Improved Approximation Guarantees Through Higher Levels of SDP Hierarchies

Eden Chlamtac[1⋆] and Gyanit Singh[2⋆⋆]

[1] Department of Computer Science, Princeton University, Princeton NJ 08544, USA
chlamtac@cs.princeton.edu
[2] Departmen of Computer Science & Engineering, University of Washington, Box 352350, Seattle, WA 98195-2350, USA
gyanit@cs.washington.edu

Abstract. For every fixed $\gamma \geq 0$, we give an algorithm that, given an n-vertex 3-uniform hypergraph containing an independent set of size γn, finds an independent set of size $n^{\Omega(\gamma^2)}$. This improves upon a recent result of Chlamtac, which, for a fixed $\varepsilon > 0$, finds an independent set of size n^{ε} in any 3-uniform hypergraph containing an independent set of size $(\frac{1}{2} - \varepsilon)n$. The main feature of this algorithm is that, for fixed γ, it uses the $\Theta(1/\gamma^2)$-level of a hierarchy of semidefinite programming (SDP) relaxations. On the other hand, we show that for at least one hierarchy which gives such a guarantee, $1/\gamma$ levels yield no non-trivial guarantee. Thus, this is a first SDP-based algorithm for which the approximation guarantee improves indefinitely as one uses progressively higher-level relaxations.

1 Introduction

Semidefinite Programming (SDP) has been one of the key tools in the development of approximation algorithms for combinatorial optimization problems since the seminal work of Goemans and Williamson [12] on MAXCUT. For a number of problems, including MAXCUT [12], MAX-3SAT [16, 29], and Unique Games [6], SDPs lead to approximation algorithms which are essentially optimal under certain complexity-theoretic assumptions [13, 18]. Howeve, for a host of other problems, large gaps between known hardness of approximation and approximation algorithmic guarantee persist.

One possibility for improvement on the approximation side is the use of so-called SDP hierarchies. In general, Linear Programming (LP) and SDP hierarchies give a sequence of nested (increasingly tight) relaxations for an integer $(0-1)$ program on n variables, where the nth level of the hierarchy is equivalent to the original integer program. Such hierarchies include LS and LS_+ (LP and SDP hierarchies, respectively), proposed by Lovász and Schrijver [22], a stronger

⋆ Supported by Sanjeev Arora's NSF grants MSPA-MCS 0528414, CCF 0514993, ITR 0205594.
⋆⋆ Supported by NSF grant CCF 0514670.

LP hierarchy proposed by Sherali and Adams [26], and the Lasserre [21] SDP hierarchy (see [20] for a comparison).

SDP hierarchies have been studied more generally in the context of optimization of polynomials over semi-algebraic sets [8, 23]. In the combinatorial optimization setting, there has been quite a large number of negative results [2, 1, 25, 28, 11, 5]. This body of work focuses on combinatorial problems for which the quality of approximation (integrality gap) of the hierarchies of relaxations (mostly LS, LS_+, and more recently Sherali-Adams) is poor (often showing no improvement over the simplest LP relaxation) even at very high levels.

On the other hand, there have been few positive results. For random graphs, Feige and Krauthgamer [9] have shown that $\Theta(\log n)$ rounds of LS_+ give a tight relaxation (almost surely) for Maximum Independent Set (a quasi-polynomial time improvement). De la Vega and Kenyon-Mathieu [28] showed that one obtains a polynomial time approximation scheme (PTAS) for MAXCUT in dense graphs using Sherali-Adams. SDP hierarchies at a constant level (where one can optimize in polynomial time) were used recently by Chlamtac [7], who examined the use of the Lasserre hierarchies for Graph Coloring and for Maximum Independent Set in 3-uniform hypergraphs. However, Chlamtac [7] used only the third level of the Lasserre hierarchy, whereas we exploit increasingly higher levels to get better approximation guarantees.

Our focus is on Maximum Independent Set 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. The first SDP-based approximation algorithm for this problem was given by Krivelevich et al. [19], 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 yielded no nontrivial guarantee for $\gamma \leq \frac{1}{2}$. Subsequently, it was shown by Chlamtac [7] that the SDP rounding of [19] finds an independent set of size $\Omega(n^{\varepsilon})$ whenever $\gamma \geq \frac{1}{2} - \varepsilon$, for some fixed $\varepsilon > 0$, if one uses the third level of the Lasserre SDP hierarchy.

We improve upon [7] by giving two algorithms with a non-trivial approximation guarantee for every $\gamma > 0$. In 3-uniform hypergraphs containing an independent set of size γn, both algorithms find an independent set of size $\geq n^{\Omega(\gamma^2)}$. Our result is novel in that for every fixed $\gamma > 0$, the approximation guarantee relies on the $\Theta(1/\gamma^2)$-level of an SDP hierarchy (which can be solved in time $n^{O(1/\gamma^2)}$), and thus gives an infinite sequence of improvements at increasingly high (constant) levels.

For the first of the two hierachies we use, we also show that this guarantee cannot be achieved using a fixed constant level by giving a sequence of integrality gaps. The second hierarchy we consider, the Lasserre hierarchy, allows us to give a slightly better approximation guarantee, by use of an SDP rounding algorithm which uses vectors in the higher levels of the SDP relaxation (in contrast to the approach in [7], where the rounding algorithm was identical to that of [19], and the analysis only relied on the *existence* of vectors in the second and third level).

Note the discrepancy between our result, and the corresponding problem for *graphs*, where Halperin et al. [14] have shown how to find an independent set of size $n^{f(\gamma)}$ for some $f(\gamma) = 3\gamma - O(\gamma^2)$ when the graph contains an independent set of size γn.

The rest of the paper is organized as follows. In Section 2 we define the SDPs used in the various algorithms, and discuss some useful properties of these relaxations. In section 3 we describe a simple integrality gap, followed by a description of the various algorithms and their analyses. Finally, in Section 4, we discuss the possible implications of this result for SDP-based approximation algorithms.

2 SDP Relaxations and Preliminaries

2.1 Previous Relaxation for MAX-IS in 3-Uniform Hypergraphs

The relaxation proposed in [19] 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 following relaxation (where vector v_i represents x_i, and $v_\emptyset$ represents 1:
MAX-KNS(H)

$$\text{Maximize } \textstyle\sum_i \|v_i\|^2 \text{ s.t. } \quad v_\emptyset^2 = 1 \tag{1}$$
$$\forall i \in V \quad v_\emptyset \cdot v_i = v_i \cdot v_i \tag{2}$$
$$\forall (i, j, l) \in E \quad \|v_i + v_j + v_l - v_\emptyset\|^2 \leq 1 \tag{3}$$

2.2 Hypergraph Independent Set Relaxations Using LP and SDP Hierarchies

The Sherali-Adams Hierarchy The Sherali-Adams hierarchy [26] is a sequence of nested linear programming relaxations for $0-1$ polynomial programs. These LPs may be expressed as a system of linear constraints on the variables $\{y_I \mid I \subseteq [n]\}$. To obtain a relaxed (non-integral) solution to the original problem, one takes $(y_{\{1\}}, y_{\{2\}}, \ldots, y_{\{n\}})$.

Suppose $\{x_i^*\}$ is a sequence of n random variables over $\{0, 1\}$, and for all $I \subseteq [n]$ we have $y_I = \mathbb{E}[\prod_{i \in I} x_i^*] = \Pr[\forall i \in I : x_i^* = 1]$. Then by the inclusion-exclusion principle, for any disjoint sets $I, J \subseteq [n]$ we have

$$y_{I,-J} \stackrel{\text{def}}{=} \sum_{J' \subseteq J} (-1)^{|J'|} y_{I \cup J'} = \Pr[(\forall i \in I : x_i^* = 1) \wedge (\forall j \in J : x_j^* = 0)] \geq 0.$$

In fact, it is not hard to see that the constraints $y_{I,-J} \geq 0$ are a necessary and sufficient condition for the existence of a corresponding distribution on $\{0, 1\}$ variables $\{x_i^*\}$. Thinking of the intended solution $\{x_i^*\}$ as a set of indicator variables for a random independent set in a hypergraph $H = (V, E)$ motivates

the following hierarchy of LP relaxations (assume $k \geq \max\{|e| \mid e \in E\}$):
$\mathrm{IS}_k^{\mathrm{SA}}(H)$

$$y_\emptyset = 1 \quad (4)$$

$$\forall I, J \subseteq V \text{ s.t. } I \cap J = \emptyset \text{ and } |I \cup J| \leq k \quad \sum_{J' \subseteq J} (-1)^{|J'|} y_{I \cup J'} \geq 0 \quad (5)$$

$$\forall e \in E \quad y_e = 0 \quad (6)$$

Note that if $\{y_I \mid |I| \leq k\}$ satisfy $\mathrm{IS}_k^{\mathrm{SA}}(H)$, then for any set of vertices $S \subseteq V$ of size k, there is a distribution over independent sets in H for which $\Pr[\forall i \in I : i \in \text{ind. set}] = y_I$ for all subsets $I \subseteq S$.

The Lasserre Hierarchy The relaxations for maximum hypergraph independent set arising from the Lasserre hierarchy [21] are equivalent to those arising from the Sherali-Adams with one additional semidefiniteness constraint:

$$(y_{I \cup J})_{I,J} \succeq 0.$$

We will express these constraints in terms of the vectors $\{v_I | I \subseteq V\}$ arising from the Cholesky decomposition of the positive semidefinite matrix. In fact, we can express the constraints on $\{v_I\}$ in a more succinct form which implies the inclusion-exclusion constraints in Sherali-Adams but does not state them explicitly:
$\mathrm{IS}_k^{\mathrm{Las}}(H)$

$$v_\emptyset^2 = 1 \quad (7)$$

$$|I|, |J|, |I'|, |J'| \leq k \text{ and } I \cup J = I' \cup J' \ \Rightarrow \ v_I \cdot v_J = v_{I'} \cdot v_{J'} \quad (8)$$

$$\forall e \in E \quad v_e^2 = 0 \quad (9)$$

For convenience, we will henceforth write $v_{i_1 \ldots i_s}$ instead of $v_{\{i_1, \ldots, i_s\}}$. We will denote by MAX-$\mathrm{IS}_k^{\mathrm{Las}}(H)$ the SDP

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

Since for any set S of size k all valid constraints on $\{v_I \mid I \subseteq S\}$ are implied by $\mathrm{IS}_k(H)$, this is, for all $k \geq 3$, a tighter relaxation than that of [19].

As in the Sherali-Adams hierarchy, for any set $S \subseteq V$ of size k, we may think of the vectors $\{v_I \mid I \subseteq S\}$ as representing a distribution on random $0-1$ variables $\{x_i^* \mid i \in S\}$, which can also be combined to represent arbitrary events (for example, we can write $v_{(x_i^*=0) \vee (x_j^*=0)} = v_\emptyset - v_{\{i,j\}}$). This distribution is made explicit by the inner-products. Formally, for any two events $\mathcal{E}_1, \mathcal{E}_2$ over the values of $\{x_i^* \mid i \in S\}$, we have $v_{\mathcal{E}_1} \cdot v_{\mathcal{E}_2} = \Pr[\mathcal{E}_1 \wedge \mathcal{E}_2]$.

Moreover, as in the Lovász-Schrijver hierarchy, lower-level relaxations may be derived by "conditioning on $x_i^* = \sigma_i$" (for $\sigma_i \in \{0,1\}$). In fact, we can condition on more complex events. Formally, for any event $\mathcal{E}_0$ involving $k_0 < k$ variables for which $\|v_{\mathcal{E}_0}\| > 0$, we can define

$$v_{\mathcal{E}}|_{\mathcal{E}_0} \stackrel{\text{def}}{=} v_{\mathcal{E} \wedge \mathcal{E}_0} / \|v_{\mathcal{E}_0}\|,$$

and the vectors $\{\, v_I|_{\mathcal{E}_0} \mid |I| \leq k - k_0\}$ satisfy $\mathrm{IS}_{k-k_0}(H)$.

An Intermediate Hierarchy We will be primarily concerned with a hierarchy which combines the power of SDPs and Sherali-Adams local-integrality constraints in the simplest possible way: by imposing the constraint that the variables from the first two levels of a Sherali-Adams relaxation form a positive-semidefinite matrix. Formally, for all $k \geq 3$ and vectors $\{v_\emptyset\} \cup \{v_i \mid i \in V\}$ we have the following system of constraints:
$\mathrm{IS}_k^{\mathrm{mix}}(H)$

$$\exists\{y_I \mid |I| \leq k\} \text{ s.t.} \tag{10}$$
$$\forall I, J \subseteq V, |I|, |J| \leq 1 \ : \ v_I \cdot v_J = y_{I \cup J} \tag{11}$$
$$\{y_I\} \text{ satisfy } \mathrm{IS}_k^{\mathrm{SA}}(H) \tag{12}$$

As above, we will denote by MAX-$\mathrm{IS}_k^{\mathrm{mix}}(H)$ the SDP

$$\text{Maximize } \textstyle\sum_i \|v_i\|^2 \ \text{ s.t. } \{v_\emptyset\} \cup \{v_i\} \text{ satisfy } \mathrm{IS}_k^{\mathrm{mix}}(H).$$

2.3 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 $\zeta \cdot v$ has the standard normal distribution.

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 ([10], Chapter VII) will be crucial.

Lemma 1. *For $s > 0$, we have* $\frac{1}{\sqrt{2\pi}} \left(\frac{1}{s} - \frac{1}{s^3}\right) e^{-s^2/2} \leq N(s) \leq \frac{1}{\sqrt{2\pi} s} e^{-s^2/2}$.

This implies the following corollary, which is at the core of the analysis of many SDP rounding schemes:

Corollary 1. *For any fixed constant $\kappa > 0$, we have $N(\kappa s) = \tilde{O}(N(s)^{\kappa^2})$.*

3 Integrality Gap and Algorithms

3.1 A Simple Integrality Gap

As observed in [26, 7], MAX-KNS$(H) \geq \frac{n}{2}$ for any hypergraph H (even the complete hypergraph). In this section we will show the necessity of using increasingly many levels of the SDP hierarchy MAX-$\mathrm{IS}^{\mathrm{mix}}$ to yield improved approximations, by demonstrating a simple extention of the above integrality gap:

Theorem 1. *For every integer $k \geq 3$ and any 3-uniform hypergraph H, we have* MAX-$\mathrm{IS}_k^{\mathrm{mix}} \geq \frac{1}{k-1} n$.

Proof. Suppose $V(H) = [n]$ and let $v_\emptyset, u_1, \ldots, u_n$ be $n+1$ mutually orthogonal unit vectors. For every $i \in V$ let $v_i = \frac{1}{k-1} v_\emptyset + \sqrt{\frac{1}{k-1} - \frac{1}{(k-1)^2}} u_i$, and $y_{\{i\}} = \frac{1}{k-1}$. Let $y_\emptyset = 1$ and for every pair of distinct vertices $i, j \in V$ let $y_{\{i,j\}} = \frac{1}{(k-1)^2}$. For all sets $I \subseteq V$ s.t. $3 \leq |I| \leq k$, let $y_I = 0$.

It is immediate that constraint (11) and the Sherali-Adams constraint (4) are satisfied. Since $y_I = 0$ for all sets I of size 3, Sherali-Adams constraint (6) is also satisfied. To verify Sherali-Adams constraints (5), it suffices to show, for any set $S \subseteq [n]$ of size k, a corresponding distribution on $0-1$ variables $\{x_i^* \mid i \in S\}$. Indeed, the following is such a distribution: Pick a pair of distinct vertices $i, j \in S$ uniformly at random. With probability $\frac{k}{2(k-1)}$, set $x_i^* = x_j^* = 1$ and for all other $l \in S$, set $x_l^* = 0$. Otherwise, set all $x_l^* = 0$. □

3.2 The Algorithm of Krivelevich, Nathaniel and Sudakov

We first review the algorithm and analysis given in [19]. Let us introduce the following notation: For all $l \in \{0, 1, \ldots, \lceil \log n \rceil\}$, let $T_l \stackrel{\text{def}}{=} \{i \in V \mid l/\log n \leq \|v_i\|^2 < (l+1)/\log n\}$. 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 2. *If the optimum of* $\mathrm{KNS}(H)$ *is* $\geq \gamma n$*, there exists an index* $l \geq \gamma \log n - 1$ *s.t.* $|T_l| = \Omega(n/\log^2 n)$.

Lemma 3. *For index* $l = \beta \log n$ *and hyperedge* $(i,j,k) \in E$ *s.t.* $i,j,k \in T_l$*, constraint* (3) *implies*

$$\|u_i + u_j + u_k\|^2 \leq 3 + (3 - 6\beta)/(1-\beta) + O(1/\log n). \tag{13}$$

Note that constraint (13) becomes unsatisfiable for constant $\beta > 2/3$. Thus, for such β, if $\mathrm{KNS}(H) \geq \beta n$, one can easily find an independent set of size $\tilde{\Omega}(n)$. Using the above notation, we can now describe the rounding algorithm in [19], which is applied to the subhypergraph induced on T_l, where l is as in Lemma 2.

KNS-Round$(H, \{u_i\}, t)$

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

The expected size of the remaining independent set can be bounded from below by $\mathbb{E}[|V_\zeta(t)|] - 3\mathbb{E}[|\{e \in E : e \subseteq V_\zeta(t)\}|]$, since each hyperedge contributes at most three vertices to $V_\zeta(t)$. If hyperedge (i,j,k) is fully contained in $V_\zeta(t)$, then we must have $\zeta \cdot (u_i + u_j + u_k) \geq 3t$, and so by Lemma 3, $\zeta \cdot \frac{u_i + u_j + u_k}{\|u_i + u_j + u_k\|} \geq (\sqrt{(3-3\gamma)/(2-3\gamma)} - O(1/\log n))t$. By Corollary 1, and linearity of expectation, this means the size of the remaining independent set is at least

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

Choosing t appropriately then yields the guarantee given in [19]:

Theorem 2. *Given a 3-uniform hypergraph H on n vertices and m hyperedges 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 $\gamma \leq \frac{1}{2}$. Chlamtac [7] showed that when the vectors satisfy $\mathrm{IS}_3^{\mathrm{Las}}(H)$, the same rounding algorithm does give a non-trivial guarantee (n^{ε}) for $\gamma \geq \frac{1}{2} - \varepsilon$ (for some fixed $\varepsilon > 0$). However, it is unclear whether this approach can work for arbitrarily small $\gamma > 0$.

Let us note the following Lemma which was implicitly used in the above analysis, and which follows immediately from Corollary 1. First, we introduce the following notation for hyperedges e along with the corresponding vectors $\{u_i \mid i \in e\}$:

$$\alpha(e) \stackrel{\text{def}}{=} \tfrac{1}{|e|(|e|-1)} \sum_{i\in e} \sum_{j \in e \setminus \{i\}} u_i \cdot u_j$$

Lemma 4. *In algorithm KNS-Round, the probability that a hyperedge e is fully contained in $V_{\zeta}(t)$ is at most $\tilde{O}(N(t)^{|e|/(1+(|e|-1)\alpha(e))})$.*

3.3 Improved Approximation Via Sherali-Adams Constraints

Before we formally state our rounding algorithm, let us motivate it with an informal overview.

Suppose $\|v_i\|^2 = \gamma$ for all $i \in V$. A closer examination of the above analysis reveals the reason the KNS rounding works for $\gamma > \frac{1}{2}$: For every hyperedge $e \in E$ we have $\alpha(e) < 0$. Thus, the main obstacle to obtaining a large independent set using KNS-Round is the presence of many pairs i, j with large inner-product $u_i \cdot u_j$. As we shall see in section 3.4, we can use higher-moment vectors in the Lasserre hierarchy to turn this into an advantage. However, just using local integrality constraints, we can efficiently isolate a large set of vertices on which the induced subhypergraph has few hyperedges containing such pairs, allowing us to successfully use KNS-Round.

Indeed, suppose that some pair of vertices $i_0, j_0 \in V$ with inner-product $v_{i_0} \cdot v_{j_0} \geq \gamma^2/2$ participates in many hyperedges. That is, the set $S_1 = \{k \in V \mid (i, j, k) \in E\}$ is very large. In that case, we can recursively focus on the subhypergraph induced on S_1. According to our probabilistic interpretation of the SDP, we have $\Pr[x^*_{i_0} = x^*_{j_0} = 1] \geq \gamma^2/2$. Moreover, for any $k \in S_1$ the event "$x^*_k = 1$" is disjoint from the event "$x^*_{i_0} = x^*_{j_0} = 1$". Thus, if we had to repeat this recursive step due to the existence of bad pairs $(i_0, j_0), \ldots, (i_s, j_s)$, then the events "$x^*_{i_l} = x^*_{j_l} = 1$" would all be pairwise exclusive. Since each such event has probability $\Omega(\gamma^2)$, the recursion can have depth at most $O(1/\gamma^2)$, after which point there are no pairs of vertices which prevent us from using KNS-Round.

We are now ready to describe our rounding algorithm. It takes an n-vertex hypergraph H for which $\text{MAX-IS}_k^{\text{mix}}(H) \geq \gamma n$, where $k = \Omega(1/\gamma^2)$ and $\{v_i\}$ is the corresponding SDP solution.

H-Round$(H=(V,E),\{v_i\},\gamma)$

1. Let $n=|V|$ and for all $i,j\in V$, let $\Gamma(i,j)\stackrel{\text{def}}{=}\{k\in V\mid (i,j,k)\in E\}$.
2. If for some $i,j\in V$ s.t. $v_i\cdot v_j\geq \gamma^2/2$ we have $|\Gamma(i,j)|\geq\{n^{1-v_i\cdot v_j/2}\}$, then find an ind. set using H-Round$(H|_{\Gamma(i,j)},\{v_k\mid k\in\Gamma(i,j)\},\gamma)$.
3. Otherwise,
   (a) Define unit vectors $\{w_i\mid i\in V\}$ s.t. for all $i,j\in V$ we have $w_i\cdot w_j=\frac{\gamma}{24}(u_i\cdot u_j)$ (outward rotation).
   (b) Let t be s.t. $N(t)=n^{-(1-\gamma^2/16)}$, and return the independent set found by KNS-Round$(H,\{w_i\mid i\in V\},t)$.

Theorem 3. *For any constant $\gamma>0$, given an n-vertex 3-uniform hypergraph $H=(V,E)$, and vectors $\{v_i\}$ satisfying $\mathrm{IS}^{\mathrm{mix}}_{4/\gamma^2}(H)$ and $|\,\|v_i\|^2-\gamma|\leq 1/\log n$ (for all vertices $i\in V$), algorithm* H-Round *finds an independent set of size $\Omega(n^{\gamma^2/32})$ in H in time $O(n^{3+2/\gamma^2})$.*

Combining this result with Lemma 2 (applying Theorem 3 to the induced subhypergraph $H|_{T_l}$), we get:

Corollary 2. *For all constant $\gamma>0$, there is a polynomial time algorithm which, given an n-vertex 3-uniform hypergraph H containing an independent set of size $\geq\gamma n$, finds an independent set of size $\tilde{\Omega}(n^{\gamma^2/32})$ in H.*

Before we prove Theorem 3, let us first see that algorithm H-Round makes only relatively few recursive calls in Step 2, and that when Step 3b is reached, the remaining hypergraph still contains a large number of vertices.

Proposition 1. *For constant $\gamma>0$, n-vertex hypergraph $H=(V,E)$, and vectors $\{v_i\}$ as in Thereom 3:*

1. *Algorithm* H-Round *makes at most $2/\gamma^2$ recursive calls in Step 2.*
2. *The hypergraph in the final recursive call to* H-Round *contains at least $\sqrt{n}$ vertices.*

Proof. Let $(i_1,j_1),\ldots,(i_s,j_s)$ be the sequence of vertices (i,j) in the order of recursive calls to H-Round in Step 2. Let us first show that for any $s'\leq\min\{s,2/\gamma^2\}$ we have

$$\sum_{l=1}^{s'} v_{i_l}\cdot v_{j_l}\leq 1. \tag{14}$$

Indeed, let $T=\bigcup\{i_l,j_l\mid 1\leq l\leq s'\}$. Since vectors $\{v_i\}$ satisfy $\mathrm{IS}^{\mathrm{mix}}_{4/\gamma^2}(H)$, and $|T|\leq 2s'\leq 4/\gamma^2$, there must be some distribution on independent sets $S\subseteq T$ satisfying $\Pr[k,k'\in S]=v_k\cdot v_{k'}$ for all pairs of vertices $k,k'\in T$. Note that by choice of vertices i_l,j_l, we have $i_{l_2},j_{l_2}\in\Gamma(i_{l_1},j_{l_1})$ for all $l_1<l_2$. Thus, the events "$i_l,j_l\in S$" are pairwise exclusive, and so

$$\sum_{l=1}^{s'} v_{i_l}\cdot v_{j_l}=\Pr[\exists l\leq s': i_l,j_l\in S]\leq 1.$$

Similarly, if $s' \le \min\{s, 2/\gamma^2 - 1\}$, then for any $k \in \bigcap_{l \le s'} \Gamma(i_l, j_l)$ we have $\sum_{l=1}^{s'} v_{i_l} \cdot v_{j_l} + v_k \cdot v_k \le 1$. However, by choice of i_l, j_l, we also have $\sum_{l=1}^{s'} v_{i_l} \cdot v_{j_l} + v_k \cdot v_k \ge |s'|\gamma^2/2 + \gamma - (1/\log n)$. Thus, we must have $s \le 2/\gamma^2 - 1$, otherwise letting $k = i_{2/\gamma^2}$ above, we would derive a contradiction. This proves part 1.

For part 2, it suffices to note that the number of vertices in the final recursive call is at least $n^{\prod(1 - v_{i_l} \cdot v_{j_l}/2)}$, and that by (14) we have $\prod(1 - v_{i_l} \cdot v_{j_l}/2) \ge 1 - \sum v_{i_l} \cdot v_{j_l}/2 \ge \frac{1}{2}$.

□

We are now ready to prove Theorem 3.

Proof (of Theorem 3). For the sake of simplicity, let us assume that for all vertices $i \in V$, $\|v_i\|^2 = \gamma$. Violating this assumption can adversely affect the probabilities of events or sizes of sets in our analysis by at most a constant factor, whereas we will ensure that all inequalities have at least polynomial slack to absorb such errors. Thus, for any $i, j \in V$, we have

$$v_i \cdot v_j = \gamma^2 + (\gamma - \gamma^2) u_i \cdot u_j. \tag{15}$$

For brevity, we will write $v_i \cdot v_j = \theta_{ij}\gamma$ for all $i, j \in V$ (note that all $\theta_{ij} \in [0, 1]$). Moreover, we will use the notation $\alpha(e)$ introduced earlier, but this time in the context of the vector solution $\{w_i\}$:

$$\alpha(e) = \frac{1}{3} \sum_{\substack{i,j \in e \\ i<j}} w_i \cdot w_j.$$

The upper-bound on the running time follows immediately from part 1 of Proposition 1. By part 2 of Proposition 1, it suffices to show that if the condition for recursion in Step 2 of H-Round does not hold, then in Step 3b, algorithm KNS-Round finds an independent set of size $\Omega(N(t)n) = \Omega(n^{\gamma^2/16})$ (where n is the number of vertices in the current hypergraph).

Let us examine the performance of KNS-Round in this instance. Recall that for every $i \in V$, the probability that $i \in V_\zeta(t)$ is exactly $N(t)$. Thus, by linearity of expectation, the expected number of nodes in $V_\zeta(t)$ is $N(t)n$. To retain a large fraction of $V_\zeta(t)$, we must show that few vertices participate in hyperedges fully contained in this set, that is $\mathbb{E}[|\{i \in e \mid e \in E \wedge e \subseteq V_\zeta(t)\}|] = o(N(t)n)$. In fact, since every hyperedge contained in $V_\zeta(t)$ contributes at most three vertices, it suffices to show that $\mathbb{E}[|\{e \in E \mid e \subseteq V_\zeta(t)\}|] = o(N(t)n)$. We will consider separately two types of hyperedges, as we shall see.

Let us first consider hyperedges which contain some pair i, j for which $\theta_{ij} \ge \gamma/2$. We denote this set by E_+. We will assign every hyperedge in E_+ to the pair of vertices with maximum inner-product. That is, for all $i, j \in V$, define $\Gamma_+(i, j) = \{k \in \Gamma(i, j) \mid \theta_{ik}, \theta_{jk} \le \theta_{ij}\}$. By (15), for all $i, j \in V$ and $k \in \Gamma_+(i, j)$ we have

$$\alpha(i, j, k) \le w_i \cdot w_j = \frac{\gamma}{24}(u_i \cdot u_j) = \frac{\gamma(\theta_{ij} - \gamma)}{24(1 - \gamma)} \le \frac{\theta_{ij}\gamma}{24}. \tag{16}$$

Now, by our assumption, the condition for recursion in Step 2 of H-Round was not met. Thus, for all $i, j \in V$ s.t. $\theta_{ij} \geq \gamma/2$, we have

$$|\Gamma_+(i,j)| \leq |\Gamma(i,j)| \leq n^{1-\theta_{ij}\gamma/2}. \tag{17}$$

By linearity of expectation, we have

$$\begin{aligned}
\mathbb{E}[|\{e \in E_+ \mid e \subseteq V_\zeta(t)\}|] &\leq \sum_{e \in E_+} \Pr[e \subseteq V_\zeta(t)] \\
&\leq \sum_{e \in E_+} \tilde{O}(N(t)^{3/(1+2\alpha(e))}) && \text{by Lemma 4} \\
&\leq \sum_{\substack{i,j \in V \\ \theta_{ij} \geq \gamma/2}} \sum_{k \in \Gamma_+(i,j)} \tilde{O}(N(t)^{3/(1+\frac{1}{12}\theta_{ij}\gamma)}). && \text{by (16)}
\end{aligned}$$

By (17), this gives

$$\begin{aligned}
\mathbb{E}[|\{e \in E_+ \mid e \subseteq V_\zeta(t)\}|] &\leq \sum_{\substack{i,j \in V \\ \theta_{ij} \geq \gamma/2}} \tilde{O}(n^{1-\frac{1}{2}\theta_{ij}\gamma} N(t)^{3/(1+\frac{1}{12}\theta_{ij}\gamma)}) \\
&= N(t) \sum_{\substack{i,j \in V \\ \theta_{ij} \geq \gamma/2}} \tilde{O}(n^{1-\frac{1}{2}\theta_{ij}\gamma} N(t)^{(2-\frac{1}{12}\theta_{ij}\gamma)/(1+\frac{1}{12}\theta_{ij}\gamma)}) \\
&= N(t) \sum_{\substack{i,j \in V \\ \theta_{ij} \geq \gamma/2}} \tilde{O}(n^{1-\frac{1}{2}\theta_{ij}\gamma-(1-\frac{1}{16}\gamma^2)(2-\frac{1}{12}\theta_{ij}\gamma)/(1+\frac{1}{12}\theta_{ij}\gamma)}) \\
&\leq N(t) \sum_{\substack{i,j \in V \\ \theta_{ij} \geq \gamma/2}} \tilde{O}(n^{1-\frac{1}{2}\theta_{ij}\gamma-(1-\frac{1}{8}\theta_{ij}\gamma)(2-\frac{1}{12}\theta_{ij}\gamma)/(1+\frac{1}{12}\theta_{ij}\gamma)}) \\
&= N(t)\frac{1}{n} \sum_{\substack{i,j \in V \\ \theta_{ij} \geq \gamma/2}} \tilde{O}(n^{-\frac{5}{96}\theta_{ij}^2\gamma^2/(1+\frac{1}{12}\theta_{ij}\gamma)}) \\
&\leq N(t) n \tilde{O}(n^{-\frac{5}{384}\gamma^4/(1+\frac{1}{24}\gamma^2)}) = o(N(t)n).
\end{aligned}$$

We now consider the remaining hyperedges $E_- = E \setminus E_+ = \{e \in E \mid \forall i, j \in e : \theta_{ij} \leq \gamma/2\}$. By (15), and by definition of $\{w_i\}$, we have

$$\alpha(e) \leq -\frac{\gamma^2}{48(1-\gamma)} \tag{18}$$

for every hyperedge $e \in E_-$. Thus we can bound the expected cardinality of $E_- \cap \{e \subseteq V_\zeta(t)\}$ as follows:

$$\begin{aligned}
\mathbb{E}[|\{e \in E_- \mid e \subseteq V_\zeta(t)\}|] &\leq \sum_{e\in E_+} \Pr[e \subseteq V_\zeta(t)] \\
&\leq \sum_{e\in E_-} \tilde{O}(N(t)^{3/(1+2\alpha(e))}) && \text{by Lemma 4} \\
&= N(t) \sum_{e\in E_-} \tilde{O}(N(t)^{(2-2\alpha(e))/(1+2\alpha(e))}) \\
&\leq N(t) n^3 \tilde{O}(N(t)^{(2-2\gamma+\frac{1}{24}\gamma^2)/(1-\gamma-\frac{1}{24}\gamma^2)}) && \text{by (18)}
\end{aligned}$$

By our choice of t, this gives

$$\begin{aligned}
\mathbb{E}[|\{e \in E_- \mid e \subseteq V_\zeta(t)\}|] &\leq N(t)\tilde{O}(n^{3-(1-\frac{1}{16}\gamma^2)(2-2\gamma+\frac{1}{24}\gamma^2)/(1-\gamma-\frac{1}{24}\gamma^2)}) \\
&= N(t)\tilde{O}(n^{1-(\frac{1}{8}\gamma^3-\frac{1}{384}\gamma^4)/(1-\gamma-\frac{1}{24}\gamma^2)}) = o(N(t)n).
\end{aligned}$$

This completes the proof. □

3.4 A Further Improvement Using The Lasserre Hierarchy

Here, we present a slightly modified algorithm which takes advantage of the Lasserre hierarchy, and gives a slightly better approximation guarantee. As before, the algorithm takes an n-vertex hypergraph H for which MAX-IS$_k^{\text{Las}}(H) \geq \gamma n$, where $k = \Omega(1/\gamma^2)$ and $\{v_I\}_I$ is the corresponding SDP solution.

H-Round$^{\text{Las}}(H = (V, E), \{v_I \mid |I| \leq k\}, \gamma)$

1. Let $n = |V|$ and let $l = \gamma' \log n - 1$ be as in Lemma 2 (where $\gamma' \geq \gamma$). If $\gamma' > 2/3 + 2/\log n$, output T_l.
2. Otherwise, set $H = H|_{T_l}$, and $\gamma = \gamma'$.
3. If for some $i, j \in T_l$ s.t. $\rho_{ij} = v_i \cdot v_j \geq \gamma^2/2$ we have $|\Gamma(i,j)| \geq \{n^{1-\rho_{ij}}\}$, then find an independent set using H-Round$(H|_{\Gamma(i,j)}, \{v_I|_{x_i^*=0 \vee x_j^*=0} \mid I \subseteq \Gamma(i,j), |I| \leq k-2\}, \gamma/(1-\rho_{ij}))$.
4. Otherwise,
   (a) Define unit vectors $\{w_i \mid i \in V\}$ s.t. for all $i, j \in V$ we have $w_i \cdot w_j = \frac{\gamma}{12}(u_i \cdot u_j)$ (outward rotation).
   (b) Let t be s.t. $N(t) = n^{-(1-\gamma^2/8)}$, and return the independent set found by KNS-Round$(H, \{w_i \mid i \in V\}, t)$.

For this algorithm, we have the following guarantee:

Theorem 4. *For any constant $\gamma > 0$, given an n-vertex 3-uniform hypergraph $H = (V, E)$ for which* MAX-IS$^{\text{Las}}_{8/(3\gamma^2)}(H) \geq \gamma n$ *and vectors $\{v_I\}$ the corresponding solution, algorithm* H-Round$^{\text{Las}}$ *finds an independent set of size $\Omega(n^{\gamma^2/8})$ in H in time $O(n^{3+8/(3\gamma^2)})$.*

We will not prove this theorem in detail, since the proof is nearly identical to that of Theorem 3. Instead, we will highlight the differences from algorithm H-Round, and the reasons for the improvement. First of all, the shortcut in step 1 (which accounts for the slightly lower level needed in the hierarchy) is valid since (as can be easily checked) constraint (3) cannot be satisfied (assuming (2) holds) when $\|v_i\|^2, \|v_j\|^2, \|v_l\|^2 > 2/3$.

The improvement in the approximation guarantee can be attributed to the following observation. Let $\{(i_1, j_1), \ldots, (i_s, j_s)\}$ be the pairs of vertices chosen for the various recursive invocations of the algorithm in Step 3. Then in the probabilistic interpretation of the SDP solution, we have carved an event of probability $\rho = \rho_{i_1 j_1} + \ldots + \rho_{i_s j_s}$ out of the sample space, and thus the SDP solution is conditioned on an event of probability $1-\rho$. Hence, the hypergraph in the final call contains $n_\rho \geq \tilde{\Omega}(n^{1-\rho})$ vertices, and the SDP value is $\gamma_\rho n_\rho$ where $\gamma_\rho \geq \gamma/(1-\rho)$. Thus one only needs to show that assuming the condition in Step 3 does not hold, the call to KNS-Round in Step 4b returns an independent set of size at least

$$n_\rho^{\gamma_\rho^2/8} \geq n^{\gamma^2/(8(1-\rho))} \geq n^{\gamma^2/8}.$$

The proof of this fact is identical to the proof of Theorem 3.

4 Discussion

Theorem 3, together with the integrality gap of Theorem 1, demonstrate that the hierarchy of relaxations MAX-IS$_k^{\text{mix}}$ gives an infinite sequence of improved approximations for higher and higher levels k. We do not know if similar integrality gaps hold for the Lasserre hierarchy, though we know that at least the integrality gap of Theorem 1 cannot be lifted even to the second level in the Lasserre hierarchy. In light of our results, we are faced with two possible scenarios:

1. For some fixed k, the kth level of the Lasserre hierarchy gives a better approximation than MAX-IS$_l^{\text{mix}}$ for any (arbitrary large constant) l, or
2. The approximation curve afforded by the kth level Lasserre relaxation gives strict improvements for infinitely many values of k.

While the second possibility is certainly the more exciting of the two, a result of either sort would provide crucial insights into the importance of lift-and-project methods for approximation algorithms. Recently Schoenebeck [27] has produced strong integrality gaps for high-level Lasserre relaxations for random 3XOR formulas, which rely on properties of the underlying 3-uniform hypergraph structure. It will be very interesting to see whether such results can be extended to confirm the second scenario, above.

Finally, we note that the existence of provably improved approximations at infinitely many constant levels of an SDP hierarchy is surprising in light of the recent work of Raghavendra [24]. One implication of that work is that if the Unique Games Conjecture [17] is true, then for every k-CSP, the kth level of a mixed hierarchy (such as MAX-IS$^{\text{mix}}$) suffices to get the best possible approximation (achievable in polynomial time). Our result, when combined with the work of Raghavendra [24], does not refute the Unique Games Conjecture (essentially, since the guaranteed optimality of the relaxations in [24] is only up to any arbitrary additive linear error). However, it may help shed light on the characteristics of combinatorial optimization problems which stand to benefit from the use of lift-and-project techniques.

References

1. M. Alekhnovich, S. Arora, and I. Tourlakis. *Towards strong nonapproximability results in the Lovász-Schrijver hierarchy.* In Proceedings of 37th Annual ACM Symposium on Theory of Computing, 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 Symposium of Theory of Computing, pp. 215-224, 2006.
4. S. Arora, S. Rao, and U. Vazirani. *Expander flows, geometric embeddings and graph partitioning.* In Proceedings of the 36th Annual ACM Symposium on Theory of Computing, pp. 222–231, 2004.
5. M. Charikar, K. Makarychev, and Y. Makarychev. *Integrality Gaps for Sherali-Adams Relaxations.* Manuscript.
6. M. Charikar, K. Makarychev, and Y. Makarychev. *Near-Optimal Algorithms for Unique Games.* In Proceedings of the 38th Annual ACM Symposium on Theory of Computing, pp. 205-214, 2006.
7. E. Chlamtac. *Approximation Algorithms Using Hierarchies of Semidefinite Programming Relaxations.* In Proceedings of 48th IEEE Symposium on Foundations of Computer Science, pp. 691-701, 2007.
8. 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.
9. U. Feige, R. Krauthgamer *The Probable Value of the Lovász-Schrijver Relaxations for Maximum Independent Set.* SIAM Journal on Computing, 32(2):345–370, 2003.
10. W. Feller. *An Introduction to Probability Theory and Its Applications,* volume 1. John Wiley & Sons, 3rd edition, 1968.
11. 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 Symposium on Foundations of Computer Science, pp. 702-712, 2007.
12. 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.
13. J. Håstad. *Some optimal inapproximability results.* Journal of the ACM, vol. 48, no. 4, pp. 798–859, July 2001.
14. E. Halperin, R. Nathaniel, and U. Zwick. *Coloring k-colorable graphs using relatively small palettes. J. Algorithms*, 45(1):72–90, 2002.
15. D. Karger, R. Motwani, and M. Sudan. *Approximate graph coloring by semidefinite programming.* Journal of the ACM, 45(2):246–265, 1998.

16. H. Karloff and U. Zwick. *A 7/8-approximation algorithm for max 3SAT?* In Proceedings of 38th IEEE Symposium on Foundations of Computer Science, pages 406–415, 1997.
17. S. Khot. *On the power of unique 2-prover 1-round games.* In Proceedings of the 34th ACM Symposium on Theory of Computing, pp. 767–775, 2002.
18. S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. *Optimal inapproximability results for MAX-CUT and other 2-variable CSPs?* SIAM Journal on Computing, 37(1):319–357, 2007. ECCC Report TR05-101, 2005.
19. 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.
20. 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.
21. J.B. Lasserre. *An explicit exact SDP relaxation for nonlinear 0-1 programs.* Lecture Notes In Computer Science, vol. 2081, pp. 293–303, 2001.
22. 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.
23. J. Nie, M. Schweighofer. *On the complexity of Putinar's Positivstellensatz.* Journal of Complexity, vol. 23, no. 1, pp. 135–150, Feb. 2007.
24. P. Raghavendra. *Optimal Algorithms and Inapproximability Results For Every CSP?* In Proceedings of 30th ACM Symposium on Theory of Computing, pp. 245–254, 2008.
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 Symposium on Theory of Computing, 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. G. Schoenebeck. Personal communication.
28. F.W. de la Vega, C. Kenyon-Mathieu. *Linear programming relaxations of maxcut.* In Proceedings of 18th ACM Symposium on Discrete Algorithms, pp. 53–61, 2007.
29. U. Zwick. *Computer assisted proof of optimal approximability results.* In Proceedings of 13th ACM Symposium on Discrete Algorithms, pp. 496–505, 2002.