a_0 & c_1 & & & \\ b_0 & a_1 & c_2 & & \\ & \cdot & \cdot & \cdot & \\ & & b_{\Delta-2} & a_{\Delta-1} & c_{\Delta} \\ & & & b_{\Delta-1} & a_{\Delta} \end{pmatrix}.$$

Moreover, suppose the i-th eigenvalue is λ_i ($\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{\Delta}$), with left eigenvector u_i and right eigenvector v_i ($u_i^T B = \lambda_i u_i^T$, $Bv_i = \lambda_i v_i$), we have $v_i(j) = n_j u_i(j)$. When v_i is normalized so that $v_i(0) = 1$, the multiplicity of λ_i in the original graph is $n/\langle u_i, v_i\rangle$.

Note that the eigenvalues in the above Lemma are for the adjacency matrix of G. There is a normalization factor d (the degree) between these eigenvalues and eigenvalues we were using for threshold rank.
Now we are ready to prove the following theorem.

Theorem 3. *If a distance transitive graph is 3-colorable, then there is an algorithm that colors it with $O(\log n)$ colors in time $n^{2^{O(\Delta)}}$.*

Proof. Due to space limit please see the full version of this paper for the proof. The main idea is to prove that for each eigenvalue smaller than $-C$, its multiplicity must be smaller than $(10/C^2)^{\Delta}$ (the multiplicity can be computed by Lemma 5), then Theorem 1 gives the algorithm.

6 Conclusion

In this paper we explored the relationship between threshold rank and graph coloring. Unlike other problems such as Unique Games and MAX-CUT considered by Arora et.al.[1], we show that 3-Coloring is actually related to the negative side of the spectrum. We give an algorithm that can find linear size independent set when the graph is 3-colorable and has threshold rank D. The efficiency of our algorithm depends on the threshold rank of a graph. Known integrality gap examples [11,12] all have threshold rank that is polylog in the number of vertices. Thus our algorithm can detect in quasipolynomial time that they are not 3-Colorable.The relationship between global correlation and rounding and the convex relaxation for threshold rank are inspired by Barak e.al.[3] and we believe these techniques can be useful in other problems.

If our approach is combined with combinatorial tools, it could possibly lead to good subexponential (or even quasipolynomial-time) coloring algorithms. In particular, if the following conjecture is true for any constant C and $D = n^{\delta}$, we get a $\exp(n^{\delta})$ time algorithm for coloring 3-Colorable graph with n^{ϵ} colors (see full version). We have no counterexamples for the Conjecture when C is a constant and D is more than n^{ϵ}.

Conjecture 1. There exists an algorithm such that for any graph G, can either

- Find a subset S of vertices. The vertex expansion is at most $\Phi_V(S) \leq (n/|S|)^{1/C}$.
- Certify the existence of doubly stochastic matrix M with same support as G such that $Rank_{-1/16}(M) \leq D$.

We also give efficient algorithm to color 3-colorable distance transitive graphs with low diameter. These graphs have properties that seem to make it hard for previously known algorithms.

Acknowledgement. We thank David Steurer, Grant Schoenebeck and Aravindan Vijayaraghavan for valuable discussions. We especially thank David Steurer for sharing the result [3] with us.

References

1. Arora, S., Barak, B., Steurer, D.: Subexponential algorithms for unique games and related problems. In: Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science, FOCS 2010, pp. 563–572. IEEE Computer Society, Washington, DC, USA (2010)

2. Arora, S., Chlamtac, E., Charikar, M.: New approximation guarantee for chromatic number. In: Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing, STOC 2006, pp. 215–224. ACM, New York (2006)
3. Barak, B., Raghavendra, P., Steurer, D.: Rounding semidefinite programming hierarchies via global correlation (2011)
4. Biggs, N.L.: Intersection matrices for linear graphs. In: Welsh, D.J.A. (ed.) Combinatorial Mathematics and its Applications, pp. 15–23. Academic Press, London (1971)
5. Blum, A.: New approximation algorithms for graph coloring. J. ACM 41, 470–516 (1994)
6. Blum, A., Karger, D.: An $\tilde{O}(n^{3/14})$-coloring algorithm for 3-colorable graphs. Information Processing Letters 61, 49–53 (1996)
7. Brouwer, A.E., Cohen, A.M., Neumaier, A.: Distance Regular Graphs. Springer, Heidelberg (1989)
8. Chlamtac, E.: Non-Local Analysis of SDP-Based Approximation Algorithms. PhD thesis, Princeton University (2009)
9. Chlamtac, E., Tulsiani, M.: Convex relaxations and integrality gaps. Unpublished manuscript
10. Dinur, I., Mossel, E., Regev, O.: Conditional hardness for approximate coloring. In: Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing, STOC 2006, pp. 344–353. ACM, New York (2006)
11. Feige, U., Langberg, M., Schechtman, G.: Graphs with tiny vector chromatic numbers and huge chromatic numbers. In: Proceedings of the 43rd Symposium on Foundations of Computer Science, FOCS 2002, pp. 283–292. IEEE Computer Society, Washington, DC, USA (2002)
12. Frankl, P., Rodl, V.: Forbidden intersections. Transactions of the American Mathematical Society 300(1), 259–286 (1987)
13. Karger, D., Motwani, R., Sudan, M.: Approximate graph coloring by semidefinite programming. J. ACM 45, 246–265 (1998)
14. Lasserre, J.B.: An explicit exact SDP relaxation for nonlinear 0-1 programs. In: Aardal, K., Gerards, B. (eds.) IPCO 2001. LNCS, vol. 2081, pp. 293–303. Springer, Heidelberg (2001)
15. Laurent, M.: A comparison of the sherali-adams, lovasz-schrijver and lasserre relaxations for 0-1 programming. Mathematics of Operations Research 28, 470–496 (2001)
16. Lovász, L., Schrijver, A.: Cones of matrices and set-functions and 0-1 optimization. SIAM Journal on Optimization 1, 166–190 (1991)
17. Sherali, H.D., Adams, W.P.: A hierarchy of relaxation between the continuous and convex hull representations. SIAM J. Discret. Math. 3, 411–430 (1990)
18. Szegedy, M.: A note on the θ number of lovasz and the generalized delsarte bound. In: Proceedings of the 35th Annual Symposium on Foundations of Computer Science, pp. 36–39. IEEE Computer Society, Washington, DC, USA (1994)
19. Wigderson, A.: Improving the performance guarantee for approximate graph coloring. J. ACM 30, 729–735 (1983)
20. Zuckerman, D.: Linear degree extractors and the inapproximability of max clique and chromatic number. In: Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing, STOC 2006, pp. 681–690. ACM, New York (2006)