Approximation Algorithms Using Hierarchies of Semidefinite Programming Relaxations

Eden Chlamtac*
Princeton University
E-mail: chlamtac@cs.princeton.edu

Abstract

We introduce a framework for studying semidefinite programming (SDP) relaxations based on the Lasserre hierarchy in the context of approximation algorithms for combinatorial problems. As an application of our approach, we give improved approximation algorithms for two problems. We show that for some fixed constant $\varepsilon > 0$, given a 3-uniform hypergraph containing an independent set of size $(\frac{1}{2} - \varepsilon)n$, we can find an independent set of size $\Omega(n^{\varepsilon})$. This improves upon the result of Krivelevich, Nathaniel and Sudakov, who gave an algorithm finding an independent set of size $\tilde{\Omega}(n^{6\gamma-3})$ for hypergraphs with an independent set of size γn (but no guarantee for $\gamma \leq \frac{1}{2}$). We also give an algorithm which finds an $O(n^{0.2072})$-coloring given a 3-colorable graph, improving upon the work of Arora, Chlamtac and Charikar. Our approach stands in contrast to a long series of inapproximability results in the Lovász Schrijver linear programming (LP) and SDP hierarchies for other problems.

1 Introduction

Semidefinite Programming (SDP) has been one of the most important tools in designing approximation algorithms for combinatorial optimization problems for the last several years. Starting with the seminal work of Goemans and Williamson [14] on MAXCUT, there has been a series of results on a diverse range of combinatorial problems. While for a number of problems, including MAXCUT [14], MAX-3SAT [17, 28], and Unique Games [10], SDPs lead to approximation algorithms which are essentially optimal under certain complexity-theoretic assumptions [15, 19], for a host of other problems the gap between known hardness of approximation and approximation algorithmic guarantee remains quite large.

One possible avenue of improvement on the approximation side is the use of so-called SDP hierarchies. Such hierarchies have been proposed by Lovász and Schrijver [23], Sherali and Adams [26], and Lasserre [22] (see [21] for a comparison). Each hierarchy is characterized by a method (such methods are known collectively as "lift-and-project") by which one can take a semidefinite relaxation for an integer $(0 - 1)$ program, and strengthen it repeatedly, thus constructing the various levels of the hierarchy. These hierarchies have the joint property that for an integer program on n variables, the nth level of the hierarchy is equivalent to the original integer program.

The quality of approximation of SDP hierarchies has been studied more generally in the context of optimization of polynomials over semi-algebraic sets [12, 6, 24]. In the combinatorial optimization setting, there has been a series of negative results, starting with [2], showing that the Lovász-Schrijver hierarchies LS (a *linear programming* hierarchy) and LS_+ (the SDP variant) do not yield good approximations for certain problems. For Vertex Cover, Schoenebeck et al. [25] showed that the integrality gap of the standard LP relaxation is $2 - o(1)$ even after $\Omega(n)$ rounds of LS, while Georgiou et al. [13] have shown that an integrality gap of $2 - o(1)$ survives $\Omega(\sqrt{\log n / \log\log n})$ rounds of LS_+. For MAX-3SAT, Hypergraph Vertex Cover and Set Cover, Alekhnovich et al. [1] showed that $\Omega(n)$ rounds of LS_+ do not give any nontrivial approximations.

Given that the kth level of any of these hierarchies is only known to be solvable in polynomial time for constant k, it is natural to ask whether for any combinatorial optimization problem, SDP hierarchies yield improved approximations at a constant level. One reason to expect that this should be the case is the following property common to the three SDP hierarchies mentioned above. For any set of k variables (for example, indicator functions for whether vertices are in an independent set), any solution to the kth level of the hierarchy projected onto this set is a convex combination of legal $0 - 1$ solutions. For this reason and others, a good candidate problem is one for which local properties propagate to a global scale without much loss.

More concretely, we propose the following heuristic. Given an algorithm which rounds an SDP solution, ana-

*Supported by a Francis Upton fellowship.

0272-5428/07 $25.00 © 2007 IEEE
DOI 10.1109/FOCS.2007.72

lyze this algorithm under the assumption that the solution is in fact a convex combination of legal $0-1$ solutions (as a thought experiment). If the analysis is sufficiently local, then it should also apply to the kth level of an SDP hierarchy. To this end, we offer a tool to apply certain kinds of analyses of (convex combinations of) integral solutions to the setting of SDP hierarchies based on the Lasserre relaxation of Independent Set. As an application, we investigate two problems for which the status of their approximability is still open, namely, Maximum Independent Set in 3-uniform hypergraphs and Graph Coloring. For both problems we show an improvement in the quality of the guaranteed approximation. In the case of the first problem, this is also a proven improvement in the integrality gap.

We start with the problem of finding large independent sets in 3-uniform hypergraphs. k-uniform hypergraphs are collections of sets of size k ("hyperedges") over a vertex set. An independent set is a subset of the vertices which does not fully contain any hyperedge. This problem was previously explored by Krivelevich et al. [20], who showed that for any 3-uniform hypergraph on n vertices containing an independent set of size γn, one can find an independent set of size $\tilde{\Omega}(\min\{n, n^{6\gamma-3}\})$. This does not yield any nontrivial guarantee for $\gamma \leq \frac{1}{2}$, and in fact one can construct tight integrality gaps for this range of γ showing that their SDP relaxation is satisfied (i.e. has an optimum value at least γn) even when the hypergraph contains no independent sets larger than 2. In contrast, we show that using the third level of the hierarchy, one can find an independent set of size $\Omega(n^{\varepsilon})$ whenever $\gamma \geq \frac{1}{2} - \varepsilon$, for some fixed $\varepsilon > 0$.

The second problem we consider is that of coloring 3-colorable graphs with as few colors as possible. This problem has a long history of study starting with the $O(\sqrt{n})$-coloring algorithm of Wigderson [27], through the sophisticated combinatorial approach of Blum [7], the SDP approach of Karger, Motwani and Sudan [16], and finally the $\tilde{O}(n^{3/14})$ approximation of Blum and Karger [8]. Recently, this result was improved by Arora, Chlamtac and Charikar [3], who gave an $O(n^{0.2111})$-coloring algorithm using a new geometric analysis of the SDP rounding similar to the SPARSEST CUT result of Arora, Rao and Vazirani [5]. While we borrow some of the basic terminology and tools introduced in [3], we deviate significantly from their approach, in that our analysis does not involve any event-chaining or measure-concentration results. Reducing the problem to Max Independent Set, and using the third level of the corresponding Lasserre relaxation, we find a legal coloring of the graph using $O(n^{0.2072})$ colors.

In related work, the present author and Singh [9] have also shown that 2-colorable 4-uniform hypergraphs can be colored using at most $O(n^{3/4-\varepsilon})$ colors (for some constant $\varepsilon > 0$), improving upon the previous $O(n^{3/4})$-coloring algorithm of Chen and Frieze [11].

The rest of the paper is organized as follows. In Section 2 we define the SDPs used in the various algorithms, and show some useful properties of relaxations of this form which are used in the analysis. In sections 3 and 4 we give the results for hypergraph independent sets, and graph coloring, respectively. The analysis of the coloring algorithm relies on some unpublished lemmas which were a continuation of the work done by Arora, Chlamtac and Charikar [3] (see Appendix A). We stress that our current improvement also gives a better guarantee than that which is achievable by applying these tools to the analysis in [3].

2 SDP relaxations and preliminaries

2.1 Independent Set relaxations using the Lasserre Hierarchy

The Lasserre hierarchy [22] is a sequence of nested semidefinite relaxations for certain $0-1$ polynomial programs. These SDPs may be expressed as a system of constraints on the vectors $\{v_I | I \subseteq [n]\}$. To obtain a relaxed (non-integral) solution to the original problem, one takes $(v_{\{1\}}^2, v_{\{2\}}^2, \ldots, v_{\{n\}}^2)$. (For convenience, we will henceforth write $v_{i_1\ldots i_s}$ instead of $v_{\{i_1,\ldots,i_s\}}$.)

When the $0-1$ polytope is the convex hull of all (indicator functions of) independent sets in a hypergraph (or graph) $H = (V, E)$, the constraints in the kth level of the hierarchy may be expressed as follows (see [21]):
$IS_k(H)$

$$v_\emptyset^2 = 1 \qquad (1)$$

$$\forall I, J, I', J' \subseteq V \text{ s.t. } |I|, |J|, |I'|, |J'| \leq k \text{ and } I \cup J = I' \cup J' \qquad v_I \cdot v_J = v_{I'} \cdot v_{J'} \qquad (2)$$

$$\forall e \in E \qquad v_e^2 = 0 \qquad (3)$$

We will denote by MAX-IS$_k(H)$ the SDP

$$\text{Maximize } \textstyle\sum_i \|v_i\|^2 \text{ s.t. } \{v_I\}_I \text{ satisfy } IS_k(H).$$

As shown in [21], these constraints imply $v_I = 0$ for any set I of at most k vertices containing at least one hyperedge. As a relaxation of the integer program over $0-1$ variables $\{x_i\}$, the vector v_I may be interpreted as a representing the value $\prod_{i \in I} x_i$. However, it will be more useful to think of $\{x_i\}$ as random $0-1$ variables. We then think of the PSD matrix $M = (v_I \cdot v_J)_{I,J}$ as the expectation over the corresponding random $0-1$ matrices, and the values $v_I \cdot v_J$ represent the probability that $x_i = x_j = 1$ for all $i \in I$ and $j \in J$ (in fact, if we limit ourselves to any fixed index set of size $\leq k$, then this interpretation is correct for the kth level of the Lasserre hierarchy). The picture may be completed (to include variables representing mixed $\{0,1\}$ partial assignments) by defining, for all $I \subseteq [n]$ and $J \subseteq$

$\{-i|i \in [n]\}$,

$$v_{I\cup J} \stackrel{\text{def}}{=} \sum_{J'\subseteq\{j|-j\in J\}} (-1)^{|J'|} v_{I\cup J'}.$$

The following lemma (whose proof is straightforward) relates these variables to the above SDP.

Lemma 1. *Constraints* (1) *and* (2) *above for $k = 2l$ imply all the constraints in $IS_l(V', E')$ where $V' = [n] \cup \{-i|i \in [n]\}$, and $E' = \{(i, -i)|i \in [n]\}$.*

As a thought experiment, we will always think of the vectors v_I as representing the distributions of the random variables $\prod_{i\in I} x_i$, as discussed above. This intuition will allow us to deduce certain properties of the geometry, which can then be proven rigorously using the Lasserre constraints (2). Let us consider the following crucial example (which will also motivate the following lemma). Consider some event A relating to partial assignments of $\{x_i\}$ (e.g. "$\forall i \in I : x_i = 1$"). Suppose that $\Pr[A] = p$ and that we have many events B_j, sub-events of A, such that $\Pr[B_j|A] = q$. Since most pairs of events cannot be too anti-correlated, for most pairs B_j, B_l we have $\Pr[B_j \wedge B_l|A] \geq q^2 - o(1)$. If we think of the vectors representing these events, we have $v_{B_j} \cdot v_{B_l} \geq pq^2 - o(1)$. Since this is true for most pairs B_j, B_l, one would imagine that they all share a common component of length $\sqrt{pq^2}$. That is, that there exists some unit vector v'_A such that $v_{B_j} \cdot v'_A \geq \sqrt{pq^2}$. Similarly, if for some A', a super-event of A, we were guaranteed that the vectors v_{B_j} had the form $v_{B_j} = \sqrt{p'} \cdot \frac{v_{A'}}{\|v_{A'}\|} + w_{B_j}$ for some $w_{B_j} \perp v_{A'}$, we could argue that the vectors w_{B_j} should have a common component of length $\geq \sqrt{pq^2 - p'}$. Using the Lasserre hierarchy, we can guarantee the existence of such a vector, as demonstrated by the following lemma (in this case think of the mutually exclusive events "$\forall l \in I_i : x_l = 1$" and the respective sub-events "$(\forall l \in I : x_l = 1) \wedge (\forall j \in J : x_j = 1)$").

Lemma 2. *Let $\{v_I\}$ be a set of vectors satisfying* (2), *let subsets $I_i \subset [n]$ and $J \subseteq [n]$ of size at most k be such that $\forall i$, $I_i \cap J = \emptyset$ and $\forall i \neq j$, $v_{I_i} \cdot v_{I_j} = 0$ and let $p_i = \|v_{I_i}\|^2$, and $q_i = \|v_{I_i\cup J}\|^2 / \|v_{I_i}\|^2$. Then*

1. *There exists a unit vector $x_0 \in$ Span$(\{v_I | I \subseteq \bigcup_i I_i\})$ such that $x_0 \cdot v_J = \sqrt{\sum_i p_i q_i^2}$.*
2. *If, moreover, for every i there are subsets I_{ij} satisfying $I_i \subseteq I_{ij} \subseteq [n] \setminus I_J$ such that the vectors $v_{I_{ij}}$ are mutually orthogonal, and $v_{I_i} = \sum_j v_{I_{ij}}$, then if v'_J is the component of v_J orthogonal to x_0 (i.e. $v_J = \sqrt{\sum_i p_i q_i^2} x_0 + v'_J$), then there exists a unit vector $x'_0 \in$ Span$(\{v_I | I \subseteq \bigcup_{i,j} I_{ij}\})$ such that $x'_0 \cdot v'_J = \sqrt{\sum_{i,j} p_{ij} q_{ij}^2 - \sum_i p_i q_i^2}$ (where $p_{ij} = \|v_{I_{ij}}\|^2$ and $p_{ij} q_{ij} = \|v_{I_{ij}\cup J}\|^2$).*

Proof. (sketch) It suffices to check, by computing inner products, and using constraint (2), that $v_J = \sum_i q_i v_{I_i} + v'_J$ (where $v'_J \cdot v_{I_i} = 0$), and that for all i, $\sum_j q_{ij} v_{I_{ij}} = q_i v_{I_i} + v''_i$, where v''_i is a vector of length $\sqrt{\sum_j p_{ij} q_{ij}^2 - p_i q_i^2}$ orthogonal to $v_{I_{ij}}$ for all j. □

The above lemma motivates the following definition:

Definition 3. We will call a set of unit vectors X a *ρ-cluster* if there exists a unit vector x_0 such that $x_0 \cdot x \geq \sqrt{\rho}$ for all $x \in X$.

These clusters will be crucial in obtaining a more refined analysis of rounding algorithms, as we shall see in section 2.4.

2.2 SDP relaxations for MAX-IS in 3-uniform hypergraphs

The relaxation proposed in [20] may be derived as follows. Given an independent set $I \subseteq V$ in a 3-uniform hypergraph $H = (V, E)$, for every vertex $i \in V$ let $x_i = 1$ if $i \in I$ and $x_i = 0$ otherwise. For any hyperedge $(i, j, l) \in E$ it follows that $x_i + x_j + x_l \in \{0, 1, 2\}$ (and hence $|x_i + x_j + x_l - 1| \leq 1$). Thus, we have the relaxation
$KNS(H)$

$$\text{Maximize } \sum_i \|v_i\|^2 \text{ s.t. } \quad v_\emptyset^2 = 1 \tag{4}$$
$$\forall i \in V \quad v_\emptyset \cdot v_i = v_i \cdot v_i \tag{5}$$
$$\forall (i, j, l) \in E \quad \|v_i + v_j + v_l - v_\emptyset\|^2 \leq 1 \tag{6}$$

One can check that in fact for $k \geq 3$ constraint (6) is implied by $IS_k(H)$.

2.3 SDP relaxation for 3-coloring

We reduce the 3-coloring problem to an Independent Set problem as follows. Given a graph $G = (V, E)$, construct graph $G' = (V', E')$ where $V' = V \times \{R, B, Y\}$ and $((i, C_1), (j, C_2)) \in E'$ if $(i, j) \in E$ and $C_1 = C_2$, or if $i = j$ and $C_1 \neq C_2$. Note that any independent set of size $|V|$ in G' induces a 3-coloring of G (since every vertex $i \in V$ appears in exactly one of the three copies of V in G'). It is not hard to see that if MAX-IS$_k(G') = n$, then in an optimal solution, for all $i \in V$ we have $v_\emptyset = v_{(i,R)} + v_{(i,B)} + v_{(i,Y)}$. Moreover, since the constraints of $IS_k(G')$ are symmetric with respect to $\{R, B, Y\}$, for any matrix $M \in IS_k(G')$, the matrix $\frac{1}{6} \sum_{\pi\in Sym(\{R,B,Y\})} \pi(M)$ also satisfies $IS_k(G')$, where $Sym(X)$ is the group of permutations on X, and $\pi(M)$ is defined as follows: $\pi(M)_{I,J} = M_{\pi(I),\pi(J)}$ where for any $I \subseteq V'$, $\pi(I) = \{(i, \pi(C))|(i, C) \in I\}$. Thus, we arrive at the following SDP relaxation for 3-coloring:

$3COL_k(G)$

$$\{v_I | I \subseteq V'\} \in IS_k(G') \quad (7)$$

$$\forall i \in V \quad v_\emptyset = \sum_{C \in \{R,B,Y\}} v_{(i,C)} \quad (8)$$

$$\forall \pi \in Sym(\{R,B,Y\})$$
$$\forall I, J \subseteq V', |I|, |J|, \leq k \quad v_I \cdot v_J = v_{\pi(I)} \cdot v_{\pi(J)} \quad (9)$$

We now show that the relaxation $3COL_1(G)$ is equivalent to the standard SDP relaxation for 3 coloring. This will be useful later on, as we will use an SDP rounding algorithm very similar to the ones in [16] and [3]. For all $i \in V$, by constraints (2) (8) and (9), we have $v_\emptyset \cdot v_{(i,R)} = \left\|v_{(i,R)}\right\|^2 = \frac{1}{3}$. Thus every $v_{(i,R)}$ can be written

$$v_{(i,R)} = \tfrac{1}{3} v_\emptyset + \tfrac{\sqrt{2}}{3} u_i, \quad (10)$$

where u_i is a unit vector orthogonal to $v_\emptyset$. We claim that the vectors $\{u_i\}$ are a *vector 3-coloring* of G, that is, that they satisfy

$$\forall (i,j) \in E \quad u_i \cdot u_j = -\tfrac{1}{2}. \quad (11)$$

Indeed, this follows immediately from (10), since $v_{i,R} \cdot v_{j,R} = \frac{1}{9} v_\emptyset^2 + \frac{2}{9} u_i \cdot u_j$. It is not hard to see that one can similarly construct a solution to $3COL_1(G)$ given any vector 3-coloring $\{u_i\}$.

2.4 Gaussian vectors and SDP rounding

Recall that the *standard normal distribution* has density function $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. A random vector $\zeta = (\zeta_1, \ldots, \zeta_n)$ is said to have the *n-dimensional standard normal distribution* if the components ζ_i are independent and each have the standard normal distribution. Note that this distribution is invariant under rotation, and its projections onto orthogonal subspaces are independent. In particular, for any unit vector $v \in \Re^n$, the projection $\langle \zeta, v \rangle$ has the standard normal distribution. Moreover, for any orthogonal subspaces $U, W \subset \Re^n$, the projections of ζ onto U, W, respectively, are independent.

We use the following notation for the tail bound of the standard normal distribution: $N(x) \stackrel{\text{def}}{=} \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$. The following property of the normal distribution will be crucial.

Lemma 4. *For $s > 0$, we have*

$$\tfrac{1}{\sqrt{2\pi}} \left(\tfrac{1}{s} - \tfrac{1}{s^3}\right) e^{-s^2/2} \leq N(s) \leq \tfrac{1}{\sqrt{2\pi} s} e^{-s^2/2}.$$

The analysis of SDP rounding algorithms frequently involves expressions of the form $\Pr_\zeta[\exists x \in X : \zeta \cdot x \geq t]$, for random vector ζ as above, and set of unit vectors X. It is easy to see that $|X| N(t)$ is an upper-bound on this probability. However, when the set X is a ρ-cluster, we can give a much better bound, as the following lemma shows.

Lemma 5. *Let X be a ρ-cluster for some fixed constant $\rho \in (0,1)$. Then for sufficiently large t, and all positive $s \leq \sqrt{\rho}$, we have*

$$\Pr[\exists x \in X : \zeta \cdot x \geq t] \leq |X| \operatorname{poly}(t) N(t)^{1+(\sqrt{\rho}-s)^2/(1-\rho)} + 2N(st).$$

Proof. Suppose, w.l.o.g. that every $x \in K$ is of the form $x = \sqrt{\rho} x_0 + \sqrt{1-\rho} x'$ (if $x_0 \cdot x > \rho$ the following analysis would only be improved). Note that since $x' \cdot x_0 = 0$, the random projection $\zeta \cdot x_0$ is independent of all projections $\zeta \cdot x'$. Thus, we can bound $\Pr_\zeta[\exists x \in K : \zeta \cdot x \geq t]$ from above using a convolution on the random variables $\zeta \cdot x_0$ and $\max_{x \in K} \zeta \cdot x'$. In the following estimate the variable ξ represents $\zeta \cdot x_0$.

$$\begin{aligned}
&\Pr_\zeta[\exists x \in X : \zeta \cdot x \geq t] \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} \Pr\left[\exists x \in X : \zeta \cdot x' \geq \frac{t - \sqrt{\rho}\xi}{\sqrt{1-\rho}}\right] d\xi \\
&\leq 2 \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} \Pr\left[\exists x \in X : \zeta \cdot x' \geq \frac{t - \sqrt{\rho}\xi}{\sqrt{1-\rho}}\right] d\xi \\
&\leq 2 \int_0^{st} \frac{e^{-\xi^2/2}}{\sqrt{2\pi}} |X| N\left(\frac{t - \sqrt{\rho}\xi}{\sqrt{1-\rho}}\right) d\xi + 2 \int_{st}^{\infty} \frac{e^{-\xi^2/2}}{\sqrt{2\pi}} d\xi \qquad (12) \\
&\leq \operatorname{poly}(t) \cdot \max_{0 \leq a \leq s} |X| N(t)^{a^2 + (1-\sqrt{\rho}a)^2/(1-\rho)} + 2N(st) \qquad (13) \\
&= \operatorname{poly}(t) \cdot \max_{0 \leq a \leq s} |X| N(t)^{1+(\sqrt{\rho}-a)^2/(1-\rho)} + 2N(st) \\
&= |X| \cdot \operatorname{poly}(t) N(t)^{1+(\sqrt{\rho}-s)^2/(1-\rho)} + 2N(st)
\end{aligned}$$

Inequality (12) is a union bound and (13) follows from Lemma 4. □

3 Finding large independent sets in 3-uniform hypergraphs

We first review the algorithm and analysis given in [20]. Let us introduce the following notation: For all $t \in \{1, \ldots, \lceil \log n \rceil\}$, let $S_t \stackrel{\text{def}}{=} \{i \in V | t/\log n \leq \|v_i\|^2 < (t+1)/\log n\}$. Also, since $\|v_i\|^2 = v_\emptyset \cdot v_i$, we can write $v_i = (v_\emptyset \cdot v_i) v_\emptyset + \sqrt{v_\emptyset \cdot v_i (1 - v_\emptyset \cdot v_i)} u_i$, where u_i is a unit vector orthogonal to $v_\emptyset$. They show the following two lemmas, slightly rephrased here:

Lemma 6. *If the optimum of $KNS(H)$ is $\geq \gamma n$, there exists an index $t \geq \gamma \log n - 1$ s.t. $|S_t| = \Omega(n/\log^2 n)$.*

Lemma 7. *For index $t = \beta \log n$ and hyperedge $(i,j,k) \in E$ s.t. $i,j,k \in S_t$, constraint (6) implies*

$$\|u_i + u_j + u_k\|^2 \leq 3 + (3 - 6\beta)/(1-\beta) + O(1/\log n). \quad (14)$$

Using the above notation, we can now describe the rounding algorithm in [20], which is applied to the subhypergraph induced on S_t, where t is as in Lemma 6.

> **HIS-Round**$(H, \{u_i\}, r)$
>
> - Choose $\zeta \in \mathbb{R}^n$ from the n-dimensional standard normal distribution.
> - Let $V_\zeta(r) \stackrel{\text{def}}{=} \{i | \zeta \cdot u_i \geq r\}$. Remove all vertices in hyperedges fully contained in $V_\zeta(r)$, and return the remaining set.

The expected size of the remaining independent set can be bounded from below by $\mathbb{E}[|V_\zeta(r)|] - 3\mathbb{E}[|e \in E : e \subseteq V_\zeta(r)|]$, since each hyperedge contributes at most three vertices to $V_\zeta(r)$. If hyperedge (i,j,k) is fully contained in $V_\zeta(r)$, then by Lemma 7 we have $\zeta \cdot \frac{u_i+u_j+u_k}{\|u_i+u_j+u_k\|} \geq (3\sqrt{(1-\gamma)/(6-9\gamma)} - O(1/\log n))r$. By Lemma 4, and linearity of expectation, this means the size of the remaining independent set is at least

$$\tilde{\Omega}(N(r)n) - \tilde{O}(N(r)^{(3-3\gamma)/(2-3\gamma)}\,|E|).$$

Choosing r appropriately then yields the guarantee given in [20].

Theorem 8. *Given a 3-uniform hypergraph H on n vertices and m edges containing an independent set of size $\geq \gamma n$, one can find, in polynomial time, an independent set of size $\tilde{\Omega}(\min\{n, n^{3-3\gamma}/m^{2-3\gamma}\})$.*

Note that m can be as large as $\Omega(n^3)$, giving no non-trivial guarantee for $\gamma \leq \frac{1}{2}$. In fact, for $\gamma = \frac{1}{2}$, not only is there no non-trivial approximation guarantee, the integrality gap is $\Omega(n)$. To see this, note that taking $v_\emptyset, u_1, \ldots, u_n$ to be an orthonormal set, the vectors $v_i = \frac{1}{2}v_\emptyset + \frac{1}{2}u_i$ satisfy the constraints of KNS(H), with $\sum_i \|v_i\|^2 = \frac{n}{2}$. This solution is legal regardless of the underlying hypergraph.

To see why SDP hierarchies should be of some use in the case of $\gamma = \frac{1}{2}$, suppose that the SDP solution derives from some distribution on independent sets as discussed earlier. Let S be an independent set chosen according to this distribution. It is not hard to see that for $\gamma = \frac{1}{2}$, the only tight case in the above analysis is when $\|v_i\|^2 = \frac{1}{2}$ for every vertex $i \in V$, and $v_i \cdot v_j = v_i \cdot v_k = v_j \cdot v_k = \frac{1}{4}$ for every hyperedge $(i,j,k) \in E$. This means that every vertex is in S with probability $\frac{1}{2}$, and the vertices of any hyperedge (i,j,k) are chosen to be in S pairwise independently. Since all three vertices are never simultaneously in S, it must be the case that $k \in S$ precisely when exactly one of i,j is in S. In particular, this means that for $i,j,k,k' \in V$ such that $(i,j,k),(i,j,k') \in E$, the vertices k and k' are always in S at the same time, which implies $v_k = v'_k$.

Returning to the above analysis (for this particular setup), fix i,j, and let $\Gamma(\{i,j\}) = \{k | (i,j,k) \in E\}$. Since u_i, u_j, u_k are mutually orthogonal for $(i,j,k) \in E$, we have $\Pr[(i,j,k) \subseteq V_\zeta(r)] = N(r)^3$. Therefore, the expected number of hyperdges in $V_\zeta(r)$ containing i and j is $N(r)^3\,|\Gamma(\{i,j\})|$, which in the worst case might be $\Omega(N(r)^3 n)$. However since u_k are all the same vector for $k \in \Gamma(\{i,j\})$, vertices i,j only participate simultaneously in a hyperedge in $V_\zeta(r)$ with probability $N(r)^3$. Since there are at most $\binom{n}{2}$ vertex pairs, and each one contributes at most two verices to $V_\zeta(r)$, the expected number of *vertices participating in edges in* $V_\zeta(r)$ is at most $N(r)^3 n^2$ (possibly much less than the expected number of edges in $V_\zeta(r)$). Therefore, choosing r such that $N(r) = 1/\sqrt{2n}$, the rounding produces an independent set of expected size $\Omega(\sqrt{n})$.

Of course, this substantial improvement only occurs in the tight case of the previous analysis. Once we slightly relax the condition that all $\|v_i\|^2 = \frac{1}{2}$ or that for all edges (i,j,k) we have $v_i \cdot v_j = \frac{1}{4}$, we can no longer deduce that (fixing a particular pair (i,j)) the vectors v_k are all equal for $k \in \Gamma(\{i,j\})$. However, we can deduce that the vectors v_k should be highly clustered, using Lemma 2, and then we may use Lemma 5 to obtain an improvement in the analysis of the rounding algorithm.

We will now formalize this intuitive explanation. Our main result of this section is the following improvement in the integrality gap.

Theorem 9. *There is some fixed constant $\varepsilon > 0$ such that any 3-uniform hypergraph H on n vertices for which the optimum of MAX-IS$_3(H)$ is at least $(\frac{1}{2} - \varepsilon)n$ contains an independent set of size $\Omega(n^\varepsilon)$. Moreover, such an independent set can be found in polynomial time.*

Corollary 10. *For some fixed $\varepsilon > 0$, there is a polynomial time algorithm which, given an n-vertex 3-uniform hypergraph H containing an independent set of size $\geq (\frac{1}{2} - \varepsilon)n$, finds an independent set of size $\Omega(n^\varepsilon)$ in H.*

Proof of Theorem 9. Let $\{v_I | I \subseteq V, |I| \leq 3\}$ be a vector solution satisfying $IS_3(H)$ s.t. $\sum_i \|v_i\|^2 \geq (\frac{1}{2} - \varepsilon)$n. By Lemma 6, there is some some subset of vertices $S \subseteq V$ of size $\tilde{\Omega}(n)$ and some $\gamma \geq \frac{1}{2} - \varepsilon$ s.t. for all vertices $i \in S$, $|v_0 \cdot v_i - \gamma| \leq 1/\log n$. For the sake of simplicity, let us assume that $v_0 \cdot v_i = \gamma$ for all $i \in [n]$ (this will only affect the analysis by an additional polylogarithmic factor). Let r be such that $N(r) = n^{-(1-\varepsilon)}$. Note that if $\|u_i + u_j + u_k\|^2 \leq 3 - \delta$ then arguing as before $\Pr_\zeta[i,j,k \in V_\zeta(r)] = \tilde{O}(N(r)^{9/(3-\delta)})$. Let us define $E_\delta^- \stackrel{\text{def}}{=} \{(i,j,k) \in E |\, \|u_i + u_j + u_k\|^2 \leq 3 - \delta\}$, and $E_\delta^+ \stackrel{\text{def}}{=} E \setminus E_\delta^-$. When ε is sufficiently small, there is some $\delta = O(\varepsilon)$ such that $\mathbb{E}[|e \in E_\delta^- : e \subseteq V_\zeta(r)|] \leq \tilde{O}(N(r)^{9/(3-\delta)} n^3) = o(N(r)n)$. Therefore, we may assume that all hyperedges are in fact in E_δ^+. In particular, for every such hyperedge this implies $u_i \cdot u_j + u_i \cdot u_k + u_j \cdot u_k \geq -\delta/2$, and so, since $\gamma \geq \frac{1}{2} - \varepsilon$,

$$v_i \cdot v_j + v_i \cdot v_k + v_j \cdot v_k \geq 3\gamma^2 + (\gamma - \gamma^2)\frac{\delta}{2} \geq \frac{3}{2}\gamma - \eta, \quad (15)$$

for some $\eta = O(\delta + \varepsilon) = O(\varepsilon)$.

Now, fix $i, j \in [n]$, and let $k \in [n]$ be such that $(i,j,k) \in E_\delta^+$. Note that (by constraint (2)) $\|v_{\{i,-j\}}\|^2 = \|v_i - v_{ij}\|^2 = \gamma - v_i \cdot v_j$, and similarly $\|v_{\{j,-i\}}\|^2 = \gamma - v_i \cdot v_j$. Crucially, we also have $v_k \cdot v_{\{i,-j\}} = v_k \cdot (v_i - v_{ij}) = v_i \cdot v_k$ (since by constraint (3), $v_k \cdot v_{ij} = 0$), and similarly $v_k \cdot v_{\{j,-i\}} = v_j \cdot v_k$. Therefore, by Lemma 2 (letting $p_0 = v_\emptyset^2 = 1$, $q_0 = v_\emptyset \cdot v_k = \gamma$, $p_{0i} = \gamma - v_i \cdot v_k$, $p_{0i}q_{0i} = v_i \cdot v_k$, and similarly for p_{0j}, q_{0j}) there is some unit vector $x_0' \in \mathrm{Span}(\{v_I | I \subseteq \{i,j\}\})$ such that $v_\emptyset \cdot x_0' = 0$, and for all k as above,

$$x_0' \cdot \sqrt{\gamma - \gamma^2} u_k \geq \sqrt{\frac{(v_i \cdot v_k)^2}{\gamma - v_i \cdot v_j} + \frac{(v_j \cdot v_k)^2}{\gamma - v_i \cdot v_j} - \gamma^2}. \quad (16)$$

By (15), we have

$$\begin{aligned}(v_i \cdot v_k)^2 + (v_j \cdot v_k)^2 &\geq (v_i \cdot v_k + v_j \cdot v_k)^2/2 \\ &\geq (3\gamma/2 - v_i \cdot v_j)^2/2 - O(\varepsilon) \\ &= \gamma(\gamma - v_i \cdot v_j) \\ &\quad + (\gamma - 2v_i \cdot v_j)^2/8 - O(\varepsilon).\end{aligned}$$

Together with (16), this implies

$$x_0' \cdot u_k \geq \sqrt{1 + \frac{(\gamma - 2v_i \cdot v_j)^2 - O(\varepsilon)}{8(\gamma - \gamma^2)(\gamma - v_i \cdot v_j)}} \quad (17)$$

This implies that $\left|v_i \cdot v_j - \frac{\gamma}{2}\right| = O(\sqrt{\varepsilon})$ (since otherwise, we would have $x_0' \cdot u_k > 1$), which in turn implies that $|u_i \cdot u_j| = O(\sqrt{\varepsilon})$ (assuming $\left|\gamma - \frac{1}{2}\right| = O(\sqrt{\varepsilon})$). By symmetry, we also have $\max\{|u_i \cdot u_k|, |u_j \cdot u_k|\} = O(\sqrt{\varepsilon})$. Thus, $\|u_i + u_j + u_k\|^2 = 3 - O(\sqrt{\varepsilon})$, and letting $\tilde{u}_k = (u_i + u_j + u_k)/\|u_i + u_j + u_k\|$, the vectors $\{u_k\}$ are a $(1-\varepsilon')$-cluster for some $\varepsilon' = O(\sqrt{\varepsilon})$ (by (17)). Note that if (for hyperedge (i,j,k)) $i,j,k \in V_\zeta(r)$, then $\zeta \cdot \tilde{u}_k \geq \sqrt{3} - O(\sqrt{\varepsilon})$. Therefore, by Lemma 5 (with $s = \sqrt{1-\varepsilon'} - \sqrt{\varepsilon'/3}$), Lemma 4, and choice of r,

$$\begin{aligned}&\Pr_\zeta[\exists k : (i,j,k) \in V_\zeta(r)] \\ &\quad \leq \Pr_\zeta[\exists k : \zeta \cdot \tilde{u}_k \geq \sqrt{3} - O(\sqrt{\varepsilon})] \\ &\quad \leq n \cdot \mathrm{poly}(r) N((\sqrt{3} - O(\sqrt{\varepsilon}))r)^{1+\varepsilon'/(3\varepsilon')} \\ &\qquad + N((1 - O(\sqrt{\varepsilon'}))(\sqrt{3} - O(\sqrt{\varepsilon}))r) \\ &\quad = \mathrm{poly}(r) N(r)^{-1/(1-\varepsilon)} N((\sqrt{3} - O(\sqrt{\varepsilon}))r)^{4/3} \\ &\qquad + N((\sqrt{3} - O(\sqrt{\varepsilon'}))r) \\ &\quad = N(r)^{3-O(\sqrt{\varepsilon'})} = N(r)^{3-O(\sqrt[4]{\varepsilon})}.\end{aligned}$$

Therefore, the expected number of vertices participating in edges contained in $V_\zeta(r)$ is at most $n^2 N(r)^{3-O(\sqrt[4]{\varepsilon})} = o(N(r)n)$ for sufficiently small $\varepsilon > 0$, and the theorem follows.

□

4 Coloring 3-colorable graphs

As is standard, we assume that in order to find colorings with $\tilde{O}(f(n))$ colors, it suffices to find independent sets of size $n/f(n)$.

We concentrate only on the case where there is a bound Δ on the maximum degree (see below). As was discussed in [3], one can use the technique of [8] to obtain an algorithm with approximation guarantee in terms of n:

Theorem 11. *Let $\mathcal{A}$ be a polynomial time algorithm that takes an n-vertex 3-colorable graph with maximum degree at most Δ as input, and returns an independent set of size $\geq n/f(n,\Delta)$ (where f is monotonically increasing in n and Δ). Then there is a polynomial time algorithm which, for any n-vertex 3-colorable graph, finds an $\tilde{O}(\min_{1 \leq \Delta \leq n}(f(n/4, 2\Delta) + (n/\Delta)^{3/5}))$ coloring.*

The KMS rounding algorithm is as follows:

KMS$(G, \{u_i\}, r)$

- Choose $\zeta \in \mathbb{R}^n$ from the n-dimensional standard normal distribution.
- Let $V_\zeta(r) \stackrel{\mathrm{def}}{=} \{i | \zeta \cdot u_i \geq r\}$. Return all $i \in V_\zeta(r)$ with no neighbors in $V_\zeta(r)$.

Theorem 12 (KMS). *For any graph G on n vertices with maximum degree $\leq \Delta$ and vector 3-coloring $\{u_i\}$ of G, there exists some $r = r(n,\Delta) > 0$ such that the expected size of the independent set returned by algorithm KMS $(G, \{u_i\}, r)$ is $\tilde{\Omega}(\Delta^{-1/3} \cdot n)$.*

To describe our algorithm, we need one more piece of notation. Given a solution $\{v_I\}$ of $3COL_3(G)$, and vertices $i, k \in V$ s.t. $v_{(i,R),(k,R)} \neq 0$, define w_{ik} to be the unit vector satisfying

$$\begin{aligned}v_{(i,R),(k,R)} =& (\|v_{(i,R),(k,R)}\| / \|v_{(i,R)}\|)^2 v_{(i,R)} \\ &+ \sqrt{\|v_{(i,R),(k,R)}\|^2 - \frac{\|v_{(i,R),(k,R)}\|^4}{\|v_{(i,R)}\|^2}}\, w_{ik}.\end{aligned} \quad (18)$$

By (2), $\|v_{(i,R),(k,R)}\|^2 = v_{(i,R),(k,R)} \cdot v_{(i,R)}$, hence such a vector exists, and is orthogonal to $v_{(i,R)}$.

Our algorithm is as follows:

KMS$_2(G)$

1. Solve the SDP $3COL_3(G)$ to obtain vectors $\{v_I\}$.
2. For "all" $r > 0$,
   - Choose $\zeta \in \mathbb{R}^n$ from the n-dimensional standard normal distribution.
   - Let $V_\zeta(r) \stackrel{\text{def}}{=} \{i | \zeta \cdot u_i \geq r\}$ (u_i is as in (10)). Pick any edge (i, j) with both endpoints in $V_\zeta(r)$, and eliminate both i and j. Repeat until no such edges are left. Let $V'_\zeta(r)$ be the remaining independent set.
3. For all i, let
   $V_i = \left\{k | \left\|v_{(i,R),(k,R)}\right\|^2 > \frac{1}{6} - 1/\sqrt{\log\log n}\right\}$,
   and obtain independent set W_i from
   KMS$(V_i, \{w_{ik} | k \in V_i\}, \sqrt{\log n}/(\log\log n)^{1/4})$.
4. Output the largest set among $V'_\zeta(r)$, W_i.

Remark 13. Equivalently, in step 2 we can first choose ζ, and then enumerate over all relevant values of r (that is, over $r_i = \zeta \cdot u_i$). However, for the purposes of the analysis, we consider the first formulation.

Note that in KMS$_2$, the set $V'_\zeta(r)$contains the set returned by KMS(G, r), so Theorem 12 holds also for KMS$_2$. Step 2 is the rounding algorithm KMS$'$ proposed in [3]. As mentioned earlier, it was shown in [16] that for vector 3-colorable graphs on n vertices with maximum degree Δ, KMS produces an independent set of size $\tilde{\Omega}(n\Delta^{-1/3})$ (and thus an $\tilde{O}(\Delta^{1/3})$-coloring). We quantify the improvement in our analysis of KMS$_2$ using the following definition from [3].

Definition 14. Given a graph G with maximum degree Δ, the parameter $r > 0$ is at most *c-inefficient* if $\Delta \leq N(\sqrt{3}r)^{-(1+c)}$.

Note that by Lemma 4, if $r > 0$ is exactly c-inefficient, then $N(r) = \tilde{\Theta}(\Delta^{-\frac{1}{3+3c}})$. Thus, our objective is to find the largest possible $c = c(\Delta)$ for which KMS$_2$ is guaranteed to return an independent set of size $\Omega(N(r) \cdot n)$ for a c-inefficient threshold r. Using this terminology, we give the following explicit guarantee for the performance of KMS$_2$.

Theorem 15. *For every $\tau > \frac{6}{11}$ there exists $c_1(\tau) > 0$ such that for $0 < c < c_1(\tau)$, and any n vertex graph G with maximum degree $\leq n^\tau$, if the parameter r is (at most) c-inefficient for G, then KMS$_2(G)$ returns an independent set of size $\Omega(N(r)n)$.*

Furthermore, $c_1(\tau)$ satisfies

$$c_1(\tau) \stackrel{\text{def}}{=} \min\left\{\frac{1}{2}, \sup\left\{c \,\middle|\, \min_{0 \leq \alpha \leq \frac{c}{1+c}} \lambda_c(\alpha) > 0\right\}\right\}, \quad (19)$$

where $\lambda_c(\alpha) \stackrel{\text{def}}{=} 7/3 + c + \alpha^2/(1-\alpha^2) - (1+c)/(\tau) - \left(\sqrt{(1+\alpha)/2} + \sqrt{c(1-\alpha)/2}\right)^2$.

Corollary 16. *For any n-vertex 3-colorable graph G with maximum degree $\leq \Delta = n^{0.6546}$, KMS$_2(G)$ returns an independent set of size $\Omega(\Delta^{-0.3166} \cdot n)$.*

Together with Theorem 11, this proves the following.

Theorem 17. *For 3-colorable graphs, one can find an $O(n^{0.2072})$ coloring in polynomial time.*

4.1 Overview of analysis

In this section we give an informal description of the analysis of KMS$_2$, which will be formalized in the next section. Following [3], our analysis will focus on the vectors u'_{ij}, which, for every edge $(i, j) \in E$, are defined to be the unit vectors satisfying

$$u_j = -\tfrac{1}{2}u_i + \tfrac{\sqrt{3}}{2}u'_{ij}. \quad (20)$$

Since $u_i \cdot u_j = -\frac{1}{2}$, we have $u'_{ij} \cdot u_i = 0$. The following lemma from [3] (adapted to the above notation) relates these vectors to the performance of the rounding algorithm.

Lemma 18. *For all nodes $i \in V$,*

$$\Pr[i \textit{ is eliminated } | i \in V_\zeta(r)] \leq \Pr[\exists j : \zeta \cdot u'_{ij} \geq \sqrt{3}r].$$

Our goal is to show that either the probability on the right is small for many vectors (and thus, in expectation an $\Omega(N(r))$-fraction of them will be in $V'_\zeta(r)$), or we can extract a large independent set from the 2-neighborhood $\Gamma(\Gamma(i))$ of some vertex i (The second part is covered by step 3 in KMS$_2$). For the purposes of the current discussion, we will make a few simplifying assumptions. First, we assume that the SDP solution corresponds to a distribution over legal 3-colorings. Let col$(\cdot)$ be a random assignment of 3-colorings chosen according to this distribution. Then, for example, $u_i \cdot u_j = \Pr_{\text{col}}[\text{col}(i) = \text{col}(j)] - \frac{1}{2}\Pr_{\text{col}}[\text{col}(i) \neq \text{col}(j)]$. Secondly, we assume that the vectors do not display any statistically significant behavior other than the above constraints. This roughly corresponds to the case where the parameter r is chosen such that $N(r) \approx \Delta^{-1/3}$, and step 2 fails (in fact, we make the stronger assumption that the right-hand-side of the inequality in Lemma 18 is large for all $i \in V$).

We would first like to show that joint neighborhoods (intersections of two neighborhoods) are clustered. Consider two vertices $j, j' \in V$ which participate in some joint neighborhood. Conditioning on the choice of color col(j), any neighbor of j is assigned a random choice of one of the two remaining colors. Thus, our assumption about statistically significant behavior implies that for any two distinct neighbors $i, k \in \Gamma(j)$ (and similarly for $i, k \in \Gamma(j')$),

$\Pr[\mathrm{col}(i) = \mathrm{col}(k)] \approx \frac{1}{2}$. Now consider i and k as fixed, and think of j, j' as a random pair of vertices in $\Gamma(i) \cap \Gamma(k)$. Then $\mathrm{col}(j) = \mathrm{col}(j')$ whenever $\mathrm{col}(i) \neq \mathrm{col}(k)$ (since in a legal 3-coloring, the joint neighborhood of two distinctly colored vertices must be monochromatic). On the other hand, conditioning on the event $\mathrm{col}(i) = \mathrm{col}(k) = C$ for some color C, we have $\Pr[\mathrm{col}(j) = \mathrm{col}(j') | \mathrm{col}(i) = \mathrm{col}(k) = C] \geq \frac{1}{2} - o(1)$ for many pairs $j, j' \in \Gamma(i) \cap \Gamma(k)$ (see the discussion preceeding Lemma 2). Summarizing, for such pairs we have

$$\begin{aligned}\Pr[\mathrm{col}(j) = \mathrm{col}(j')] \geq \Pr[\mathrm{col}(i) \neq \mathrm{col}(k)] \\ + (1/2 - o(1))\Pr[\mathrm{col}(i) = \mathrm{col}(k)] \\ \approx \tfrac{1}{2} + \tfrac{1}{2}(\tfrac{1}{2} - o(1)) = \tfrac{3}{4} - o(1).\end{aligned}$$

Now, by definition of $\mathrm{col}(\cdot)$, this implies that $u_j \cdot u_{j'} = \Pr[\mathrm{col}(j) = \mathrm{col}(j')] - \frac{1}{2}\Pr[\mathrm{col}(j) \neq \mathrm{col}(j')] \geq \frac{5}{8} - o(1)$. This, in turn, implies $u'_{ij} \cdot u'_{ij'} \geq \frac{1}{2} - o(1)$, so the vectors $\{u'_{ij} | j \in \Gamma(i) \cap \Gamma(k)\}$ form a $1/2 - cluster$. This intuition can be formalized using Lemma 2. The cardinality of such clusters must be small, since otherwise, by the bound in Lemma 5, they would have a disproportionately small contribution to the probability in Lemma 18. This is made precise in Lemma 24, which in this case implies that for $i, k \in V$ as above, $|\Gamma(i) \cap \Gamma(k)| \leq \sqrt{\Delta}$.

This suffices to show that the number of vertices at distance 2 from i is large. Indeed,

$$\begin{aligned}\Delta^2 = |\{(j,k) \in E | j \in \Gamma(i)\}| = \textstyle\sum_{k \in \Gamma(\Gamma(i))} |\Gamma(i) \cap \Gamma(k)| \\ \leq |\Gamma(\Gamma(i))| \sqrt{\Delta},\end{aligned}$$

and thus $|\Gamma(\Gamma(i))| \geq \Delta^{3/2}$. On the other hand, as we mentioned earlier, for most $k \in \Gamma(\Gamma(i))$, $\Pr[\mathrm{col}(i) = \mathrm{col}(k)] \approx \frac{1}{2}$. Thus the expected number of vertices in $\Gamma(\Gamma(i))$ with the same color as i is $\frac{1}{2} |\Gamma(\Gamma(i))|$. In particular, the set $\Gamma(\Gamma(i))$ contains an independent set which is nearly half of all its vertices. In this case we can use any of a number of Vertex Cover approximations to extract an independent set of size $\tilde{\Omega}(|\Gamma(\Gamma(i))|) = \tilde{\Omega}(\Delta^{3/2})$. This gives the following trade-off: For r s.t. $N(r) \approx \Delta^{-1/3}$, either step 2 produces an independent set of size $N(r)n \approx \Delta^{-1/3}n$, or step 3 produces an independent set of size $\tilde{\Omega}(\Delta^{3/2})$.

Slightly relaxing the above argument (by decreasing r, hence increasing the size of the independent set when step (2) succeeds), gives a better trade-off in the worst case, as long as $\Delta^{-1/3}n < \Delta^{3/2}$, i.e. $\Delta > n^{6/11}$. However, decreasing r introduces error-terms at every step of the argument, possibly decreasing the guaranteed size of $\Gamma(\Gamma(i))$. The subtle trade-off between these two parameters is the main focus of the analysis.

4.2 Analysis of current improvement

In this section we prove Theorem 15. The goal of the analysis is to show that if KMS_2 does not find a large independent set in step (2), then one of the sets V_i is large. We first note that this is sufficient.

Lemma 19. *Let V_i be as in algorithm KMS$_2$. Then KMS$(V_i, \{w_{ik} | k \in V_i\}, \sqrt{\log n}/(\log\log n)^{1/4})$ returns an independent set of size $\Omega(|V_i| N(\sqrt{\log n}/(\log\log n)^{\frac{1}{4}})) = \tilde{\Omega}(|V_i| N(n^{-1/(2\sqrt{\log\log n})}))$.*

Proof. For any $k, k' \in V_i$ s.t. $(k, k') \in E$ we have $v_{(i,R),(k,R)} \cdot v_{(i,R),(k',R)} = 0$, and hence by equation (18), $w_{ik} \cdot w_{ik'} = -\|v_{(i,R),(k,R)}\|^2 / (\|v_{(i,R)}\|^2 - \|v_{(i,R),(k,R)}\|^2) < -1 + O(1/\sqrt{\log\log n})$. In particular, $\|w_{ik} + w_{ik'}\|^2 = O(1/\sqrt{\log\log n})$. Thus, for $r = \sqrt{\log n}/(\log\log n)^{1/4}$, the probability that both $k, k' \in V_\zeta(r)$ is at most $\Pr_\zeta[\zeta \cdot (w_{ik} + w_{ik'}) \geq 2r] = N(2r/\|w_{ik} + w_{ik'}\|) = o(N(r)/n^2)$, where the final equality follows from Lemma 4. In particular, the expected number of edges contained in $V_\zeta(r)$ is at most $o(N(r))$, whereas the expected number of vertices is $\Omega(|V_i| N(r))$. □

The following theorem, together with Lemma 19 immediately implies Theorem 15.

Theorem 20. *For every $\tau > \frac{6}{11}$ and $0 < c < c_1(\tau)$, there exists some $\varepsilon = \varepsilon(\tau, c) > 0$ s.t. for sufficiently large n, any n vertex graph G with max degree $\leq n^\tau$, and r s.t. $N(\sqrt{3}r)^{-(1+c)}$, either*

1. *Step (2) of KMS$_2(G)$ returns an independent set of expected size $\Omega(N(r)n)$, or*

2. *There exists some vertex i for which $|V_i| \geq (N(r)n^{1+\varepsilon})$.*

The rest of this section is devoted to proving Theorem 20. We will use the following definition from [5] and [3].

Definition 21. A set of unit vectors $\{x_1, \ldots, x_k\}$ is said to be a *(t, δ)-cover*, if for $\zeta \in \mathbb{R}^n$ chosen from the standard normal distribution, $\Pr[\exists i : \zeta \cdot x_i \geq t] \geq \delta$.
The cover $\{x_1, \ldots, x_k\}$ is said to be (at most) *c-inefficient*, if $k \leq N(t)^{-(1+c)}$.

To motivate the above definition, we note that, by lemma 18, for any vertex i for which $\Pr_\zeta[i \in V'_\zeta(r)] \leq \frac{1}{2}N(r)$, we have a $(\frac{1}{2}, \sqrt{3}r)$-cover $\{u'_{ij}\}_j$, and moreover, this cover is at most c-inefficient if the parameter r is c-inefficient for G. We further refine the above definition as follows:

Definition 22. A set of unit vectors $\{x_1, \ldots, x_k\}$ is said to be a *uniformly c-inefficient* (t, δ)-cover, if $k \geq \delta N(t)^{-(1+c)}$, and every subset $S \subseteq [k]$ is itself a $(t, N(t)^{1+c} |S|)$-cover.

Using this definition, we will show that every cover which has bounded inefficiency, contains a large core which has bounded *uniform* inefficiency.

Lemma 23. *Let X be a c-inefficient (t, δ)-cover. Then*

1. *For some $0 \leq b \leq c + O(\ln(1/\delta)/t^2)$, there exists a subset $X' \subseteq X$ which is a uniformly b-inefficient $(t, \Omega(\delta/t^2))$-cover.*
2. *If, in addition, X is a ρ-cluster and $\delta = \Omega(1/\mathrm{poly}(t))$, then $b \geq \rho/(1-\rho) - \tilde{O}(1/t)$.*

Proof. We assign to the elements in X some additive measure $\mu(\cdot)$ s.t. $\mu(X) \geq \delta$ and every subset $S \subset X$ is a $(t, \mu(S))$-cover, i.e. $\Pr_\zeta[\exists x \in S : \zeta \cdot x \geq t] \geq \mu(S)$. A natural choice is given by $\mu(x) \stackrel{\text{def}}{=} \Pr_\zeta[\zeta \cdot x \geq t \text{ and } \zeta \cdot x = \max_{x' \in X} \zeta \cdot x']$. Let $X_+ = \{x | \mu(x) > \delta N(t)^{1+c}/2\}$, and $X_- = X \setminus X_+$. Then, by the efficiency and cover properties of X, we have

$$\delta \leq \mu(X) = \mu(X_-) + \mu(X_+) \leq \tfrac{1}{2} |X| \delta N(t)^{1+c} + \mu(X_+) \leq \delta/2 + \mu(X_+).$$

Thus, $\mu(X_+) \geq \delta/2$, and, by Lemma 4 and definition of X_+, for every $x \in X_+$, $\mu(x) = N(t)^{1+b_x}$ for some $b_x \in [0, c + O(\ln(1/\delta)/t^2]$. Divide this range into t^2 subintervals I_i of length $(c + O(\ln(1/\delta)/t^2))/t^2$, and divide X_+ into bins accordingly, so that $x \in X_i$ iff $b_x \in I_i$. Thus, some such bin must have measure $\mu(X_i) = \Omega(\delta/t^2)$. This X_i satisfies the required properties in part (1), where the lower bound on $|X_i|$ follows immediately from the upper bound on $\mu(x)$ for all $x \in X_i$.

For part (2), let $I_i = [b_1, b_2]$ be the interval chosen above, and X_i the corresponding subset of X_+. First, note that, by definition of X_i, $N(t)^{1+b_2} |X_i| \leq \mu(X_i) \leq 1$, and thus $|X_i| \leq N(t)^{-(1+b_2)}$. Let s be such that $N(st) = o(\delta/t^2)$. By Lemma 4 there is some $s = O(\sqrt{\log(t^2/\delta)}/t) = O(\sqrt{\log t}/t)$ satisfying this property. Therefore, by Lemma 5, we have

$$\delta/t^2 \leq \mu(X_i) \leq \mathrm{poly}(t) N(t)^{-(1+b_2)+1+\rho/(1-\rho)-O(s)} + o(\delta/t^2).$$

And so the desired lower bound on b follows, since

$$N(t)^{-b_2+\rho/(1-\rho)-O(\sqrt{\log t}/t)} \geq \frac{1}{\mathrm{poly}(t)} = N(t)^{O(\log t/t^2)}.$$

□

We now show that uniformly efficient covers of cardinality k do not contain ρ-clusters significantly larger than $k^{1-\rho}$.

Lemma 24. *Let X be a uniformly b-inefficient (t, δ)-cover, then for all $\rho \geq b/(1+b)$ any ρ-cluster in X has cardinality at most $O(\mathrm{poly}(t) N(t)^{-(\sqrt{1-\rho}+\sqrt{b\rho})^2})$.*

Proof. Let $K \subset X$ be a ρ-cluster of cardinality $N(t)^{-\beta}$, and let $s = \sqrt{\rho} - \sqrt{b(1-\rho) - \eta}$ for some $\eta = o(1)$ (specified later). Then by Lemma 5 (with the above choice of s), we have

$$\Pr_\zeta[\exists x \in K : \zeta \cdot x \geq t] \leq \mathrm{poly}(t) N(t)^{-\beta+1+b+O(\eta)} + 2N(st).$$

Thus, by the uniform b-inefficiency assumption, and by Lemma 4, for some $\eta = O(\sqrt{\log t}/t^2)$, we have

$$\begin{aligned} N(t)^{-\beta+1+b} &\leq \Pr_\zeta[\exists x \in K : \zeta \cdot x \geq t] \\ &\leq o(N(t)^{-\beta+1+b}) \\ &\quad + \mathrm{poly}(t) N(t)^{(\sqrt{\rho}-\sqrt{b(1-\rho)})^2}. \end{aligned}$$

Hence, the required bound on $|K| = N(t)^{-\beta}$ follows immediately. □

We are now ready to prove Theorem 20.

Proof of Theorem 20. If for at least $n/2$ vertices $i \in V$ we have $\Pr_\zeta[i \in V'_\zeta(r)] \geq \frac{1}{2} N(r)$, then clearly, by linearity of expectation, we find an independent set of expected size $\Omega(N(r)n)$ in step (2). Assume this is not the case. Then we can prune as in Lemma 25. Let $G' = (V', E')$ be the remaining graph, and fix some vertex $i \in V'$. Then, by Lemma 25, we have that $\{u'_{ij} | j \in \Gamma_{G'}(i)\}$ and and the sets $\{u'_{jk} | k \in \Gamma_{G'}(j)\}$ for every $j \in \Gamma_{G'}(i)$ are all uniform $\left(\sqrt{3}r, \frac{1}{8} - O\left(\frac{1}{\log r}\right)\right)$-covers which are at most c-inefficient. Moreover, there exists some constant $C > 0$ such that the sets $U_{ji} = \left\{u'_{jk} \;\middle|\; -\frac{C}{r^2} \leq u'_{ji} \cdot u'_{jk} \leq \frac{c}{1+c} + \frac{C}{\log r}\right\}$ (for every $j \in \Gamma_{G'}(i)$) are $\left(\sqrt{3}r, \Omega(\frac{1}{r^3})\right)$-covers. Note that for all such j, $\{k | u'_{jk} \in U_{ji}\} \subseteq V_i$. Therefore, we need to give a lower bound on $\left|\bigcup_{j \in \Gamma_{G'}(i)} \{k | u'_{jk} \in U_{ji}\}\right|$.

Now, subdivide the interval $[-\frac{C}{r^2}, \frac{c}{1+c} + \frac{C}{\log r}]$ into $O(r)$ subintervals I_l of length $1/r$. For every $j \in \Gamma_{G'}(i)$ there is some $l = l(j)$ such that the set $\{u'_{jk} | u'_{ji} \cdot u'_{jk} \in I_{l(j)}\}$ is a $(\sqrt{3}r, \Omega(1/r^4))$-cover. Moreover, there is some l_0 such that the set $U'_i = \{u'_{ij} | l(j) = l_0\}$ is a $(\sqrt{3}r, \Omega(1/r))$-cover. Let α be such that $I_{l_0} = [\alpha, \alpha + 1/r)$. By Lemma 23, there is some subset $U''_i \subseteq U'_i$ which is a uniformly b-inefficient $(\sqrt{3}r, \Omega(1/r^3))$-cover for some $0 \leq b \leq c + o(1)$. Similarly, for every $j \in U''_i$, there is some set $W_j \subset \{u'_{jk} | u'_{ji} \cdot u'_{jk} \in I_{l(0)}\}$ which is a uniformly a_j-inefficient $(\sqrt{3}r, \Omega(1/r^5))$-cover, where $a_j \geq \alpha^2/(1-\alpha^2) - o(1)$ (since the sets $\{u'_{jk} | u'_{ji} \cdot u'_{jk} \in I_{l_0}\}$ are α^2-clusters for all $j \in \Gamma_{G'}(i)$). Let us summarize the situation:

1. U''_i is a uniformly b-inefficient $(\sqrt{3}r, \Omega(1/r^3))$-cover for some $0 \leq b \leq c + o(1)$.

2. $\forall j \in U_i''$, W_j is a uniformly a_j-inefficient $(\sqrt{3}r, \Omega(1/r^5))$-cover for some $a_j \geq \frac{\alpha^2}{1-\alpha^2} - o(1)$.

3. $\forall j, k$ s.t. $u'_{jk} \in W_j$, $v_{(i,R)} \cdot v_{(k,R)} \in [(1+\alpha)/6, (1+\alpha)/6 + o(1)]$.

4. $-o(1) \leq \alpha \leq c/(1+c) + o(1)$.

Property 3 follows easily from the definition of $\{u'_{jk}\}$. However, for the sake of simplicity, let us assume that for all such j, k, $\left\|v_{(i,R),(k,R)}\right\|^2 = v_{(i,R)} \cdot v_{(k,R)} = \frac{1+\alpha}{6}$, as the error term will have a negligible effect. By constraint (9), this also implies $\left\|v_{(i,B),(k,B)}\right\|^2 = \left\|v_{(i,Y),(k,Y)}\right\|^2 = (1+\alpha)/6$. Moreover, since (as can be easily checked) $\left\|v_{(i,B)}\right\|^2 = \sum_{C \in R,B,Y} \left\|v_{(i,B),(k,C)}\right\|^2$, we have (again by (9)), $\left\|v_{(i,B),(k,Y)}\right\|^2 = \left\|v_{(i,Y),(k,B)}\right\|^2 = (1-\alpha)/12$. Furthermore, for $j \in \Gamma(i) \cap \Gamma(k)$ and $(C_1, C_2) \in \{(B,Y),(Y,B)\}$, we have $v_{(i,C_1),(k,C_1)} \cdot v_{(j,R)} = \frac{1}{2}\left\|v_{(i,C_1),(k,C_1)}\right\|^2$, and $v_{(i,C_1),(k,C_2)} \cdot v_{(j,R)} = \left\|v_{(i,C_1),(k,C_2)}\right\|^2$. Finally, we note that for all $(i,j) \in E$, $v_{(j,R)} = \frac{1}{2}(v_{(i,B)} + v_{(i,Y)}) + \frac{1}{\sqrt{6}} u'_{ij}$ (by definition of u_i, u'_{ij}, and by constraint (8)). We now fix $i, k \in [n]$ as above (i.e. $v_{(i,R)} \cdot v_{(k,R)} = (1+\alpha)/6$), and apply Lemma 2, where for all $C_1, C_2 \in \{B, Y\}$, we let $p_{C_1} = \left\|v_{(i,C_1)}\right\|^2 = 1/3$, $p_{C_1} q_{C_1} = v_{(i,C_1)} \cdot v_{(j,R)}$, $p_{C_1,C_2} = \left\|v_{(i,C_1),(k,C_2)}\right\|^2$, and $p_{C_1,C_2} q_{C_1,C_2} = v_{(i,C_1),(k,C_2)} \cdot v_{(j,R)}$. Thus, there is some unit vector $x'_0 \in \text{Span}(\{v_I | I \subseteq \{i,k\} \times \{B,Y\}\})$ such that

$$\begin{aligned} x'_0 \cdot \tfrac{1}{\sqrt{6}} u'_{ij} &= \sqrt{2 \cdot \tfrac{1+\alpha}{6} \cdot \left(\tfrac{1}{2}\right)^2 + 2 \cdot \tfrac{1-\alpha}{12} - 2 \cdot \tfrac{1}{3} \cdot \left(\tfrac{1}{2}\right)^2} \\ &= \sqrt{(1-\alpha)/12}. \end{aligned}$$

Thus, for all k, the set $\{u'_{ij} \in U_i'' | u'_{jk} \in W_j\}$ is in fact a $(1-\alpha)/2$-cluster, and so by property 1 above and Lemma 24, we have $\left|\{u'_{ij} \in U_i'' | u'_{jk} \in W_j\}\right| \leq N(\sqrt{3}r)^{-(\sqrt{(1+\alpha)/2} + \sqrt{b(1-\alpha)/2})^2 - o(1)}$. Hence,

$$\begin{aligned} \sum_{j : u'_{ij} \in U_i''} |W_j| &= \left|\{(j,k) | u'_{ij} \in U_i'' \text{ and } u'_{jk} \in W_j\}\right| \\ &\leq \Big|\bigcup_{j \in U_i''} \{k | u'_{jk} \in W_j\}\Big| \\ &\quad \times N(\sqrt{3}r)^{-\left(\sqrt{(1+\alpha)/2} + \sqrt{b(1-\alpha)/2}\right)^2 - o(1)} \end{aligned}$$

Yet, by properties 1 and 2, $\sum_{j : u'_{ij} \in U_i''} |W_j| \geq N(\sqrt{3}r)^{-(1+b)-(1+\alpha^2/(1-\alpha^2))+o(1)}$.

Thus we have shown

$$\Big|\bigcup_{j \in U_i''} \{k | u'_{jk} \in W_j\}\Big| \geq N(\sqrt{3}r)^{-\left(\lambda_b(\alpha) + \frac{1}{3} - (1+b)/\tau\right) + o(1)}. \tag{21}$$

As a final simplification, we note that the function above is monotonically decreasing in b for all $b \leq (1-\alpha)/(1+\alpha)$. This is consistent with the range of b (up to $o(1)$), since the fact that $c \leq \frac{1}{2}$ and property 4 imply $(1-\alpha)/(1+\alpha) \geq c$. Therefore, w.l.o.g. $b = c$. Substituting $b = c$ in (21), by our choice of c, and the inefficiency of r, we have that for some constant $\varepsilon' > 0$, $\left|\bigcup_{j \in U_i''} \{k | u'_{jk} \in W_j\}\right| \geq N(\sqrt{3}r)^{1/3 - \varepsilon' + o(1)} n$. □

Acknowledgements

The author wishes to thank Sanjeev Arora, Moses Charikar and Venkatesan Guruswami for helpful conversations.

The author is especially grateful to Elad Verbin for discussions on graph coloring, and to Gyanit Singh, for discussion on hypergraph independent sets, and with whom related work on hypergraph coloring was a direct inspiration for this paper.

References

[1] M. Alekhnovich, S. Arora, and I. Tourlakis. *Towards strong nonapproximability results in the Lovász-Schrijver hierarchy.* In Proceedings of 37th ACM STOC, pp. 294–303, 2005.

[2] S. Arora, B. Bollobás, L. Lovász, and I. Tourlakis. *Proving integrality gaps without knowing the linear program.* Theory of Computing, vol. 2, pp. 19-51, Feb. 2006.

[3] S. Arora, E. Chlamtac, and M. Charikar. *New approximation guarantee for chromatic number.* In Proceedings of 38th ACM STOC, pp. 215-224, 2006.

[4] S. Arora, J. R. Lee, and A. Naor. *Euclidean distortion and the sparsest cut.* In Proceedings of the 37th ACM STOC, pp. 553-562, 2005.

[5] S. Arora, S. Rao, and U. Vazirani. *Expander flows, geometric embeddings and graph partitioning.* In Proceedings of the 36th ACM STOC, pp. 222–231, 2004.

[6] A. Barvinok. *Optimization of polynomials on the unit sphere.*

[7] A. Blum. *New approximations for graph coloring.* Journal of the ACM, 41(3):470–516, 1994.

[8] A. Blum and D. Karger. *An $\tilde{O}(n^{3/14})$-coloring algorithm for 3-colorable graphs.* Information Processing Letters, 61(1):49–53, 1997.

[9] E. Chlamtac, G. Singh. Manuscript in preparation.

[10] M. Charikar, K. Makarychev, and Y. Makarychev. *Near-Optimal Algorithms for Unique Games.* In Proceedings of the 38th ACM STOC, pp. 205-214, 2006.

[11] H. Chen, F. Frieze *Coloring bipartite hypergraphs.* In Proceedings of the 5th IPCO Conference, pp. 345–358, 1996.

[12] E. De Klerk, M. Laurent, and P. Parrilo. *A PTAS for the minimization of polynomials of fixed degree over the simplex.* Theoretical Computer Science, vol. 361, issues 2-3, pp. 210-225, Sep. 2006.

[13] K. Georgiou, A. Magen, T. Pitassi, I. Tourlakis. *Integrality gaps of $2-o(1)$ for vertex cover in the Lovász-Schrijver hierarchy.* In Proceedings of 48th IEEE FOCS, 2007.

[14] M. Goemans and D. Williamson. *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming.* Journal of the ACM, vol. 42, no. 6, pp. 1115–1145, Nov. 1995.

[15] J. Håstad. Some optimal inapproximability results. Journal of the ACM, vol. 48, no. 4, pp. 798–859, July 2001.

[16] D. Karger, R. Motwani, and M. Sudan. *Approximate graph coloring by semidefinite programming.* Journal of the ACM, 45(2):246–265, 1998.

[17] H. Karloff and U. Zwick. *A 7/8-approximation algorithm for max 3SAT?* In Proceedings of 38th IEEE FOCS, pages 406–415, 1997.

[18] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. *Optimal inapproximability results for MAX-CUT and other 2-variable CSPs?* Proceedings of the 45th IEEE FOCS, pp. 146-154, 2004.

[19] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. *Optimal inapproximability results for MAX-CUT and other 2-variable CSPs?* ECCC Report TR05-101, 2005.

[20] M. Krivelevich, R. Nathaniel, and B. Sudakov. *Approximating coloring and maximum independent sets in 3-uniform hypergraphs.* Journal of Algorithms, vol. 41, no. 1, pp. 99–113, Oct. 2001.

[21] M. Laurent. *A comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre relaxations for 0-1 programming.* Mathematics of Operations Research, vol. 28, no. 3, pp. 460-496, Aug. 2003.

[22] J.B. Lasserre. *An explicit exact SDP relaxation for nonlinear 0-1 programs.* Lecture Notes In Computer Science, vol. 2081, pp. 293–303, 2001.

[23] L. Lovász and A. Schrijver. *Cones of matrices and set-functioins and 0-1 optimization.* SIAM Journal on Optimization, vol. 1, no. 2, pp. 166-190, May 1991.

[24] J. Nie, M. Schweighofer. *On the complexity of Putinar's Positivstellensatz.* Journal of Complexity, vol. 23, no. 1, pp. 135–150, Feb. 2007.

[25] G. Schoenebeck, L. Trevisan, and M. Tulsiani. *Tight integrality gaps for Lovász-Schrijver LP relaxations of Vertex Cover and Max Cut.* In Proceedings of 29th ACM STOC, pp. 302-310, 2007.

[26] H.D. Sherali and W.P. Adams. *A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems.* SIAM Journal on Discrete Mathematics, vol. 3, issue 3, pp. 411–430, Aug. 1990.

[27] A. Wigderson. *Improving the performance guarantee of approximate graph coloring.* Journal of the ACM, 30(4):729–735, 1983.

[28] U. Zwick. *Computer assisted proof of optimal approximability results.* In Proceedings of 13th ACM SODA, pp. 496–505, 2002.

A Pruning efficient covers

We use the following lemma from the extended version of [3]. For completeness, the proof will appear in the full version of this paper.

Lemma 25. *For any $r, \delta > 0$, if in step 2 of KMS$_2(G)$*

$$\Pr[x \textit{ is eliminated} \mid i \in V_\zeta(r)] \geq \delta$$

for at least $n/2$ vertices $i \in V$, and r is c-inefficient for G, then there is a non-empty subgraph $G' = (V', E')$ of G such that for all $j \in V'$ we have (for some universal constant C):

1. *$\{u'_{jk} | k \in \Gamma_{G'}(j)\}$ is a $(\sqrt{3}r, \frac{\delta}{4} - O(1/\log r))$-cover.*
2. *For every $i \in \Gamma_{G'}(j)$ the set*

$$\left\{ u'_{jk} \,\middle|\, -\frac{C}{r^2} \leq u'_{ji} \cdot u'_{jk} \leq \frac{c}{1+c} \cdot \left(1 + \frac{C}{\log r}\right) \right\}$$

is a $\left(\sqrt{3}r, \Omega(\frac{1}{r^3})\right)$-cover.