Non-Local Analysis of SDP-Based Approximation Algorithms

Eden Chlamtac

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
By the Department of
Computer Science

Advisor: Sanjeev Arora

January 2009

Abstract

In this work, we study approximation algorithms based on semidefinite programming (SDP) for which the performance guarantee involves a non-local analysis, and in some instances a non-local SDP relaxation.

We examine two such approaches. The first of these is inspired by recent work of Arora, Rao and Vazirani on Sparsest Cut. Using a geometric intuition similar to theirs, we give an algorithm for coloring 3-colorable graphs which is nearly identical to that of Blum and Karger, and finds a legal coloring which uses roughly $O(n^{0.2130})$ as opposed to the original $O(n^{0.2143})$ guarantee in that paper.

The second approach makes use of SDP hierarchies, on which prior work has yielded mostly negative results. Using this method, we give an algorithm for coloring 3-colorable graphs which finds a legal $O(n^{0.2072})$-coloring.

As an additional application of this approach, in 3-uniform hypergraphs containing an independent set of size γn (for any constant $\gamma > 0$), we describe an algorithm which finds an independent set of size $n^{\Omega(\gamma^2)}$ using the $\Theta(1/\gamma^2)$-level of an SDP hierarchy. We also present integrality gaps for this hierarchy which imply improved performance guarantees as one uses progressively higher-level SDP relaxations.

Acknowledgments

First of all, I would like to thank my advisor, Sanjeev Arora, for all his guidance and support throughout my time at Princeton. Without his continuing direction and help during this time, this thesis would not have been possible.

I am also grateful to the other members of my committee for their helpful comments both before and after the writing of this thesis: Boaz Barak, Moses Charikar, Bernard Chazelle and Ken Steiglitz. I would like to thank the coauthors who collaborated with me on the various results presented here, Sanjeev Arora, Moses Charikar, and Gyanit Singh, as well as my other coauthors: Uri Feige, Hugues Hoppe, Konstantin Makarychev, Yury Makarychev, Diego Nehab and Pedro Sander.

I would like to thank the Computer Science department at Princeton University, and especially Melissa Lawson, whose handling of all the administrative work made my PhD studies flow smoothly and seamlessly. I also thoroughly enjoyed one semester visiting the University of Washington and Microsoft Research, Redmond, and one semester visiting the Weizmann Institute of Science in Israel, and I am grateful to them for the wonderful opportunities they provided.

I gratefully acknowledge a Francis Upton fellowship, and NSF grants MSPA-MCS 0528414, CCF 0514993 and ITR 0205594 for providing the financial support for my research at Princeton University.

Finally, I would like to thank my parents, Merav and Imrich Chlamtac, and my sister Noga Chlamtac, who have stood by me through good times and bad, and without whose love and support I would not be where I am today. I also thank all my relatives, who have waited long for my return, and especially my grandmother, Klára Chlamtáčová, who through her unwavering love, affection and guidance, has always been much more than a grandmother to me. I dedicate this thesis to her.

To Oma

Contents

List of Figures

Chapter 1

Introduction

1.1 Preface

In this work, we study approximation algorithms based on semidefinite programming (SDP). Semidefinite programming has been one of the central tools in approximation algorithm design since the seminal work of Goemans and Williamson [19] on MAXCUT. In all instances, the SDP-based approach to approximation algorithms follows the same general lines: as in the linear programming (LP) approach, one takes a combinatorial (discrete) optimization problem which is NP-hard, relaxes it to a convex optimization problem (in this case, a *vector* problem) which is tractable, and then "rounds" the relaxed solution to a discrete one.

The core of the analysis of the performance guarantee of such algorithms lies in examining the rounding algorithm. Traditionally, the analysis of the rounding algorithm involves examining the behavior of the algorithm on local configurations of vectors (often, only pairs of vectors) related to local constraints in the combinatorial optimization problem. While for a large number of problems (e.g. [19, 41, 10]), this method yields approximation guarantees which are optimal under certain complexity theoretic assumptions, for several other problems the gap between known hardness of approximation and approximation algorithmic guarantee remains quite large. Thus there is a need to expand the SDP-based algorithmic toolkit, both in terms of algorithm design and in terms of proof techniques for the performance guarantee.

We examine two avenues for improvement involving non-local analysis of SDP rounding algorithms. The first of these is based on a geometric intuition pioneered in recent work of Arora, Rao and Vazirani [5]. Using this approach, we are able to show that an algorithm for coloring 3-colorable graphs which is nearly identical to that of Blum and Karger [7] finds a legal coloring which uses roughly $O(n^{0.2130})$ colors as opposed to the original $O(n^{0.2143})$ guarantee in that paper. While the improvement may not seem quantitatively substantial, the techniques introduced have paved the way to further improvements and set this problem apart from others for which a simple SDP-based approximation algorithm and analysis give the best-possible result (up to certain complexity-theoretic assumptions).

The second approach involves the use of SDP hierarchies. These give a sequence of nested (increasingly tight) relaxations for any integer $(0-1)$ program on n variables, where the nth level of the hierarchy is equivalent to the original integer program, and the kth level produces an SDP the optimum of which can be found in time $n^{O(k)}$. Both LP and SDP hierarchies lend themselves quite naturally to non-local analysis, since each level introduces constraints involving progressively larger sets of variables.

While most previous work on LP and SDP hierarchies has focused on negative results [2, 1, 34, 38, 18, 9, 36], we will investigate two positive applications. Starting with the aforementioned coloring problem, we show that, using the third level of a certain SDP hierarchy, we can find a legal $O(n^{0.2072})$-coloring in any 3-colorable graph. Moreover, in 3-uniform hypergraphs containing an independent set of size γn (for any constant $\gamma > 0$), we describe an algorithm which finds an independent set of size $n^{\Omega(\gamma^2)}$ using the $\Theta(1/\gamma^2)$-level of another SDP hierarchy. On the other hand, using a simpler SDP, no non-trivial guarantee is possible for $\gamma \leq \frac{1}{2}$ (in fact, in the SDP hierarchy we consider, no such guarantee can be obtained at any level up to $\frac{1}{\gamma} + 1$).

1.2 Previous Publications

The results covered in this thesis, or preliminary versions thereof, have appeared previously in the following publications (listed in chronological order):

- S. Arora, M. Charikar and E. Chlamtac. *New Approximation Guarantee for Chromatic Number.* In Proceedings of the 38th ACM Symposium on Theory of Computing (STOC), pp. 215–224, 2006.

- E. Chlamtac. *Approximation Algorithms Using Hierarchies of Semidefinite Programming Relaxations.* In Proceedings of the 48th IEEE Symposium on Foundations of Computer Science (FOCS), pp. 691–701, 2007.

- E. Chlamtac and G. Singh. *Improved Approximation Guarantees Through Higher Levels of SDP Hierarchies.* In Proceedings of the 11th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX), pp. 49–62, 2008.

1.3 Graph Coloring

In the graph k-coloring problem we wish to assign each vertex one of k colors such that every pair of vertices connected with an edge are assigned different colors. Finding the minimum k for which a k-coloring exists (the chromatic number of the graph) is a classical NP-complete problem.

In general, it is NP-hard to approximate the chromatic number to within $n^{1-\varepsilon}$ for any constant $\varepsilon > 0$ [15, 40], though the problem becomes much easier if we are guaranteed that the chromatic number of the graph is small. Specifically, we concentrate on approximation algorithms of the following form: Given a 3-colorable graph (the coloring is not known), find a legal coloring using as few colors as possible. Dinur, Mossel and Regev [14] have shown that, assuming some variant of the Unique Games Conjecture [25], this cannot be done using any constant number of colors, though this does not rule out the possibility of finding, say, an $O(\log\log n)$ coloring.

In terms of positive results, there has been the focus of a long line of work. Prior to the results described here, the best known algorithm was due to Blum and Karger [7], who gave a $\tilde{O}\left(n^{3/14}\right)$ coloring (we use $\tilde{O}\left(f(n)\right)$ to mean $O(f(n)\log^C n)$ for some constant C). Their work combined the SDP-based approach of Karger, Motwani and Sudan [23] (which gives a $\tilde{O}\left(\Delta^{1/3}\right)$ coloring in graphs with maximum degree Δ) with an earlier combinatorial approach of Blum [6]. While the results presented here rely on improving the performance of SDP-based algorithms, we also formalize the method of Blum and Karger [7] in Section 3.7, giving a general method to combine the combinatorial tools of Blum [6] with an SDP-based algorithm.

In Chapter 3, we first review the algorithm of Karger, Motwani and Sudan [23], and then give an improved analysis for an algorithm which uses the same SDP relaxation proposed in [23], and a nearly identical rounding algorithm. This improves the $\tilde{O}\left(\Delta^{1/3}\right)$ guarantee in [23], and when combined with the combinatorial approach in [6] gives a $O(n^{0.2130})$ coloring. The analysis is inspired by a geometric approach used in the Sparsest Cut algorithm of Arora, Rao and Vazirani [5].

In Chapter 4, we present an algorithm which makes use of a tighter SDP relaxation arising from the Lasserre hierarchy [30]. While the running time for this algorithm is higher (due to the higher cost of finding an SDP optimum for this relaxation), the algorithm presented gives the current best guarantee, namely a legal coloring using $O\left(n^{0.2072}\right)$ colors.

1.4 Hypergraphs and Hypergraph Independent Sets

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. Finding a maximum size independent set in a hypergraph is a natural generalization of the Maximum Independent Set problem in graphs, where hyperedges have size 2. Moreover, a wide range of $0-1$ optimization problems with local constraints, for instance all binary Constraint Satisfaction Problems (CSPs), can be naturally expressed as Hypergraph Independent Set problems.

We focus on the case of 3-uniform hypergraphs. The first SDP-based approximation algorithm for this problem was given by Krivelevich, Nathaniel and Sudakov [28], 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 $\gamma \leq \frac{1}{2}$. On the hardness side, Khot and Regev [27] have shown that for any constant $\varepsilon > 0$, it is hard to find an independent set of size $\geq \varepsilon n$ in a 3-uniform hypergraph containing an independent set of size $\frac{2}{3} - \varepsilon$ assuming the Unique Games Conjecture [25]. However, this still leaves much room for improvement.

In Chapter 5, we present two algorithms which, for every $\gamma > 0$, in any n-vertex 3-uniform hypergraph containing an independent set of size γn, find an independent set of size $n^{\Omega(\gamma^2)}$. Each of these algorithms relies on an SDP relaxation arising from the $\Theta(1/\gamma^2)$-level of some SDP hierarchy. For the hierarchy used in the first algorithm, we also present an integrality gap which implies that this performance guarantee cannot be achieved at any level up to $\frac{1}{\gamma} + 1$. This implies a sequence of improving approximation guarantees as one uses progressively higher-level relaxations.

The various hierarchies used are detailed in Section 2.1, where we also present some useful properties of the corresponding relaxations, and provide some intuition for analyzing the structure of SDP hierarchies which will motivate the later analysis in Chapters 4 and 5.

Chapter 2

SDP and SDP Rounding

To illustrate the use of Semidefinite Programming (SDP) in approximation algorithms, let us consider the Maximum Independent Set problem in graphs. This is a special case of the Hypergraph Maximum Independent Set problem, which will be discussed in Chapter 5. The most common approach is to first formulate the problem as a Quadratic Program, and then find a corresponding SDP relaxation. The following is a natural Quadratic Programming formulation for Maximum Independent Set in a graph $G = (V, E)$:

$$\max_{\{x_i \mid i \in V \cup \{0\}\}} \sum_{i \in V} x_i^2 \qquad \text{s.t.}$$

$$x_0^2 = 1 \tag{2.1}$$

$$\forall i \in V \quad x_i^2 = x_i x_0 \tag{2.2}$$

$$\forall (i, j) \in E \quad x_i x_j = 0 \tag{2.3}$$

We now arrive at an SDP relaxation by replacing the linear variables $\{x_i\}$ above with vectors $\{v_i\}$, and replacing all products with dot products of vectors. This gives the following relaxation:

$$\max_{\{v_i | i \in V \cup \{0\}\}} \sum_{i \in V} \|v_i\|^2 \qquad \text{s.t.}$$

$$\|v_0\|^2 = 1 \tag{2.4}$$

$$\forall i \in V \quad \|v_i\|^2 = v_i \cdot v_0 \tag{2.5}$$

$$\forall (i,j) \in E \qquad v_i \cdot v_j = 0 \tag{2.6}$$

The benefit of working with an SDP relaxation as above is that the optimum, along with the corresponding vector solution $\{v_i\}$, can be found in polynomial time to within arbitrary precision (as was first shown by Grötschel, Lovász, and Schrijver [20]). The final step involves finding an independent set by using the graph structure and the vector solution $\{v_i\}$. This is known as the *rounding algorithm.* The design of rounding algorithms and especially the analysis of their performance guarantee (in this case, the size of the independent set found, as a function of the SDP optimum) is the most challenging and most interesting part, and will be the main focus of later chapters. We first examine a number of the relaxations which will be used later, as well as some important components of SDP rounding algorithms.

2.1 Hierarchies of LP and SDP Relaxations

LP and SDP hierarchies give a sequence of relaxations for an integer program on n variables, where the nth level of the hierarchy is equivalent to the original integer program. These include LS and LS_+ (LP and SDP hierarchies, respectively), proposed by Lovász and Schrijver [31], a stronger LP hierarchy proposed by Sherali and Adams [35], and the Lasserre [30] SDP hierarchy (see [29] for a comparison).

SDP hierarchies have been studied more generally in the context of optimization of polynomials over semi-algebraic sets [13, 32]. In the combinatorial optimization setting, there has been quite a large number of negative results [2, 1, 34, 38, 18, 9, 36]. 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 [15] 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 [38] showed that one obtains a polynomial time approximation scheme (PTAS) for MAXCUT in dense graphs using Sherali-Adams.

We will consider various LP and SDP hierarchies in this section through the lens of relaxations for Hypergraph Maximum Independent Set. This problem is sufficiently general to capture all $0-1$ optimization problems with local constraints, yet it will allow us to present these hierarchies in a clear and intuitive manner.

2.1.1 The Sherali-Adams Hierarchy

The Sherali-Adams hierarchy [35] 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 $(y_{\{1\}}, y_{\{2\}}, \ldots, y_{\{n\}})$.

As a gedankenexperiment, 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 $k \geq \max\{|e| \mid e \in E\}$):
$\mathrm{IS}_k^{\mathrm{SA}}(H)$

$$y_\emptyset = 1 \tag{2.7}$$

$$\forall I, J \subseteq V \text{ s.t. } I \cap J = \emptyset \text{ and } |I \cup J| \leq k \qquad \sum_{J' \subseteq J} (-1)^{|J'|} y_{I \cup J'} \geq 0 \tag{2.8}$$

$$\forall e \in E \qquad y_e = 0 \tag{2.9}$$

As noted above, if $\{y_I \mid I \subseteq V\}$ satisfy $\mathrm{IS}_n^{\mathrm{SA}}(H)$ (where $n = |V|$), 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 $I \subseteq V$ (that is, the nth level of the hierarchy, $\mathrm{IS}_n^{\mathrm{SA}}(H)$, corresponds to the *integer* polytope). In particular, for any integer $1 \leq k \leq n$, this implies that if $\{y_I \mid |I| \leq k\}$ satisfy $\mathrm{IS}_k^{\mathrm{SA}}(H)$, then for any set $S \subseteq V$ 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 $I \subseteq S$.

2.1.2 The Lasserre Hierarchy

The relaxations for maximum hypergraph independent set arising from the Lasserre hierarchy [30] 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: $\mathrm{IS}_k^{\mathrm{Las}}(H)$

$$v_\emptyset^2 = 1 \tag{2.10}$$

$$|I|, |J|, |I'|, |J'| \leq k \text{ and } I \cup J = I' \cup J' \Rightarrow \quad v_I \cdot v_J = v_{I'} \cdot v_{J'} \tag{2.11}$$

$$\forall e \in E \quad v_e^2 = 0 \tag{2.12}$$

For convenience, whenever possible, 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

$$\text{Maximize } \textstyle\sum_i \|v_i\|^2 \text{ s.t. } \{v_I \mid I \subseteq V \wedge |I| \leq k\} \text{ satisfy } \mathrm{IS}_k^{\mathrm{Las}}(H).$$

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 be combined to represent arbitrary events. Formally, the vector corresponding to the event $\mathcal{E}_{I,-J} =$ "$\forall i \in I, j \in J : (x_i^* = 1) \wedge (x_j^* = 0)$" is

$$v_{\mathcal{E}_{I,-J}} \stackrel{\text{def}}{=} \textstyle\sum_{J' \subseteq J} (-1)^{|J'|} v_{I \cup J'}.$$

The picture is then completed by defining $v_{\bigcup_l \mathcal{E}_l} = \sum_l v_{\mathcal{E}_l}$ for disjoint events $\mathcal{E}_l$ (for example, we can write $v_{(x_i^*=0)\vee(x_j^*=0)} = v_{\emptyset,-\{i,j\}} + v_{\{i\},-\{j\}} + v_{\{j\},-\{i\}} = v_\emptyset - v_{ij}$). The corresponding local distribution is made explicit by the inner-products: 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 $\mathrm{IS}_{k-k_0}(H)$.

2.1.3 An Intermediate Hierarchy

We will also use 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 a k_0-uniform hypergraph $H = (V, E)$, for $k \geq k_0$ and vectors $\{v_\emptyset\} \cup \{v_i \mid i \in V\}$ we have the following system of constraints: $\mathrm{IS}_k^{\mathrm{mix}}(H)$

$$\begin{aligned} & && \exists \{y_I \mid |I| \leq k\} \text{ s.t.} & \\ &\forall I, J \subseteq V, |I|, |J| \leq 1 \ : && v_I \cdot v_J = y_{I \cup J} & (2.13) \\ & && \{y_I\} \text{ satisfy } \mathrm{IS}_k^{\mathrm{SA}}(H) & (2.14) \end{aligned}$$

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

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

2.1.4 Relating the Lasserre Hierarchy to Distributions on $0-1$ Solutions

Recall the probabilistic interpretation of vector solutions $\{v_I \mid |I| \leq k\}$ satisfying $\mathrm{IS}_k^{\mathrm{Las}}(H)$. This intuition will allow us to deduce additional geometric properties of the SDP solution, which can then be proven rigorously using the Lasserre constraints (2.11). Let us consider the following crucial example (which will also motivate the following lemma). Consider some event A relating to partial assignments of $\{x_i^* \mid i \in V\}$ (e.g. "$\forall i \in I : x_i^* = 1$"). Suppose that $\Pr[A] = p$ and that we have many events B_l, sub-events of A, such that $\Pr[B_j \mid A] = q$. Then for most pairs $B_l, B_{l'}$ we have $\Pr[B_l \wedge B_{l'} \mid A] \geq q^2 - o(1)$ since, in principle, most pairs of events cannot be too anti-correlated. (Indeed, consider a set of indicator variables X_l. The inequality $\mathbb{E}[(\sum_l X_l)^2] \geq \mathbb{E}[\sum_l X_l]^2$ implies that $\sum_{l \neq l'} \mathbb{E}[X_l X_{l'}] \geq \sum_{l \neq l'} \mathbb{E}[X_l]\mathbb{E}[X_{l'}] - \sum_l \mathbb{E}[X_l](1 - \mathbb{E}[X_l])$.)

If we think of the vectors representing these events, we have $v_{B_l} \cdot v_{B_{l'}} \geq pq^2 - o(1)$. This would also hold true for most pairs $B_l, B_{l'}$ if the vectors $\{v_{B_l}\}$ all shared a common component of length $\sqrt{pq^2}$. That is, if there were some unit vector v'_A such that $v_{B_l} \cdot v'_A \geq \sqrt{pq^2}$. Now, suppose that for some A', a super-event of A, we were guaranteed that the vectors v_{B_l} had the form $v_{B_l} = \sqrt{p'} \cdot \frac{v_{A'}}{\|v_{A'}\|} + w_{B_l}$ for some $w_{B_l} \perp v_{A'}$. By a similar argument, we would expect that the vectors w_{B_l} have a common component of length $\geq \sqrt{pq^2 - p'}$ (that is, they have a projection of at least this magnitude on the same unit vector). Using the Lasserre hierarchy, we can guarantee the existence of such a vector, as demonstrated by the following lemma (in this case think of A as a union of the mutually exclusive events "$\forall h \in I_i : x_h = 1$" and of a single event B_l as a union of the respective sub-events "$(\forall h \in I_i : x_h = 1) \wedge (\forall j \in J : x_j = 1)$").

Lemma 2.1.1. *Let $\{v_I\}$ be a set of vectors satisfying* (2.11)*, 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 $\forall i \neq i'$, $v_{I_i} \cdot v_{I_{i'}} = 0$ 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 $x_0 \in Span(\{v_I \mid I \subseteq \bigcup_i I_i\})$ 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 $x'_0 \in Span(\{v_I \mid I \subseteq \bigcup_{i,j} I_{ij}\})$ such that $x'_0 \cdot v'_J = \sqrt{\sum_{i,j} p_{ij} q_{ij}^2 - \sum_i p_i q_i^2}$ (where $p_{ij} = \left\|v_{I_{ij}}\right\|^2$ and $p_{ij} q_{ij} = \left\|v_{I_{ij} \cup J}\right\|^2$).*

3. *The vectors x_0 and x'_0 are uniquely determined by the vectors $\{v_I \mid I \subseteq \bigcup_{i,j} I_{ij}\}$ and values $\{p_i\}$, $\{q_i\}$, $\{p_{ij}\}$ and $\{q_{ij}\}$. They do not depend on the choice of vector v_J.*

Proof. For part 1, it suffices to check, by computing inner products, and using constraint (2.11), that $v_J = \sum_i q_i v_{I_i} + v'_J$ (where $v'_J \cdot v_{I_i} = 0$).

For part 2, first observe that by orthogonality of the various $v_{I_{ij}}$, and since $v_{I_i} = \sum_j v_{I_{ij}}$, we have

$$p_i = \|v_{I_i}\|^2 = \sum_j \left\|v_{I_{ij}}\right\|^2 = \sum_j p_{ij} \tag{2.15}$$

and

$$p_i q_i = v_J \cdot v_{I_i} = \sum_j v_J \cdot v_{I_{ij}} = \sum_j p_{ij} q_{ij}. \tag{2.16}$$

Now, by constraint (2.11), and orthogonality of the various $v_{I_{ij}}$, for any i_0, j_0 we have

$$v'_J \cdot v_{I_{i_0 j_0}} = (v_J - \sum_i q_i v_{I_i}) \cdot v_{I_{i_0 j_0}} = p_{i_0 j_0} q_{i_0 j_0} - \sum_i q_i \sum_j v_{I_{ij}} \cdot v_{I_{i_0 j_0}} = p_{i_0 j_0}(q_{i_0 j_0} - q_{i_0}).$$

Thus, we can represent v'_J as $v'_J = \sum_{i,j}(q_{ij} - q_i)v_{I_{ij}} + v''_J$, where v''_J is orthogonal to all $v_{I_{ij}}$. Finally, note that

$$\begin{aligned}
\Big\| \sum_{i,j}(q_{ij} - q_i)v_{I_{ij}} \Big\|^2 &= \sum_{i,j}(q_{ij} - q_i)^2 p_{ij} \\
&= \sum_{i,j} q_{ij}(q_{ij} - q_i)p_{ij} - \sum_i q_i \sum_j (q_{ij} - q_i)p_{ij} \\
&= \sum_{i,j} q_{ij}(q_{ij} - q_i)p_{ij} - \sum_i q_i \Big(\Big(\sum_j p_{ij} q_{ij}\Big) - p_i q_i\Big) && \text{by (2.15)} \\
&= \sum_{i,j} q_{ij}(q_{ij} - q_i)p_{ij} && \text{by (2.16)} \\
&= \sum_{i,j} p_{ij} q_{ij}^2 - \sum_i q_i \sum_j p_{ij} q_{ij} \\
&= \sum_{i,j} p_{ij} q_{ij}^2 - \sum_i p_i q_i^2 && \text{by (2.16).}
\end{aligned}$$

□

2.2 Gaussian Vectors and SDP Rounding

Recall that the *standard normal distribution* has density function $\frac{1}{\sqrt{2\pi}} e^{-t^2/2}$. 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 $v \in \mathbb{R}^n$, the projection $\zeta \cdot v$ has the standard normal distribution.

We use the following notation for the corresponding tail bound:

$$N(s) \stackrel{\text{def}}{=} \int_s^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

The following property of the normal distribution ([17], Chapter VII) will be crucial.

Lemma 2.2.1. *For $s > 0$, we have* $\frac{1}{\sqrt{2\pi}} \left(\frac{1}{s} - \frac{1}{s^3}\right) e^{-s^2/2} \leq N(s) \leq \frac{1}{\sqrt{2\pi} s} e^{-s^2/2}$.

We introduce the following definition:

Definition 2.2.2. 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 $x \in X$.

The analysis of SDP rounding algorithms frequently involves expressions of the form $\Pr_\zeta[\exists x \in X : \zeta \cdot x \geq s]$, for ζ as above, and set of unit vectors X. It is easy to see that $|X|\, N(s)$ 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 2.2.3. *Let X be a ρ-cluster for some fixed constant $\rho \in (0, 1)$. Then for all $s \geq 0$ and $0 \leq r \leq \sqrt{\rho}$, we have*

$$\Pr_\zeta[\exists x \in X : \zeta \cdot x \geq s] \leq |X| \cdot \frac{r}{\pi\sqrt{1-\rho}} e^{-(1+(\sqrt{\rho}-r)^2/(1-\rho))s^2/2} + 2N(rs).$$

Proof. Since X is a ρ-cluster, each $x \in X$ is of the form $x = \sqrt{\rho_x} x_0 + \sqrt{1-\rho_x} x'$ for some $\rho_x \geq \rho$. 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 $\Pr_\zeta[\exists x \in K : \zeta \cdot x \geq s]$ 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 $\zeta \cdot x_0$.

$$\begin{aligned}
\Pr_\zeta[\exists x \in X : \zeta \cdot x \geq s] &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} \Pr[\exists x \in X : \zeta \cdot \sqrt{1-\rho_x} x' \geq s - \sqrt{\rho_x}\xi] d\xi \\
&\leq 2 \int_{0}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} \Pr[\exists x \in X : \zeta \cdot \sqrt{1-\rho_x} x' \geq s - \sqrt{\rho_x}\xi] d\xi \\
&\leq 2 \left(\int_0^{rs} \frac{e^{-\xi^2/2}}{\sqrt{2\pi}} \sum_{x \in X} N\left(\frac{s - \sqrt{\rho_x}\xi}{\sqrt{1-\rho_x}} \right) d\xi + \int_{rs}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} d\xi \right) \qquad (2.17) \\
&\leq 2 \left(\int_0^{rs} \frac{e^{-\xi^2/2}}{\sqrt{2\pi}} |X|\, N\left(\frac{s - \sqrt{\rho}\xi}{\sqrt{1-\rho}} \right) d\xi + \int_{rs}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} d\xi \right) \qquad (2.18)
\end{aligned}$$

Inequality (2.17) is a union bound and inequality (2.18) can be verified by noting that the function $f_{\xi,s}(\rho) = \frac{s-\sqrt{\rho_x}\xi}{\sqrt{1-\rho_x}}$ is monotone increasing for $\rho \geq \xi^2/s^2$ (note that we are only concerned with $\xi \leq rs$), and that $\rho_x \geq \rho \geq r^2$ for all $x \in X$.

Hence, substituting $a = \xi/s$ and applying Lemma 2.2.1, we have:

$$\begin{aligned}
&\Pr_\zeta[\exists x \in X : \zeta \cdot x \geq s] \\
&\quad \leq 2rs \left(\max_{0 \leq a \leq r} \frac{e^{-a^2 s^2/2}}{\sqrt{2\pi}} |X|\, N\left(\frac{\left(1-\sqrt{\rho}a\right) s}{\sqrt{1-\rho}} \right) \right) + 2N(rs) \\
&\quad \leq rs \left(\max_{0 \leq a \leq r} |X| \frac{\sqrt{1-\rho}}{\pi s(1-\sqrt{\rho}a)} e^{-(a^2+(1-\sqrt{\rho}a)^2/(1-\rho))s^2/2} \right) + 2N(rs) \\
&\quad \leq \frac{r}{\pi\sqrt{1-\rho}} \left(\max_{0 \leq a \leq r} |X|\, e^{-(a^2+(1-\sqrt{\rho}a)^2/(1-\rho))s^2/2} \right) + 2N(rs) \\
&\quad = \frac{r}{\pi\sqrt{1-\rho}} \left(\max_{0 \leq a \leq r} |X|\, e^{-(1+(\sqrt{\rho}-a)^2/(1-\rho))s^2/2} \right) + 2N(rs) \\
&\quad = |X| \cdot \frac{r}{\pi\sqrt{1-\rho}} e^{-(1+(\sqrt{\rho}-r)^2/(1-\rho))s^2/2} + 2N(rs)
\end{aligned}$$

□

In many cases, a simpler bound (following immediately from Lemma 2.2.1) will suffice:

Corollary 2.2.4. *Let X be a ρ-cluster for some fixed constant $\rho \in (0,1)$. Then for sufficiently large s, and all $0 \leq r \leq \sqrt{\rho}$, we have*

$$\Pr_{\zeta}[\exists x \in X : \zeta \cdot x \geq s] \leq |X| \cdot \mathrm{poly}(s) N(s)^{1+(\sqrt{\rho}-r)^2/(1-\rho)} + 2N(rs).$$

Chapter 3

Improved Analysis of Graph Coloring

In this chapter we will review the SDP relaxations and rounding algorithm of Karger, Motwani and Sudan [23], and give an improved analysis for a slight variant of their algorithm. When combined with the Blum coloring tools (discussed in Section 3.7) this gives an $O(n^{0.2130})$ coloring.

3.1 The Karger-Motwani-Sudan Algorithm

We begin by describing the algorithm described in [23] for coloring 3-colorable graphs and its analysis. We will also present some notation and terminology which will be need later for a more sophisticated analysis of this and other algorithms. The KMS algorithm uses the standard approach of finding large independent sets in order to achieve a coloring with few colors. As is well-known (see, for example, [6]), to find a legal coloring using $\tilde{O}(f(n))$ colors, it suffices to have an algorithm which can find an independent set of size $n/f(n)$). (Indeed, this follows from the following simple algorithm: apply the independent set algorithm, assign all vertices in the independent set the same color, remove these vertices from the graph, update the color counter, and repeat.) Furthermore, it suffices to concentrate on the case where there is a bound Δ on the maximum degree; see Section 3.7 for how to turn such a guarantee into an algorithm whose performance is stated in terms of n (this generalizes the approach in [6, 7]).

Consider the following relaxation for k-coloring.

Definition 3.1.1. For a graph $G = (V, E)$ with vertex set $V = \{1, 2, \ldots, n\}$, a *vector k-coloring* is an assignment of unit vectors $u_1, \ldots, u_n \in \mathbb{R}^n$ to the vertices, such that:

$$\forall (i, j) \in E: \quad u_i \cdot u_j \leq -\frac{1}{k-1}. \tag{3.1}$$

The vector k-coloring is said to be *strict* when equality holds in condition (3.1).

As is shown in [23], for any $k \geq 2$, every k-colorable graph is also vector k-colorable. While these SDP relaxations may seem inherently different from the $0-1$ relaxations discussed earlier, we shall see in Section 4.1 that strict vector k-coloring has an equivalent $0-1$ formulation.

The KMS rounding algorithm takes a graph G, vector coloring $\{u_i\}$, and a *threshold parameter* $t > 0$, and outputs an independent set in G:

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

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

Figure 3.1: Algorithm **KMS**

The performance guarantee for the KMS rounding algorithm, as shown in [23], is as follows:

Theorem 3.1.2 (KMS). *There exists some* $t = t(n, \Delta) > 0$ *such that the expected size of the independent set returned by algorithm KMS* $(G, \{u_i\}, t)$ *is* $\tilde{\Omega}(\Delta^{-1/3}n)$.

In particular, this implies an $\tilde{O}(\Delta^{1/3})$-coloring. To precisely quantify the various improvements discussed later, we need the following definition.

Definition 3.1.3. Given a graph G with vector 3-coloring $\{u_i\}$ and maximum degree Δ, the threshold parameter $t > 0$ is *c-inefficient* for $(G, \{u_i\})$ if

$$\Delta \leq N(\sqrt{3}t)^{-(1+c)}.$$

Note that by Lemma 2.2.1, if $t > 0$ is exactly c-inefficient, then $N(t) = \tilde{\Theta}(\Delta^{-\frac{1}{3+3c}})$. Thus, our objective will be to find a threshold $t > 0$ with the largest possible inefficiency c for which our algorithm is guaranteed to return an independent set of size $\Omega(N(t)n)$.

Now we recall the proof of Theorem 3.1.2 from [23] – but rephrased in our terminology. To simplify the presentation, we will only consider strict vector 3-colorings $\{u_i\}$ for now.

Note that, for any choice of threshold t, for any vertex $i \in V$, $\Pr[i \in V_\zeta(t)] = N(t)$. Say a vertex is *good* for a certain value of t if in the KMS algorithm,

$$\Pr[i \text{ is eliminated} \mid i \in V_\zeta(t)] \leq 1/2, \tag{3.2}$$

and otherwise call the vertex *bad*. If i is good, then the probability it ends up in the final independent set is at least $\frac{1}{2}\Pr[i \in V_\zeta(t)] = N(t)/2$.

Now we analyze what makes a vertex good. For any $i \in V$, let $\Gamma(i)$ be its neighborhood in G. Then

$$\Pr[i \text{ is eliminated} \mid i \in V_\zeta(t)] = \Pr[\exists j \in \Gamma(i) : \zeta \cdot u_j \geq t \mid \zeta \cdot u_i \geq t]. \tag{3.3}$$

Since $\{u_i\}$ is a strict vector 3-coloring, we can introduce the following notation: for every edge $(i, j) \in E$, we will write

$$u_j = -\tfrac{1}{2}u_i + \tfrac{\sqrt{3}}{2}u'_{ij} \tag{3.4}$$

where u'_{ij} is a unit vector orthogonal to u_i. Writing $u'_{ij} = \frac{2}{\sqrt{3}}(u_j + \frac{1}{2}u_i)$, we see that for any vector ζ:

$$\zeta \cdot u_i \geq t \quad \text{and} \quad \zeta \cdot u_j \geq t \quad \Longrightarrow \quad \zeta \cdot u'_{ij} \geq \sqrt{3}t.$$

Hence the right hand side of (3.3) is bounded from above by

$$\begin{aligned}
&\Pr[\exists j : \zeta \cdot u_j \geq t \mid \zeta \cdot u_i \geq t] \\
&\qquad \leq \Pr[\exists j : \zeta \cdot u'_{ij} \geq \sqrt{3}t \mid \zeta \cdot u_i \geq t)] \\
&\qquad = \Pr[\exists j : \zeta \cdot u'_{ij} \geq \sqrt{3}t] && \text{(independence of orthogonal projections of Gaussian)} \\
&\qquad \leq \sum_{j \in \Gamma(i)} \Pr[\zeta \cdot u'_{ij} \geq \sqrt{3}t] && \text{(union bound)} \\
&\qquad \leq \Delta N(\sqrt{3}t) \\
&\qquad = \tilde{O}(\Delta N(t)^3). && \text{(by Lemma 2.2.1)}
\end{aligned}$$

Choose the threshold t so that $N(\sqrt{3}t) = \tilde{\Theta}(N(t)^3)$ is less than $1/(2\Delta)$, in which case every vertex is good. Therefore, by linearity of expectation, the output independent set has expected size at least $N(t)n/2 = \tilde{\Omega}(\Delta^{-1/3}n)$.

Let us state one corollary of the above proof that will be useful later:

Lemma 3.1.4. *Let $\{u_i\}$ be a strict vector 3-coloring. Then for all $i \in V$ we have*

$$\Pr[\exists j \in \Gamma(i) : \zeta \cdot u_j \geq t \mid \zeta \cdot u_i \geq t] \leq \Pr[\exists j \in \Gamma(i) : \zeta \cdot u'_{ij} \geq \sqrt{3}t].$$

We adapt the following definition from [5], to use normal distributions rather than random unit vectors.

Definition 3.1.5. A set of unit vectors X is an (s, δ)-*cover*, if for $\zeta \in \mathbb{R}^n$ chosen from the standard normal distribution,

$$\Pr[\exists x \in X : \zeta \cdot x \geq s] \geq \delta.$$

The cover is said to be (at most) *c-inefficient*, if $|X| \leq N(s)^{-(1+c)}$.

Note that for any set of unit vectors X we have, by union bound

$$\Pr[\exists x \in X : \zeta \cdot x \geq s] \leq \sum_{x \in X} \Pr[\zeta \cdot x \geq s] = |X| \cdot N(s).$$

So any uniform $(s, \frac{1}{2})$-cover must contain at least $\frac{1}{2} \cdot N(s)^{-1}$ vectors, by union bound. Hence, a cover is efficient when the number of vectors is only slightly larger than the minimum required. To motivate the above definition, we note that, by lemma 3.1.4, for any vertex i for which $\Pr_\zeta[(i \in V_\zeta(t)) \wedge (i \text{ is not eliminated})] \leq \frac{1}{2}N(t)$, the vectors $\{u'_{ij} \mid j \in \Gamma(i)\}$ form a (at most) c-inefficient $(\frac{1}{2}, \sqrt{3}t)$-cover. In general, we have the following corollary of Lemma 3.1.4:

Corollary 3.1.6. *If $t > 0$ is at most c-inefficient for (G, V), then for every vertex $i \in V$ which is eliminated with probability $\geq \delta$ (conditioned on $i \in V_\zeta(t)$), the vectors $\{u'_{ij} \mid j \in \Gamma(i)\}$ corresponding to the neighbors of i form a c-inefficient $(\sqrt{3}t, \delta)$-cover.*

3.2 A Variation on the KMS Algorithm

We consider the following slight variation of the KMS rounding algorithm:

KMS′$(G, \{u_i\})$

- For "all" $t > 0$,
  - Choose $\zeta \in \mathbb{R}^n$ from the n-dimensional standard normal distribution.
  - Pick any edge $(i, j) \in E$ with both endpoints in $V_\zeta(t)$, and eliminate both i and j. Repeat until no such edges are left.
  - Let $V'_\zeta(t)$ be the set of remaining vertices in $V_\zeta(t)$.
- For t which maximizes $\left|V'_\zeta(t)\right|$, return the independent set $V'_\zeta(t)$.

Figure 3.2: Algorithm **KMS′**

Remark 3.2.1. Equivalently, we can first choose ζ, and then enumerate over all relevant values of t (that is, over $t_i = \zeta \cdot u_i$). However, for the purposes of the analysis, we will consider the first formulation.

Note that the set returned by KMS′ contains the set returned by KMS (for the corresponding value of t), so Theorem 3.1.2 holds also for KMS′. The crucial difference is that in KMS′, the vertices removed from $V_\zeta(t)$ form a matching, and this will be used in the simple "pruning" argument of Lemma 3.6.10.

Recall that our goal is to find the largest possible $c > 0$ for which the threshold parameter t is c-inefficient while still guaranteeing that the algorithm find an independent set of size $\Omega(N(t)n)$. With this in mind, we now give the following explicit guarantee for the performance of KMS$'$.

Theorem 3.2.2. *For every $\tau > \frac{5}{9}$ 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 t is (at most) c-inefficient for $(G, \{u_i\})$, then KMS$'(G, \{u_i\})$ returns an independent set of size $\Omega(N(t)n)$.*

Furthermore, $c_1(\tau)$ satisfies

$$c_1(\tau) \stackrel{\text{def}}{=} \sup \left\{ c \;\middle|\; \min_{0 \leq \alpha \leq \frac{c}{1+c}} \lambda_c(\alpha) > \sqrt{\frac{1+c}{\tau}} \right\}, \tag{3.5}$$

where

$$\lambda_c(\alpha) \stackrel{\text{def}}{=} \left(3 - \alpha - 2\sqrt{1-\alpha^2}\sqrt{c}\right) / \sqrt{5 - 2\alpha - 3\alpha^2}. \tag{3.6}$$

Corollary 3.2.3. *For any n-vertex graph G with maximum degree $\leq \Delta = n^{0.6451}$, and vector 3-coloring $\{u_i\}$, KMS$'(G, \{u_i\})$ returns an ind. set of size $\Omega(\Delta^{-0.3301} n)$.*

Combining this result with the Blum coloring tools (see Theorem 3.7.2), immediately yields the following result:

Theorem 3.2.4. *Given an n-vertex 3-colorable graph, one can find an $O(n^{0.2130})$ coloring in polynomial time.*

3.3 High-Level Description of KMS$'$ Analysis

The analysis of when a vertex is good is locally tight, even though it uses the union bound. Nevertheless, we will present a nonlocal argument that shows that the local analysis cannot be simultaneously tight for all vertices for this value of t. Thus the

KMS′ algorithm can use a less efficient (smaller) threshold t than the KMS [23] paper did, which increases $\Omega(N(t)n)$, the size of the final independent set. If fewer than $n/2$ vertices are bad, the expected size of the independent set is at least $N(t)n/4$. We show below that there is such a threshold t satisfying $\Delta > N(\sqrt{3}t)^{-(1+c)}$ for some $c > 0$. Thus we can find an independent set of size $N(t)n/4 = \tilde{\Omega}(\Delta^{-1/(3+3c)}n)$, an improvement over KMS.

Our nonlocal argument is directly inspired by the "walk" argument of Arora, Rao, and Vazirani [5]. However, our walks are only of length $O(1)$ whereas theirs were longer. To illustrate our idea, let us first assume that the vectors in the SDP solution are "nondegenerate," by which we mean that their pairwise inner products do not exhibit any statistically significant patterns apart from those implied by the SDP constraints. To give an example, for any vertex $i \in V$, the constraints of strict vector coloring require the vectors u'_{ij} defined above to be orthogonal to u_i. In a nondegenerate solution, we also expect that for any arbitrary unit vector u_0, most of the vectors u'_{ij} should only have negligible projection on u_0 (Lemma 3.6.3 gives a sufficient condition for this phenomenon). As we shall see, the nondegeneracy property corresponds to a vector coloring for which the KMS [23] analysis is tight, and so we are interested in ruling out the existence of such a solution.

We give a heuristic argument why the KMS′ algorithm should return an independent set of size $\tilde{\Omega}(n^{8/9})$ in a nondegenerate solution. This implies that whenever the original analysis of [23] gives an independent set of size less than $\tilde{O}(n^{8/9})$ (i.e., when $\Delta > n^{1/3}$), the vector coloring cannot be nondegenerate, and thus some improvement is possible. We note that the integrality gaps of [16] rule out the existence of independent sets of size $\Omega(n^{0.843})$ in certain vector 3-colorable graphs. Thus, this heuristic argument cannot guarantee an independent set of size $\tilde{\Omega}(n^{8/9})$ in general.

Let t be the smallest value s.t. at least half the vertices i are bad, that is:

$$\Pr_\zeta[i \text{ gets eliminated} \mid i \in V_\zeta(t)] \geq 1/2. \tag{3.7}$$

We will show that $N(t) \geq \tilde{\Omega}\left(n^{-1/9}\right)$, and thus (since for a slightly smaller t half the vertices must also be good) KMS′ returns an independent set of expected size $\geq \frac{1}{2}N(t)n = \tilde{\Omega}\left(n^{8/9}\right)$.

First, a simple pruning argument (see Lemma 3.6.10, the only place where we use the difference between KMS and KMS′) allows us to assume that condition (3.7) holds for *all* vertices in the graph rather than just half the vertices (with the probability 1/2 replaced by a smaller constant). Hence Corollary 3.1.4 implies that for every vertex $i \in V$, the set $\{u'_{ij} \mid j \in \Gamma(i)\}$ is a $\{\sqrt{3}t, \Omega(1)\}$-cover. By symmetry, the vectors $\{-u'_{ij} \mid j \in \Gamma(i)\}$ also form a $\{\sqrt{3}t, \Omega(1)\}$-cover. The random projection $\zeta \cdot u_i$ is negligible compared to t (i.e. $o(t)$) for all but an $o(1)$ fraction of Gaussian vectors ζ (since u_i is a fixed vector) so, since $\zeta \cdot u_j = -\frac{1}{2}\zeta \cdot u_i + \frac{\sqrt{3}}{2}\zeta \cdot u'_{ij}$, we conclude that the vectors $\{-u_j \mid j \in \Gamma(i)\}$ form a $\left(\left(\frac{3}{2} - o(1)\right)t, \Omega(1)\right)$-cover. That is, with probability $\Omega(1)$ some $j \in \Gamma(i)$ satisfies $\zeta \cdot u_j \leq -\left(\frac{3}{2} - o(1)\right)t$.

Now consider all the neighbors $k \in \Gamma(j)$ of such a vertex j. By our assumption, the vectors $\{u'_{jk} \mid k \in \Gamma(j)\}$ are also a $\{\sqrt{3}t, \Omega(1)\}$-cover. Thus for most Gaussians ζ, there is some $k \in \Gamma(j)$ such that $\zeta \cdot u'_{jk} \geq \sqrt{3}t$. Such a vertex k satisfies:

$$\begin{aligned}\zeta \cdot u_k &= -\frac{1}{2}\zeta \cdot u_j + \frac{\sqrt{3}}{2}\zeta \cdot u'_{jk} \\ &\geq \frac{3(1-o(1))t}{4} + \frac{3t}{2} = \frac{9}{4}t(1-o(1)).\end{aligned}$$

Note that here we are using nondegeneracy strongly, since we are assuming that the union of the events "$[\zeta \cdot u'_{ij} \leq -\sqrt{3}t] \wedge [\exists k \in \Gamma(j) : \zeta \cdot u'_{jk} \geq \sqrt{3}t]$" (for all various $j \in \Gamma(i)$) happens with good probability when the individual events have large probability. This would be true if the events "$\zeta \cdot u'_{ij} \leq -\sqrt{3}t$" were disjoint. It turns out (see Theorem 3.4.4) that the nondegeneracy assumption is enough to make a slightly weaker claim of the same form.

One can continue this argument with neighbors of k and so on, ultimately deducing that for a constant fraction of Gaussian vectors ζ, there is some vertex l such that $\zeta \cdot u_l \geq 3t(1 - o(1))$. (In general, given that some vector in the vector coloring has a projection s on ζ, this argument shows the existence of a vector corresponding to some neighbor whose projection is $\geq (s/2 + 3t/2)(1 - o(1))$, which is larger than s so long as $s < 3t$.)

Since the number of vertices in the graph is n, the union bound implies that for Gaussian vector ζ the expected number of l such that $\zeta \cdot u_l \geq 3t$ is at most $N(3t)n$. We conclude that $N(3t)n = \Omega(1)$, and hence (by Lemma 2.2.1) $N(t)^9 n = \tilde{\Omega}(1)$.

Of course, the above analysis ignores all conditioning between the probability calculations in successive steps of the argument, which is justifiable only when the vectors are nondegenerate. In the correct argument such conditioning cannot be ignored, and so we must exploit the connection between nondegeneracy and tightness of the KMS analysis. Namely, we must show that when $\Delta > n^{1/3}$, we can introduce a small degree of inefficiency (thus increasing $N(t)$), and still use the above argument to show that most vertices are "good" (otherwise obtaining a contradiction). The extent to which one can make the threshold t inefficient without weakening the above argument too much requires a careful quantification (and a formalization of the above argument which allows for nearly-nondenerate solutions).

3.4 Details of KMS′ Analysis

In this section we prove Theorem 3.2.2 using a two-step walk analysis of KMS′. For simplicity, we will assume that the vectors comprise a strict vector 3-coloring. We will relax the strictness condition in Section 3.5. The proof is by contradiction: if $\Delta < N(\sqrt{3}t)^{1/(1+c)}$ then, as in Section 3.3, we use a chaining argument to exhibit a high-probability event that is actually very unlikely.

We adapt the definition of (s, δ)-covers from Section 3.1 to include sets of vectors which are not necessarily unit vectors.

Definition 3.4.1. A set of vectors X is a *non-uniform* (s, δ)*-cover*, if for $\zeta \in \mathbb{R}^n$ chosen from the standard normal distribution,

$$\Pr[\exists x \in X : \zeta \cdot x \geq s] \geq \delta.$$

To make the distinction explicit, we will call regular (s, δ)-covers *uniform*.

In Section 3.6 we will show that if at least half the vertices $i \in V$ are bad, then we can identify a subgraph where *every* vertex is almost-bad (see Lemma 3.6.10), and thus has a $(\sqrt{3}t, \Omega(1))$-cover associated with its neighbors. In this graph, fix a vertex $i \in V$. Note that the set $\{u'_{ij} \mid j \in \Gamma(i)$ and the sets $\{u'_{jk} \mid k \in \Gamma(j)\}$ for all $j \in \Gamma(i)$ are (uniform) $(\sqrt{3}t, \Omega(1))$-covers. We want to compose the cover $\{-u'_{ij} \mid j \in \Gamma(i)\}$ with the various covers $\{u'_{jk} \mid k \in \Gamma(j)\}$ to attain a lower bound on the probability of the event "$\exists j \in \Gamma(i), k \in \Gamma(j) : [\zeta \cdot (-u'_{ij}) \geq \sqrt{3}t] \wedge [\zeta \cdot u'_{jk} \geq \sqrt{3}t]$".

As pointed out earlier, if such a composition were always possible, the argument in Section 3.3 would contradict known integrality gaps. Hence there must be some loss (specifically, we will only attain a lower bound on the probability of the above event by considering smaller projection for the u'_{jk}). This loss will be a monotonically increasing function of c, the inefficiency of the covers. Thus, our aim is to find the largest c for which a contradiction can still be obtained. First we need the following lemma, which shows that in any uniform (s, δ) cover $\{y_k\}$, the components of vectors y_k orthogonal to a fixed vector form a non-uniform $(s - o(s), \delta - o(1))$-cover.

Lemma 3.4.2. *Let v' be a unit vector and $\{y_k\}$ be a uniform (s, δ)-cover. Rewrite each y_k as $y_k = \alpha_k v' + \sqrt{1 - \alpha_k^2} y'_k$ for $\alpha_k = v' \cdot y_k \in [-1, 1]$ and unit vector $y'_k \perp v'$. Then for all $\rho \geq 0$, we have*

$$\Pr\left[\exists k : \zeta \cdot y'_k \geq \frac{s - \rho}{\sqrt{1 - \alpha_k^2}}\right] \geq \delta - 2N(\rho).$$

Proof.

$$\begin{aligned}
\delta \leq \Pr[\exists k : \zeta \cdot y_k \geq s] &\leq \Pr[|\zeta \cdot v'| \geq \rho] + \Pr[|\zeta \cdot v'| \leq \rho \wedge \exists k : \zeta \cdot y_k \geq s] \\
&\leq 2N(\rho) + \Pr\left[\exists k : \zeta \cdot (y_k - \alpha_k v') \geq s - |\alpha_k|\, \rho\right] \\
&= 2N(\rho) + \Pr\left[\exists k : \zeta \cdot y'_k \geq \frac{s - |\alpha_k|\, \rho}{\sqrt{1 - \alpha_k^2}}\right] \\
&\leq 2N(\rho) + \Pr\left[\exists k : \zeta \cdot y'_k \geq \frac{s - \rho}{\sqrt{1 - \alpha_k^2}}\right]
\end{aligned}$$

□

The following lemma shows how one "boosts" covers via measure concentration.

Lemma 3.4.3 (measure concentration). *Let $\{y_j\}$ be a non-uniform $(s, N(\theta))$-cover. Then for any $\zeta \in \mathbb{R}^n$ having standard normal distribution, and $a \geq 0$,*

$$\Pr[\exists j : \zeta \cdot y_j \geq s - \|y_j\|\, a] \geq N(\theta - a).$$

Proof. Let $\gamma_n(\cdot)$ denote the normalized Gaussian measure on R^n. The theorem of measure concentration for Gauss space ([8], [37]) states that for any measurable set $A \subseteq \mathbb{R}^n$, if $\gamma_n(A) = N(\theta)$ for $\theta \in R$, then for any $a \geq 0$ the set $A_a = \{\zeta \mid \exists z \in \mathbb{R}^n : (\|z\| \leq a) \wedge (\zeta + z \in A)\}$ has measure at least $N(\theta - a)$.

Let $A = \{\zeta \mid \exists j : \zeta \cdot y_j / \|y_j\| \geq s / \|y_j\|\}$. By our assumption, this set has measure at least $N(\theta)$. Since $\{y_j / \|y_j\|\}$ are unit vectors, we have in this case

$$\begin{aligned} A_a &= \{\zeta \mid \exists j, z : (\|z\| \leq a) \wedge ((\zeta + z) \cdot y_j / \|y_j\| \geq s / \|y_j\|)\} \\ &= \{\zeta \mid \exists j, z : (\|z\| \leq a) \wedge (\zeta \cdot y_j / \|y_j\| \geq s / \|y_j\| - z \cdot y_j / \|y_j\|)\} \\ &= \left\{\zeta \,\middle|\, \exists j : \zeta \cdot y_j / \|y_j\| \geq s / \|y_j\| - \left(\max_{z:\|z\|\leq a} z \cdot y_j / \|y_j\|\right)\right\} \\ &= \{\zeta \mid \exists j : \zeta \cdot y_j / \|y_j\| \geq s / \|y_j\| - a\} \end{aligned}$$

Applying measure concentration, the claim follows immediately. □

We now use this lemma to prove a cover composition theorem.

Theorem 3.4.4 (Cover composition). *Let $\{x_j \mid j \in J\}$ be a uniform c-inefficient (s_1, δ)-cover, and for each $j \in J$, let Y_j be a non-uniform $(s_2, N(\theta))$-cover such that $y \perp x_j$. Then we have*

$$\Pr\left[\exists j \in J, y \in Y_j : (\zeta \cdot x_j \geq s_1) \ \wedge \ \left(\zeta \cdot y \geq s_2 - \|y\| \cdot \left(\theta + \sqrt{c(1+\varepsilon)} \cdot s_1\right)\right)\right] \geq \delta - O\left(\frac{1}{s_1}\right)$$

for some $\varepsilon = O\left(\frac{\log s_1}{s_1^2}\right)$.

Proof. If we associate with every $j \in J$ the halfspace $\{\zeta \mid \zeta \cdot x_j \geq s_1\}$, then the (s_1, δ)-cover property implies that the union of these halfspaces has Gaussian measure $\geq \delta$. The idea is to obtain an upper-bound on the measure of points in each halfspace not participating in the relevant set (i.e. the points ζ not satisfying the event in the theorem statement). Formally, for every j we define

$$Z_j = \left\{ z \in \mathbb{R}^n \mid z \cdot x_j \geq s_1 \ \wedge \ \forall y \in Y_j : z \cdot y \leq s_2 - \|y\| \cdot \left(\theta + \sqrt{c(1+\varepsilon)} \cdot s_1\right) \right\}$$

Symmetry of the standard normal distribution implies that $N(\rho) + N(-\rho) = 1$ for all $\rho \in \mathbb{R}$. Hence, by the independence of orthogonal components of a Gaussian vector, and by Lemma 3.4.3 we have

$$\begin{aligned}
\Pr[\zeta \in Z_j] &= \Pr[\zeta \cdot x_j \geq s_1] \cdot \Pr\left[\forall y \in Y_j : \zeta \cdot y \leq s_2 - \|y\| \cdot \left(\theta + \sqrt{c(1+\varepsilon)} \cdot s_1\right)\right] \\
&= N(s_1) \cdot \Pr\left[\forall y \in Y_j : \zeta \cdot y \leq s_2 - \|y\| \cdot \left(\theta + \sqrt{c(1+\varepsilon)} \cdot s_1\right)\right] \\
&\leq N(s_1) \cdot (1 - N(-\sqrt{c(1+\varepsilon)} \cdot s_1)) \\
&= N(s_1) \cdot N(\sqrt{c(1+\varepsilon)} \cdot s_1).
\end{aligned} \tag{3.8}$$

Letting $\varepsilon = \frac{2 \ln s_1}{s_1^2}$, this gives

$$\begin{aligned}
\Pr\Big[\exists j \in J, y \in Y_j : \zeta \cdot x_j \geq s_1 \\
&\wedge \ \zeta \cdot y \geq s_2 - \|y\| \cdot \left(\theta + \sqrt{c(1+\varepsilon)} \cdot s_1\right)\Big] \\
&\geq \Pr[\exists j : \zeta \cdot x_j \geq s_1] - \Pr[\exists j : \zeta \in Z_j] \\
&\geq \delta - \sum_j \Pr[\zeta \in Z_j] \\
&\geq \delta - |J| \cdot N(s_1) \cdot N(\sqrt{c(1+\varepsilon)} \cdot s_1) && \text{by (3.8)} \\
&\geq \delta - N(s_1)^{-c} \cdot N(\sqrt{c(1+\varepsilon)} \cdot s_1) && \text{by efficiency of } \{x_j\} \\
&\geq \delta - \frac{1}{\sqrt{c(1+\varepsilon)} \cdot s_1}. && \text{by Lemma 2.2.1}
\end{aligned}$$

$\square$

We can use the composition theorem to obtain a result reminiscent of the chaining argument in [5]: it shows that whenever KMS′ fails (in expectation), for some large s we can find an $(s, \Omega(1))$-cover containing few vectors.

Theorem 3.4.5. *Let $c > 0$ be any fixed constant. Then for any graph $G = (V, E)$, strict vector 3-coloring $\{u_i\}$, and threshold $t > 0$ which is at most c-inefficient for $(G, \{u_i\})$, the following holds: If at least half the vertices in V are bad for KMS′$(G, \{u_i\})$ and threshold t, then identifying the vertices V with the vector coloring $\{u_i\}$, there is some $i \in V$ and some subset of its 2-neighborhood $W \subseteq \Gamma(\Gamma(i))$ such that*

1. *For all $k \in W$ we have $u_k = (\frac{1}{4} + \frac{3}{4}\alpha_k)u_i + w_k$ for some vector $w_k \perp u_i$, and $-\frac{C}{t^2} \le \alpha_k \le \frac{c}{1+c} + \frac{C}{\log t}$ (for some universal constant $C > 0$).*
2. *With probability at least $\frac{1}{8} - O\left(\frac{1}{\log t}\right)$, there is some $k \in W$ for which*

$$\zeta \cdot w_k \ge \left(\frac{9}{4} - \frac{3}{4}\alpha_k - \frac{3}{2}\sqrt{1 - \alpha_k^2}\sqrt{c}\right) t - O(\sqrt{\log t})$$

Proof. Prune as in Lemma 3.6.11, and for simplicity, assume $G = (V, E)$ is the remaining graph. Now, fixing some $i \in V$, we have that $\{u'_{ij} \mid j \in \Gamma(i)\}$ and the sets $\{u'_{jk} \mid k \in \Gamma(j)\}$ for every $j \in \Gamma(i)$ are all uniform $\left(\sqrt{3}t, \frac{1}{8} - O\left(\frac{1}{\log t}\right)\right)$-covers which are at most c-inefficient. Moreover, there exists some constant $C > 0$ such that letting $W_j = \left\{k \in \Gamma(j) \;\middle|\; -\frac{C}{t^2} \le u'_{ji} \cdot u'_{jk} \le \frac{c}{1+c} + \frac{C}{\log t}\right\}$ for every $j \in \Gamma(i)$, the sets $\left\{u'_{jk} \mid k \in W_j\right\}$ are $(\sqrt{3}t, \Omega(\frac{1}{t^3}))$-covers.

Note that for all $k \in W_j$, $u_i \cdot u_k = (-\frac{1}{2}u_j + \frac{\sqrt{3}}{2}u'_{ji}) \cdot (-\frac{1}{2}u_j + \frac{\sqrt{3}}{2}u'_{jk}) = \frac{1}{4} + \frac{3}{4}u'_{ji} \cdot u'_{jk}$. Hence the value $\alpha_k \stackrel{\text{def}}{=} u'_{ji} \cdot u'_{jk}$ depends only on k (and i) and not on the choice of intermediate vertex j. For all $j \in \Gamma(i), k \in \Gamma(j)$ let us write

$$u'_{jk} = \alpha_k u'_{ji} + \sqrt{1-\alpha_k^2} w'_{jk}$$

for unit vector $w'_{jk} \perp u'_{ji}$. We can now define w_k as follows:

$$w_k = -\frac{\sqrt{3}}{4}(1-\alpha_k)u'_{ij} + \frac{\sqrt{3}}{2}\sqrt{1-\alpha_k^2} w'_{jk}. \tag{3.9}$$

This definition is consistent with the decomposition of u_i in part 1, as we see here:

$$\begin{aligned} u_k = -\frac{1}{2}u_j + \frac{\sqrt{3}}{2}u'_{jk} &= \frac{1}{4}u_i - \frac{\sqrt{3}}{4}u'_{ij} + \frac{\sqrt{3}}{2}u'_{jk} \\ &= \frac{1}{4}u_i - \frac{\sqrt{3}}{4}u'_{ij} + \frac{\sqrt{3}}{2}\left(\alpha_k u'_{ji} + \sqrt{1-\alpha_k^2} w'_{jk}\right) \\ &= \frac{1}{4}u_i - \frac{\sqrt{3}}{4}u'_{ij} + \frac{\sqrt{3}}{2}\left(\alpha_k\left(\frac{\sqrt{3}}{2}u_i + \frac{1}{2}u'_{ij}\right) + \sqrt{1-\alpha_k^2} w'_{jk}\right) \\ &= \left(\frac{1}{4} + \frac{3}{4}\alpha_k\right)u_i - \frac{\sqrt{3}}{4}(1-\alpha_k)u'_{ij} + \frac{\sqrt{3}}{2}\sqrt{1-\alpha_k^2} w'_{jk}. \end{aligned} \tag{3.10}$$

For each $j \in \Gamma(i)$, we now apply Lemma 3.4.2 for the cover $\left\{u'_{jk} \mid k \in W_j\right\}$, $v' = u'_{ji}$ and $\rho = N^{-1}(\frac{1}{t^4}) = O(\sqrt{\log t})$. The lemma implies that the vectors $Y_j = \left\{\sqrt{1-\alpha_k^2} w'_{jk} \;\middle|\; k \in W_j\right\}$ form a non-uniform $\left(\sqrt{3}t - \rho, \frac{C'}{t^3}\right)$-cover for some constant $C' > 0$.

By definition of w'_{jk}, we see that $w'_{jk} \perp u'_{ij}$. Indeed, we have:

$$\begin{aligned} u'_{ij} \cdot w'_{jk} &= \left(\frac{\sqrt{3}}{2} u_j + \frac{1}{2} u'_{ji} \right) \cdot w'_{jk} \\ &= \frac{\sqrt{3}}{2} u_j \cdot w'_{jk} \\ &= \frac{\sqrt{3}}{2} u_j \cdot \frac{1}{\sqrt{1-\alpha_k^2}} \left(u'_{jk} - \alpha_k u'_{ji} \right) = 0. \end{aligned}$$

Hence, applying Theorem 3.4.4 for $x_j = -u'_{ij}$, Y_j as above, and $\theta = N^{-1}\left(\frac{C'}{t^3}\right) = O\left(\sqrt{\log t}\right)$, we get, for some $\varepsilon = O\left(\frac{\log t}{t^2}\right)$,

$$\Pr\left[\begin{array}{l} \exists j \in \Gamma(i), k \in W_j : \\ \zeta \cdot u'_{ij} \leq -\sqrt{3}t \wedge \zeta \cdot w'_{jk} \geq \left(\frac{1}{\sqrt{1-\alpha_k^2}} - \sqrt{c(1+\varepsilon)} \right) \sqrt{3}t - \frac{\rho}{\sqrt{1-\alpha_k^2}} - \theta \end{array}\right] \geq \frac{1}{8} - O\left(\frac{1}{\log t}\right).$$

By (3.9), this immediately implies part 2 of the theorem statement for $W = \bigcup_{j \in \Gamma(i)} W_j$.

□

Now we prove Theorem 3.2.2 for the case of strict vector 3-coloring.

Proof of Theorem 3.2.2 (for strict vector coloring). Let c be the degree of inefficiency of the input, and suppose, for the sake of contradiction, that at least half the vertices i are bad. Applying Theorem 3.4.5, we obtain a set of at most n vectors $\{w_k\}$ so that with constant probability, some w_k has projection at least $\frac{\sqrt{3}}{4}(3 - \alpha_k - 2\sqrt{1-\alpha_k^2}\sqrt{c} - o(1))\sqrt{3}t$ for some $-o(1) \leq \alpha_k \leq \frac{c}{1+c} + o(1)$.

Let $\hat{w}_k = w_k / \|w_k\|$. Noting that $\|w_k\| = \frac{\sqrt{3}}{4}\sqrt{5 - 2\alpha_k - 3\alpha_k^2}$, by (3.6) we have

$$
\begin{aligned}
\Omega(1) &\le \Pr[\exists k : \zeta \cdot \hat{w}_k \ge (\lambda_c(\alpha_k) - o(1))\sqrt{3}t] \\
&\le \sum_k N((\lambda_c(\alpha_k) - o(1))\sqrt{3}t) \\
&\le n \cdot \max_k N((\lambda_c(\alpha_k) - o(1))\sqrt{3}t) \\
&\le n \cdot N\left(\min_{0 \le \alpha \le \frac{c}{1+c}} (\lambda_c(\alpha) - o(1)) \cdot \sqrt{3}t \right).
\end{aligned}
$$

Let $f(c) = \min_{0 \le \alpha \le \frac{c}{1+c}} \lambda_c(\alpha)$, and note that $\lim_{c \to 0} f(c) = \frac{3}{\sqrt{5}}$. Therefore, for $\tau > \frac{5}{9}$, $c_1(\tau)$ in equation (3.5) is well-defined. Moreover, if $c \le c_1(\tau) - \Omega(1)$, we have $f(c)^2 \ge (1+c) \cdot (\frac{1}{\tau} + a)$, for some constant $a > 0$. Thus, using the efficiency of t and Lemma 2.2.1, the above inequality gives

$$
\begin{aligned}
\Omega(1) &\le n \cdot N\left(\sqrt{3}t\right)^{f(c)^2 - o(1)} \\
&\le n \cdot N\left(\sqrt{3}t\right)^{(1+c)\cdot(\frac{1}{\tau} + a - o(1))} \\
&\le n \cdot \Delta^{-\frac{1}{\tau} - (a - o(1))} \\
&= \Delta^{-(a - o(1))}
\end{aligned}
$$

which is a contradiction. □

3.5 Extending the Analysis of KMS′ to Non-Strict Vector Coloring

We sketch a generalization of the analysis in Section 3.4 which applies to KMS′ when the vector 3-coloring in the input is not necessarily strict. Specifically, we prove Theorem 3.2.2 where the set $\{u_i\}$ is a non-strict vector 3-coloring.

Let us first generalize the notation from Section 3.4. For any neighboring vertices $i, j \in V$, we write

$$u_j = (u_i \cdot u_j)u_i + \sqrt{1 - (u_i \cdot u_j)^2}u'_{ij}.$$

For a "one-step analysis" (e.g. the original KMS result), it is clear that if the inner product between neighbors is $< -\frac{1}{2}$ then the analysis only improves. Specifically, we have the following easy generalization of Lemma 3.1.4.

Lemma 3.5.1. *Let $\{u_i\}$ be a non-strict vector 3-coloring, let $i \in V$ be any vertex, and for all neighbors $j \in \Gamma(i)$, let $a_j = -u_i \cdot u_j$ (note that $a_j \geq \frac{1}{2}$). Then we have*

$$\begin{aligned} \Pr[\exists i : \zeta \cdot u_j \geq t \mid \zeta \cdot u_i \geq t] &\leq \Pr[\exists i : \zeta \cdot u'_{ij} \geq \sqrt{(1 + a_j)/(1 - a_j)}t] \\ &\leq \Pr[\exists i : \zeta \cdot u'_{ij} \geq \sqrt{3}t]. \end{aligned}$$

The difficulty in verifying that the "two-step walk analysis" extends to non-strict vector coloring seems to arise when we walk from i to j to k, where $u_i \cdot u_j < u_j \cdot u_k$. This case can be avoided using a simple binning argument.

Proof of Theorem 3.2.2 for non-strict vector coloring. As before, we assume, for the sake of contradiction, that at least half the vertices are bad, and prune as in Lemma 3.6.11 (this lemma is valid for non-strict vector coloring by Lemma 3.5.1 above). For simplicity, assume $G = (V, E)$ is the remaining graph. For any neighboring vertices $i, j \in V$, define

$$V_{i,j} = \left\{ k \in \Gamma(j) \;\middle|\; -\frac{C}{t^2} \leq u'_{ji} \cdot u'_{jk} \leq \frac{c}{1+c} + \frac{C}{\log t} \right\},$$

where $C > 0$ is the constant in Lemma 3.6.11. Then, by Lemma 3.6.11, the sets $\{u'_{jk} \mid k \in V_{i,j}\}$ are $\left(\sqrt{3}t, \Omega(\frac{1}{t^3})\right)$-covers (that is, $\left(\sqrt{3}t, \Omega\left(\log^{-\frac{3}{2}} n\right)\right)$-covers).

Now, consider a partition of the edges into $\log n$ bins $\{E_l\}$ by inner product of endpoints, i.e.

$$E_l = \left\{(i,j) \in E : -u_i \cdot u_j \in \left[\frac{1}{2} + \frac{l}{\log n}, \frac{1}{2} + \frac{l+1}{\log n}\right)\right\}.$$

For any edge (i', i) there is some $l = l(i', i)$ such that the set $\left\{u'_{ij} \mid j \in \Gamma_{E_l}(i) \cap V_{i',i}\right\}$ is a $\left(\sqrt{3}t, \Omega\left(\log^{-\frac{5}{2}} n\right)\right)$-cover. We will denote this value simply as $l(i', i)$. Let i', i be two vertices that minimize $l(i', i)$. We concentrate now only on the subgraph of G induced on the following vertices: the vertex i, neighbors $j \in \Gamma_{E_{l(i',i)}}(i) \cap V_{i',i}$, and for all such vertices j, neighbors $k \in \Gamma_{E_{l(i,j)}}(j) \cap V_{i,j}$. Finally, let $a = \frac{1}{2} + \frac{l(i',i)}{\log n}$, let $b_j = \frac{1}{2} + \frac{l(i,j)}{\log n}$ and $b_{jk} = -u_j \cdot u_k$. This choice of vertices ensures the following important facts:

- For all j, $|u_i \cdot u_j + a| \leq \frac{1}{\log n}$ and for all $k \in \Gamma_{E_{l(i,j)}}(j)$, $|b_{jk} - b_j| \leq \frac{1}{\log n}$.
- For all j, k, $u_k = -b_{jk} u_j + \sqrt{1 - b_{jk}^2}\, u'_{jk}$ and $b_{jk} \geq a \geq \frac{1}{2}$.
- For all j, k, $-\frac{C}{t^2} \leq u'_{ji} \cdot u'_{jk} \leq \frac{c}{1+c} + \frac{C}{\log t}$
- The sets $\left\{u'_{ij} \mid j \in \Gamma_{E_{l(i',i)}}(i) \cap V_{i',i}\right\}$ and $\left\{u'_{jk} \mid k \in \Gamma_{E_{l(i,j)}}(j) \cap V_{i,j}\right\}$ for all j are all $(\sqrt{3}t, \tilde{\Omega}(1))$-covers.

To simplify the argument, assume $u_i \cdot u_j = -a$ for all j, and assume $u_j \cdot u_k = -b_j$ for all j, k. We can do this because the resulting error terms (projections along vectors of norm $O(\frac{1}{\log n})$) are negligible. Arguing as before, one can show the following generalization of equation (3.10)

$$u_k = \left(ab_j + \alpha_{jk}\sqrt{(1-a^2)(1-b_j^2)}\right) u_i - \left(b_j\sqrt{1-a^2} - \alpha_{jk} a \sqrt{1-b_j^2}\right) u'_{ij} + \sqrt{(1-\alpha_{jk}^2)(1-b_j^2)}\, w'_{jk}$$

where $-\frac{C}{t^2} \le \alpha_{jk} \le \frac{c}{1+c} + \frac{C}{\log t}$ and w'_{jk} is a unit vector orthogonal to u'_{ij} and u_i.

Moreover, as in Theorem 3.4.5, we can show

$$\Pr\left[\exists i,j : \left(\zeta \cdot u'_{ij} \le -\sqrt{3}\,t\right) \wedge \left(\zeta \cdot w'_{jk} \ge \rho_{jk}\sqrt{3}\,t\right)\right] = \tilde{\Omega}(1) \tag{3.11}$$

for some

$$\rho_{jk} = 1/\sqrt{1-\alpha_{jk}^2} - \sqrt{c} - o(1).$$

Let w_k be the component of u_k orthogonal to u_i. Namely,

$$w_k = -\left(b_j\sqrt{1-a^2} - \alpha_{jk}a\sqrt{1-b_j^2}\right)u'_{ij} + \sqrt{(1-\alpha_{jk}^2)(1-b_j^2)}w'_{jk}. \tag{3.12}$$

Our goal is to show that (3.11) implies a projection of at least $(\min_{j,k}\lambda_c(\alpha_{jk}) - o(1))\sqrt{3}t$ (the corresponding projection when $a = b_j = \frac{1}{2}$) for $\hat{w}_k = w_k/\|w_k\|$. The rest of the proof then follows as before. For brevity, write $\hat{w}_k = -\theta_{jk}u'_{ij} + \sqrt{1-\theta_{jk}^2}w'_{jk}$. We will denote by κ_{jk} the corresponding value of θ_{jk} when $a = b_j = \frac{1}{2}$. Thus, it suffices to show that

$$\theta_{jk} + \sqrt{1-\theta_{jk}^2}\rho_{jk} \ge \kappa_{jk} + \sqrt{1-\kappa_{jk}^2}\rho_{jk}.$$

The comparison is facilitated by the following simple observation: For any non-negative constant $0 \le \rho \le 1$, the function $f(\theta) = \theta + \sqrt{1-\theta^2}\rho$ is monotonically increasing in the range $\theta \in [-1, \frac{1}{\sqrt{2}}]$. Thus, it suffices to show the following:

1. $\rho_{jk} < 1$.
2. $\frac{\theta_{jk}}{\sqrt{1-\theta_{jk}^2}} \ge \frac{\kappa_{jk}}{\sqrt{1-\kappa_{jk}^2}} = \frac{1}{2}\cdot\frac{1-\alpha_{jk}}{\sqrt{1-\alpha_{jk}^2}}$ (equivalent to showing $\theta_{jk} \ge \kappa_{jk}$).
3. $\frac{\theta_{jk}}{\sqrt{1-\theta_{jk}^2}} < 1$ (and thus $\theta_{jk} < \frac{1}{\sqrt{2}}$).

To show property 1, first note that we always have $c < 1/9$, since $\lambda_{1/9}(0) < \sqrt{(1+1/9)/\tau}$ for all $\tau \leq 1$, thus violating the condition in (3.5). Now, for such values of c we have

$$|\alpha_{jk}| \leq \frac{c}{1+c} + o(1) < \frac{\sqrt{c}}{1+\sqrt{c}} < \frac{\sqrt{2\sqrt{c}+c}}{1+\sqrt{c}}.$$

Equivalently, this gives $1/\sqrt{1-\alpha_{jk}^2} - \sqrt{c} < 1$, and thus property 1 follows.

Property 2 can be derived as follows:

$$\begin{aligned}
\frac{\theta_{jk}}{\sqrt{1-\theta_{jk}^2}} &= \frac{b_j\sqrt{1-a^2} - \alpha_{jk}a\sqrt{1-b_j^2}}{\sqrt{(1-\alpha_{jk}^2)(1-b_j^2)}} && \text{by (3.12)} \\
&\geq \frac{b_j - \alpha_{jk}a}{\sqrt{1-\alpha_{jk}^2}} && \text{since } b_j \geq a \\
&\geq \frac{a - \alpha_{jk}a}{\sqrt{1-\alpha_{jk}^2}} \geq \frac{1-\alpha_{jk}}{2\sqrt{1-\alpha_{jk}^2}} && \text{since } b_j \geq a \geq \frac{1}{2}
\end{aligned}$$

Finally, let us show property 3. First note that for every vertex j the set $\left\{u'_{jk} \mid k \in \Gamma_{E_{l(i,j)}}(j) \cap V_{i,j}\right\}$ is a $\left(\sqrt{3}t, \Omega\left(\log^{-\frac{5}{2}} n\right)\right)$-cover. Hence by Lemma 3.5.1 and the efficiency of t, we have

$$N(\sqrt{3}t)^{-(1+c)}N\left(\sqrt{\frac{1+b_j}{1-b_j}}\right) \geq \Delta N\left(\sqrt{\frac{1+b_j}{1-b_j}}\right) = \Omega\left(\log^{-\frac{5}{2}} n\right),$$

which by Lemma 2.2.1 implies $3 + 3c \geq (1+b_j)/(1-b_j) - o(1)$, or

$$b_j \leq \frac{2+3c}{4+3c} + o(1). \tag{3.13}$$

Moreover, recall that

$$-o(1) \le \alpha_{jk} \le \frac{c}{1+c} + o(1) < \sqrt{\frac{c}{1+c}}. \tag{3.14}$$

Hence we have

$$\begin{aligned}
\frac{\theta_{jk}}{\sqrt{1-\theta_{jk}^2}} &= \frac{b_j\sqrt{1-a^2} - \alpha_{jk} a \sqrt{1-b_j^2}}{\sqrt{(1-\alpha_{jk}^2)(1-b_j^2)}} && \text{by (3.12)} \\
&\le \frac{1}{\sqrt{1-\alpha_{jk}^2}} \left(\frac{(2+3c)\sqrt{1-a^2}}{2\sqrt{3+3c}} + o(1) - \alpha_{jk} a \right) && \text{by (3.13)} \\
&\le \frac{1}{\sqrt{1-\alpha_{jk}^2}} \left(\frac{2+3c}{4\sqrt{1+c}} + o(1) - \frac{\alpha_{jk}}{2} \right) && \text{since } a \ge \frac{1}{2} \\
&\le \frac{1}{\sqrt{1-\alpha_{jk}^2}} \left(\frac{2+3c}{4\sqrt{1+c}} + o(1) \right) && \text{by (3.14)} \\
&< \frac{2+3c}{4} && \text{by (3.14)} \\
&< \frac{7}{12} < 1. && \text{since } c < \frac{1}{9}
\end{aligned}$$

□

3.6 Pruning Efficient Covers

The purpose of this section is to prove Lemma 3.6.11, a structural lemma concerning the behavior of algorithm KMS′ which is used in the current chapter as well as Chapter 4.

3.6.1 De-Clustering Efficient Covers

We begin by showing that efficient covers cannot contain large clusters (see Definition 2.2.2). It is instructive to consider the following example: Let $x_0, x_1, \ldots, x_k$ be mutually orthogonal unit vectors, and $y_i = ax_0 + \sqrt{1-a^2}x_i$ for all $i = 1, \ldots, k$. A simple calculation shows that for $a = \sqrt{\frac{c}{1+c}} - o(1)$ and $k = N(s)^{-(1+c)}$, the vectors $\{y_i\}$ form a c-inefficient $(s, \Omega(1))$-cover. Note the following properties of this cover:

1. The vectors $\{y_i\}$ are a $(c/(1+c) - o(1))$-cluster.

2. For any $i \neq j$ we have $y_i \cdot y_j \approx \frac{c}{1+c}$.

Essentially, we will show that in terms of these properties, this is the most clustered configuration of vectors that can form a c-inefficient cover. Specifically, only small subsets of a c-inefficient cover can form a λ-cluster for λ significantly greater than $\frac{c}{1+c}$. Moreover, we show that for most vectors y_i in a c-inefficient cover, at most a small fraction of other cover vectors y_j satisfy $y_i \cdot y_j \geq \frac{c}{1+c} + o(1)$.

To make the quantification more precise, and to facilitate the pruning arguments later, we need the following definition, which extends the notion of (s, δ)-covers defined in Section 3.1.

Definition 3.6.1. Given a set of unit vectors X together with some measure μ on this set, we call (X, μ) a (s, δ)-*packing*, if

1. for $\zeta \in \mathbb{R}^n$ chosen from the standard normal distribution, and any $X' \subseteq X$,

$$\mu(X') \leq \Pr[\exists x \in X' : \zeta \cdot x \geq s],$$

   and

2. the measure of all vectors $\mu(X) \geq \delta$.

Note that if (X, μ) is an (s, δ)-packing then X is an (s, δ)-cover, and conversely, every cover has a corresponding packing for an appropriate choice of μ. Namely, if X is an (s, δ)-cover, then taking $Z(x)$ to be the Voronoi regions $Z(x) = \{\zeta \mid (\zeta \cdot x \geq s) \wedge (\zeta \cdot x = \max_{x' \in X} \zeta \cdot x')\}$, we can take $\mu(x) = \Pr[\zeta \in Z(x)]$. We'll say a packing (X, μ) is *c-inefficient* if the cover X is c-inefficient.

The next definition formalizes the "well-spread" property:

Definition 3.6.2. An (s, δ)-packing (X, μ) is said to be (λ, p)-*spread*, if for all $x \in X$ we have,

$$\mu(\{x' \in X \mid x \cdot x' \geq \lambda\}) \leq p.$$

The following lemma shows that in an efficient cover, vectors which form a $(c/(1+c) + o(1))$-cluster have a negligible contribution.

Lemma 3.6.3. *Let $\varepsilon > 0$, let X be an (s, δ)-cover which is at most c-inefficient, and let v_0 be any unit vector. Then letting $X' = \left\{x \in X \mid x_0 \cdot x \geq \sqrt{\frac{c}{1+c}(1+\varepsilon)}\right\}$, we have*

$$\Pr_\zeta[\exists x \in X' : \zeta \cdot x \geq s] \leq O(N(c'\varepsilon s))$$

for $c' = \min\{4\sqrt{c/(1+c)}, \sqrt{c(1+c)}/4\}$.

(The value of c' above is unimportant. It is some constant which depends only on c.)

Corollary 3.6.4. *Let (X, μ) be an (s, δ)-packing which is at most c-inefficient. Then for any $\varepsilon > 0$ this packing is $\left(\sqrt{\frac{c}{1+c}(1+\varepsilon)}, O(N(c'\varepsilon s))\right)$-spread.*

Proof of Lemma 3.6.3. The set X' is a $\rho = \frac{c}{1+c}(1+\varepsilon)$-cluster. Thus, for any constant $\theta > 0$, from Corollary 2.2.4 and the efficiency of X', we know that

$$\Pr_\zeta[\exists x \in X' : \zeta \cdot x \geq s] \leq \mathrm{poly}(s) N(s)^{-(1+c)} N(s)^{1+(\sqrt{\rho}-\theta\sqrt{c\varepsilon})^2/(1-\rho)} + 2N(\theta\sqrt{c}\varepsilon s).$$

Hence, by Lemma 2.2.1, it suffices to find some $\theta \geq \min\left\{\frac{4}{\sqrt{1+c}}, \frac{\sqrt{1+c}}{4}\right\}$ for which

$$(\sqrt{\rho} - \theta\sqrt{c\varepsilon})^2/(1-\rho) - c > \min\left\{\frac{16}{1+c}, \frac{1+c}{16}\right\} c\varepsilon^2.$$

Indeed, we have

$$\begin{aligned}
\frac{1}{c}\left(\frac{(\sqrt{\rho} - \theta\sqrt{c\varepsilon})^2}{1-\rho} - c\right) &= \frac{1}{1-c\varepsilon}\left((1+c)\varepsilon - 2\theta\sqrt{(1+c)(1+\varepsilon)}\varepsilon + \theta^2(1+c)\varepsilon^2\right) \\
&\geq (1+c)\varepsilon - 2\theta\sqrt{1+c}\varepsilon - \theta\sqrt{1+c}\varepsilon^2 + \theta^2(1+c)\varepsilon^2 \\
&= \sqrt{1+c}\left(\sqrt{1+c} - 2\theta - (\theta - \theta^2\sqrt{1+c})\varepsilon\right)\varepsilon.
\end{aligned}$$

Now, if $\varepsilon \leq 1$, we set $\theta = \frac{\sqrt{1+c}}{4}$, which gives

$$\begin{aligned}
\frac{1}{c}\left(\frac{\left(\sqrt{\rho} - \theta\sqrt{c\varepsilon}\right)^2}{1-\rho} - c\right) &\geq \sqrt{1+c}\left(\frac{1}{2}\sqrt{1+c} - \frac{\sqrt{1+c}}{4}\left(1 - \frac{1+c}{4}\right)\varepsilon\right)\varepsilon \\
&\geq \frac{5(1+c)}{16}\varepsilon \geq \frac{5(1+c)}{16}\varepsilon^2 \\
&> \frac{1+c}{16}\varepsilon^2.
\end{aligned}$$

Otherwise, if $\varepsilon > 1$, we set $\theta = \frac{4}{\sqrt{1+c}}$, which gives

$$\begin{aligned}
\frac{1}{c}\left(\frac{(\sqrt{\rho} - \theta\sqrt{c\varepsilon})^2}{1-\rho} - c\right) &\geq \sqrt{1+c}\left(\sqrt{1+c} - \frac{8}{\sqrt{1+c}} + \frac{12}{\sqrt{1+c}}\varepsilon\right)\varepsilon \\
&\geq (12\varepsilon - 7)\varepsilon \geq 5\varepsilon^2 \\
&> \varepsilon^2 \geq \min\left\{\frac{16}{1+c}, \frac{1+c}{16}\right\}\varepsilon^2.
\end{aligned}$$

□

We need the following combinatorial lemma:

Lemma 3.6.5. *Let μ be a measure on some set X satisfying $\mu(X) \leq 1$, and let $X_1, \ldots, X_k$ be any k subsets of X satisfying $\sum_{j=1}^{k} \mu(X_j) \geq 2\alpha k$ where $\alpha k \in \mathbb{N}$. Then if $S \subseteq \{1, \ldots, k\}$ is a random subset sampled uniformly over all subsets of cardinality αk, we have $\mathbb{E}\left[\mu\left(\bigcap_{j \in S} X_j\right)\right] \geq \binom{2\alpha k}{\alpha k} / \binom{k}{\alpha k}$. In particular, there exists a set $S_0 \subseteq \{1, \ldots, k\}$ of cardinality αk such that $\mu\left(\bigcap_{j \in S_0} X_j\right) \geq \binom{2\alpha k}{\alpha k} / \binom{k}{\alpha k}$.*

Proof. First note that the set $Y = \{x \in X \mid |\{j \mid X_j \ni x\}| \geq \alpha k\}$ has measure $\mu(Y) > \alpha$. Indeed,

$$\begin{aligned} 2\alpha k \leq \sum_{j=1}^{k} \mu(X_j) &= \sum_{x \in X} \mu(x) \, |\{j \mid X_j \ni x\}| \\ &= \sum_{x \in Y} \mu(x) \, |\{j \mid X_j \ni x\}| + \sum_{x \in X \setminus Y} \mu(x) \, |\{j \mid X_j \ni x\}| \qquad (3.15) \\ &\leq \mu(Y) k + \alpha k - 1. \end{aligned}$$

Now, let $\lambda = \frac{1}{\mu(Y)} \sum_{x \in Y} \mu(x) \, |\{j \mid X_j \ni x\}|$ (note that, by definition of Y, $\lambda \geq \alpha k$). By (3.15), we have

$$\mu(Y)\lambda + (1 - \mu(Y))(\alpha k - 1) \geq 2\alpha k,$$

and so

$$\mu(Y) \geq \frac{\alpha k + 1}{\lambda - \alpha k + 1}. \qquad (3.16)$$

Since $\mu(Y) \leq 1$, this also implies $\lambda \geq 2\alpha k$.

We show the required lower bound by summing over all subsets of $\{X_1, \ldots, X_k\}$

of cardinality αk. Scaling by $1/\mu(Y)$, we have

$$\frac{1}{\mu(Y)} \sum_{\substack{S \subseteq \{1,\ldots,k\} \\ |S| = \alpha k}} \mu\left(\bigcap_{j \in S} X_j\right) = \frac{1}{\mu(Y)} \sum_{x \in Y} \mu(x) \binom{|\{j \mid X_j \ni x\}|}{\alpha k}$$

$$\geq \binom{\frac{1}{\mu(Y)} \sum_{x \in Y} \mu(x)\, |\{j \mid X_j \ni x\}|}{\alpha k} = \binom{\curvearrowright}{\alpha k},$$

where the inequality follows from the convexity of the function $f(x) = \binom{x}{\alpha k} = \frac{1}{k!} \prod_{i=0}^{\alpha k - 1} (x - i)$ for $x \geq \alpha k$. Hence, we have

$$\begin{aligned}
\sum_{\substack{S \subseteq \{1,\ldots,k\} \\ |S| = \alpha k}} \mu\left(\bigcap_{j \in S} X_j\right) &\geq \mu(Y) \binom{\curvearrowright}{\alpha k} \\
&\geq \frac{\alpha k + 1}{\curvearrowright - \alpha k + 1} \binom{\curvearrowright}{\alpha k} && \text{by (3.16)} \\
&= \frac{\alpha k + 1}{\alpha k} \binom{\curvearrowright}{\alpha k - 1} \\
&\geq \frac{\alpha k + 1}{\alpha k} \binom{2\alpha k}{\alpha k - 1} \\
&= \binom{2\alpha k}{\alpha k}.
\end{aligned}$$

□

Remark 3.6.6. This lemma is tight. Consider the uniform distribution over subsets $T \subseteq \{1, \ldots, k\}$ of cardinality $2\alpha k$. For all $i \in \{1, \ldots, k\}$ let X_i be the event $i \in T$. Clearly, $\Pr[X_i] = 2\alpha$. On the other hand, for any set $S \subseteq \{1, \ldots, k\}$ of cardinality αk, we have

$$\Pr\left[\bigcap_{i \in S} X_i\right] = \Pr[S \subseteq T] = \binom{(1-\alpha)k}{\alpha k} \Big/ \binom{k}{2\alpha k} = \binom{2\alpha k}{\alpha k} \Big/ \binom{k}{\alpha k}.$$

We will use a slightly weaker bound which gives a simpler expression:

Corollary 3.6.7. *Let μ be a measure on some set X satisfying $\mu(X) \leq 1$, and let $X_1, \dots, X_k$ be any k subsets of X satisfying $\sum_{j=1}^{k} \mu(X_j) \geq 2\alpha k$ where $\alpha k \in \mathbb{N}$. Then there exists a set $S \subseteq \{1, \dots, k\}$ of cardinality αk such that $\mu\left(\bigcap_{j \in S} X_j\right) \geq \left(\frac{4\alpha}{e}\right)^{\alpha k}$.*

Proof. By Lemma 3.6.5, there exists such a set S with

$$\begin{aligned} \mu\left(\bigcap_{j \in S} X_j\right) &\geq \binom{2\alpha k}{\alpha k} \Big/ \binom{k}{\alpha k} \\ &= \prod_{i=0}^{\alpha k - 1} \frac{2\alpha k - i}{k - i}. \end{aligned}$$

We bound this expression from below by taking the natural logarithm:

$$\begin{aligned} \ln\left(\prod_{i=0}^{\alpha k - 1} \frac{2\alpha k - i}{k - i}\right) &= \sum_{i=0}^{\alpha k - 1} (\ln(2\alpha k - i) - \ln(k - i)) \\ &\geq \int_0^{\alpha k} (\ln(2\alpha k - x) - \ln(k - x)) dx \\ &= \alpha k \ln\left(4\alpha(1 - \alpha)^{\frac{1}{\alpha} - 1}\right) \\ &\geq \alpha k \ln\left(\frac{4\alpha}{e}\right). \end{aligned}$$

□

The following boosting lemma is the main tool in this pruning argument. It shows that efficient $(\lambda, o(1))$-spread packings contain large subpackings which are roughly $\left(\sqrt{\lambda \cdot \frac{c}{1+c}}, o(1)\right)$-spread. Applying this lemma repeatedly will yield the desired result (a large subpacking which is $\left(\frac{c}{1+c} + o(1), o(1)\right)$-spread).

Lemma 3.6.8. *Let (X, μ) be an (s, δ)-packing which is at most c-inefficient and (λ, p)-spread. Then for any $\sigma \geq 1$, $\varepsilon > 0$ and $\alpha > \frac{e}{4}(C \cdot N(c'\varepsilon s))^{1/\sigma}$ (for some universal constant C) there is a subset $X' \subseteq X$ such that (X', μ) is an $(s, \delta - \frac{\sigma}{\alpha} \cdot p)$-packing, and furthermore this packing is $\left(\sqrt{\lambda \cdot \frac{c}{1+c}(1+\varepsilon)\left(1+\frac{1}{\lambda\sigma}\right)}, 2\alpha\right)$-spread.*

Proof. For every $x \in X$, define

$$X_x = \left\{ x' \in X \,\middle|\, x \cdot x' \geq \sqrt{\lambda \cdot \frac{c}{1+c}(1+\varepsilon)\left(1+\frac{1}{\lambda\sigma}\right)} \right\}.$$

Now, while there are any vectors $x \in X$ such that $\mu(X_x\} \geq 2\alpha$, choose such an x and remove all $x' \in X$ such that $x \cdot x' \geq \lambda$ from the cover X (but not from the various sets X_x). Add the vector x to some (initially empty) set T. Since X is (λ, p)-spread, if we repeat the above procedure k times, the removed vectors can have measure (μ) at most kp. If the above procedure terminates after at most $\frac{\sigma}{\alpha}$ steps, the claim follows (where X' is the set of vectors remaining in X). Suppose by contradiction, that the procedure does not terminate after the first $\frac{\sigma}{\alpha}$ steps, and consider the set T of the various vectors chosen at different steps ($|T| = \frac{\sigma}{\alpha}$). By Corollary 3.6.7 (for $k = \frac{\sigma}{\alpha}$), there is some set $S \subseteq T$ of cardinality σ such that the set $X'' = \bigcap_{x \in S} X_x$ has measure

$$\mu(X'') \geq \left(\frac{4\alpha}{e}\right)^{\sigma}. \tag{3.17}$$

Now let $v = \sum_{x \in S} x$. The pruning ensures that for every two vectors $x_1, x_2 \in T$, we have $x_1 \cdot x_2 \leq \lambda$. Hence $\|v\|^2 \leq \sigma + \sigma(\sigma - 1)\lambda$. Letting $\hat{v} = v / \|v\|$, this implies that for all $x' \in X''$,

$$\hat{v} \cdot x' \geq \frac{\sigma\sqrt{\lambda \cdot \frac{c}{1+c}(1+\varepsilon)(1+\frac{1}{\lambda\sigma})}}{\sqrt{\sigma + \sigma(\sigma-1)\lambda}}$$
$$> \sqrt{\frac{c}{1+c}(1+\varepsilon)}.$$

By Lemma 3.6.3, this is a contradiction if $\mu(X'') \geq \Omega(N(c'\varepsilon s))$, which by (3.17) holds for $\alpha > \frac{e}{4}\Omega\left(C \cdot N(c'\varepsilon s)\right)^{1/\sigma}$. □

Now, by choosing appropriate parameters and boosting repeatedly, we can obtain the desired result.

Theorem 3.6.9. *Let (X, μ) be an (s, δ)-packing which is at most c-inefficient. Then there is a subset $X' \subseteq X$ such that (X', μ) is an $(s, \delta - \frac{1}{\log s})$-packing which is $\left(\frac{c}{1+c} \cdot (1 + O(\frac{1}{\log s})), e^{-\Omega(\log^2 s)}\right)$-spread.*

Proof. For all $k = 1, \dots, \log s$, define

$$\lambda_k = \left(\frac{c}{1+c}\right)^{1-2^{-k}} \left(1 + \frac{\log s}{s}\right)^{2^{-k}} \left[\left(1 + \frac{\log^{3/2} s}{s}\right)\left(1 + \frac{1+c}{c\log s}\right)\right]^{1-2^{1-k}}.$$

In particular $\lambda_1 = \sqrt{\frac{c}{1+c}\left(1 + \frac{\log s}{s}\right)}$. By Corollary 3.6.4, we know that (X, μ) is (λ_1, p_1)-spread, where $p_1 = C \cdot N(c' \log s)$ (for some constant $C > 0$). We extend this to the sequence $p_k = (2\log^3 s)^{k-1} \cdot p_1$ for all $k = 1, \dots, \log s$. Note that, by Lemma 2.2.1, for all $k \geq 2$ we have

$$p_k \geq p_2 = 2C \log^3 s \cdot N(c' \log s) = \omega(N(c' \log^{3/2} s)^{1/\log s}).$$

Moreover, $\lambda_k \geq \frac{c}{1+c}$ for every $k \geq 1$, and hence,

$$\lambda_{k+1} \geq \sqrt{\lambda_k \cdot \frac{c}{1+c}\left(1+\frac{\log^{3/2} s}{s}\right)\left(1+\frac{1}{\lambda_k \log s}\right)}.$$

Thus we can apply Lemma 3.6.8 inductively for $\log s$ steps with (at step k) $\lambda = \lambda_k$, $p = p_k$, $\sigma = \log s$, $\varepsilon = \log^{3/2} s/s$, and $\alpha = p_{k+1}/2$. At each step, we may lose at most measure $\log s \cdot \frac{p_k}{p_{k+1}/2} = \frac{1}{\log^2 s}$. Hence, after $\log s$ steps, we only lose measure $\frac{1}{\log s}$, giving a $(s, \delta - \frac{1}{\log s})$-packing, where X' is the set of vectors remaining. The spread property follows by noting that (using Lemma 2.2.1)

$$p_{\log s+1} = C(2\log^3 s)^{\log s} \cdot N(c' \log s) = e^{-\Omega(\log^2 s)}.$$

□

3.6.2 Pruning

Returning to the analysis of KMS′, we first argue that if at least half the vertices $i \in V$ are bad (i.e., their probability of being eliminated from $V_\zeta(t)$ is more than 1/2), then we can focus on a subgraph in which all vertices are almost-bad.

Lemma 3.6.10. *For any $t, \delta > 0$, if in KMS′$(G, \{u_i\})$ we have*

$$\Pr[i \text{ is eliminated} \mid i \in V_\zeta(t)] \geq \delta$$

for at least $n/2$ vertices $i \in V$, then there is a non-empty subgraph $G' = (V', E')$ of G such that for all $i \in V'$ we have

$$\Pr[i \text{ is eliminated with a neighbor in } G' \mid i \in V_\zeta(t)] \geq \delta/4.$$

Proof. The proof uses a pruning argument from [5]. Consider the graph of all matching edges (for all choices of ζ), with edges weighted by probability of elimination. Denote by $\gamma(i)$ the total weight of edges incident to i. Remove, one after the other, vertices with $\gamma(i) < N(t)\delta/4$ (while updating $\gamma(\cdot)$ values for vertices with removed neighbors). Since the initial total edge weight in the graph is $\frac{1}{2}\sum_{i\in V}\gamma(i) \geq \frac{1}{2}\cdot\frac{n}{2}\cdot N(t)\delta = n\cdot N(t)\delta/4$, and the total edge weight eliminated is $< n\cdot N(t)\delta/4$, there must be positive edge weight left. The remaining graph is therefore non-empty and has the desired property. □

In the analysis of $\text{KMS}'(G,\{u_i\})$, we need a pruning argument as above, but we also need the spread property of Theorem 3.6.9 to hold for all covers. This is guaranteed by the following lemma.

Lemma 3.6.11. *For any $t,\delta > 0$, if in KMS′$(G,\{u_i\})$ we have*

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

for at least $n/2$ vertices $i \in V$, and t is at most 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. $\Pr[j \textit{ is eliminated with a neighbor in } G' \mid j \in V_\zeta(t)] \geq \frac{\delta}{4} - O\left(\frac{1}{\log t}\right)$.
2. *For every $i \in \Gamma_{G'}(j)$ the set*

$$W'_{ji} \stackrel{\text{def}}{=} \left\{ u'_{jk} \,\middle|\, u'_{ji}\cdot u'_{jk} \leq \frac{c}{1+c}\cdot\left(1+\frac{C}{\log t}\right)\right\}$$

is a $\left(\sqrt{3}t, \frac{\delta}{4} - O\left(\frac{1}{\log t}\right)\right)$-cover.

3. For every $i \in \Gamma_{G'}(j)$ the set

$$W''_{ji} \stackrel{\text{def}}{=} \left\{ u'_{jk} \,\middle|\, -\frac{C}{t^2} \le u'_{ji} \cdot u'_{jk} \le \frac{c}{1+c} \cdot \left(1 + \frac{C}{\log t}\right) \right\}$$

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

Proof. Fix any vertex $j \in V$, and let $p_j = \Pr[j \text{ is eliminated} \mid j \in V_\zeta(t)]$. Define measure μ_j on the set $\Gamma(j)$ as follows: For all $K \subseteq \Gamma(j)$, let

$$\mu_j(K) = \Pr[\exists k \in K : j \text{ is eliminated along with } k \mid j \in V_\zeta(t)].$$

By Lemma 3.1.4 (and since KMS$'$ eliminates a matching), it follows that $(\{u'_{jk} \mid k \in \Gamma(j)\}, \mu_j)$ is a $(\sqrt{3}t, p_j)$-packing. By choice of t, this packing is c-inefficient. Now apply Theorem 3.6.9 to obtain a subset $K'_j \subseteq \Gamma(j)$ s.t. $\left(\left\{u'_{jk} \mid k \in K'_j\right\}, \mu_j\right)$ is a $\left(\frac{c}{1+c}\left(1 + \frac{C}{\log t}\right), e^{-C_1 \log^2 t}\right)$-spread $\left(\sqrt{3}t, p_j - \frac{1}{\log(\sqrt{3}t)}\right)$-packing (for some constants $C, C_1 > 0$). Remove all edges (j,k) for $k \notin K'_j$ from G (removing additional edges does not change the spread property of any packings). Note that for some $C' > 0$ and all $t > 0$ we have $e^{-C_1 \log^2 t} \le \frac{C'}{t^3}$, so the packing $\left(\{u'_{jk} \mid k \in K'_j\}, \mu_j\right)$ is $\left(\frac{c}{1+c}\left(1 + \frac{C}{\log t}\right), \frac{C'}{t^3}\right)$-spread.

Again, consider the graph of matching edges, and extend the measures μ_j to edge weights as follows: for edge (j,k), let $\mu(j,k) = \mu_j(k)$. Denote by $\mu(j)$ the total weight of edges incident to j (thus the probability of j being eliminated is exactly $\mu(j)N(t)$ and an edge (j,k) is eliminated with probability $\mu(j,k) \cdot N(t)$). To obtain the subgraph $G' = (V', E')$ repeatedly perform the following pruning operation (while possible):

1. If any vertex j has (weighted) degree $\mu(j) < \frac{\delta}{4} - \frac{1}{4\log(\sqrt{3}t)} - \frac{2C'}{t} - \frac{2C'}{t^3}$, eliminate j.

2. Otherwise, if for some vertex j and neighbor $i \in \Gamma_{G'}(j)$ the edge set

$$E_i(j) = \left\{ (j,k) \;\middle|\; u'_{ji} \cdot u'_{jk} \geq -\frac{1}{t^2} \right\}$$

satisfies $\mu(E_i(j)) < \frac{2C'}{t^3}$, remove all edges in $E_i(j)$.

To see that the remaining graph is non-empty, consider a vertex j for which we prune some edges $\gimel$ times in step 2 for some $\gimel \geq 1$. Denote by $i_1, \ldots, i_\gimel \in \Gamma_G(j)$ the neighbors for which edge sets $E_{i_s}(j)$ were pruned. By definition of the sets $E_i(j)$, we have

$$0 \leq \left\| \sum_{s=1}^{\gimel} u'_{ji_s} \right\|^2 = \gimel + \sum_{s=1}^{\gimel} \sum_{\substack{l=1 \\ l \neq s}}^{\gimel} u'_{ji_s} \cdot u'_{ji_l} < \gimel - \gimel(\gimel - 1) \cdot \frac{1}{t^2},$$

and so $\gimel < t^2 + 1$. The total weight eliminated is therefore strictly less than

$$n\left(\frac{\delta}{4} - \frac{1}{4\log(\sqrt{3}t)} - \frac{2C'}{t} - \frac{2C'}{t^3}\right) + n(t^2+1)\frac{2C'}{t^3} = n\left(\frac{\delta}{4} - \frac{1}{4\log(\sqrt{3}t)}\right),$$

whereas the total edge weight in the original graph is

$$\frac{1}{2}\sum_{x \in V} \mu(x) \geq \frac{1}{2} \cdot \frac{n}{2}\left(\delta - \frac{1}{\log(\sqrt{3}t)}\right).$$

Thus, the remaining graph is non-empty.

Define

$$K'_{ji} \stackrel{\text{def}}{=} \left\{ k \in \Gamma_{G'}(j) \;\middle|\; u'_{ji} \cdot u'_{jk} \leq \frac{c}{1+c} \cdot \left(1 + \frac{C}{\log t}\right) \right\}$$

and

$$K''_{ji} \stackrel{\text{def}}{=} \left\{ k \in \Gamma_{G'}(j) \;\middle|\; -\frac{C}{t^2} \leq u'_{ji} \cdot u'_{jk} \leq \frac{c}{1+c} \cdot \left(1 + \frac{C}{\log t}\right) \right\}.$$

Then we have the following:

- By pruning step (1):

$$\mu(j) \geq \frac{\delta}{4} - \frac{1}{4\log(\sqrt{3}t)} - \frac{2C'}{t} - \frac{2C'}{t^3}.$$

- By the spread property of $\big(\{u'_{jk} \mid k \in K'_j\}, \mu_j\big)$:

$$\mu_j(K'_{ji}) \geq \mu(j) - \frac{C'}{t^3}.$$

- By pruning step (2) and the spread property of $\big(\{u'_{jk} \mid k \in K'_j\}, \mu_j\big)$:

$$\mu_j(K''_{ji}) \geq \mu_j(E_i(j)) - \mu_j(K'_j \setminus K'_{ji})) \geq \frac{2C'}{t^3} - \frac{C'}{t^3} = \frac{C'}{t^3}.$$

This completes the proof.

□

3.7 A Modified Blum Karger Algorithm

In this section we give a summary of the technique of Blum and Karger [7], which relies on the coloring tools of Blum [6]. This allows us to present the results in Chapters 3 and 4 in the same framework as those of [6, 7]. In order to explain the approach, we use the notion of progress towards a coloring, as defined in [6].

Definition 3.7.1. For an n-vertex 3-colorable graph G, and monotonically increasing function $f : \mathbb{N} \to \mathbb{N}$, we define *progress towards a $f(n)$-coloring* as finding any one of the following objects:

Progress Type 1 An independent set of size $\Omega(n/f(n))$.

Progress Type 2 An independent set S which has a neighborhood of size

$$\left|\bigcup_{v\in S}\Gamma(v)\right| = O(|S|\, f(n)).$$

Progress Type 3 Two vertices that must have the same color in any legal 3-coloring of G.

The main result of this section (which we prove in Section 3.7.2) is the following.

Theorem 3.7.2. *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 makes progress towards an $f(n,\Delta)$-coloring. 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.*

This immediately implies the results of Blum [6] and Blum and Karger [7]:

Corollary 3.7.3. *For 3-colorable graphs, using the greedy $\Delta+1$-coloring approach, one can find an $\tilde{O}(n^{3/8})$ coloring in polynomial time.*

Corollary 3.7.4. *For 3-colorable graphs, using the KMS guarantee of progress towards an $\tilde{O}(\Delta^{1/3})$-coloring in Theorem 3.1.2, one can find an $\tilde{O}(n^{3/14})$ coloring in polynomial time.*

3.7.1 Making Progress in Dense Graphs

We will need the following (slightly simplified) lemma from [6].

Lemma 3.7.5. *Let $f(n)$ be any monotonically increasing function, then in order to find an $\tilde{O}(f(n))$-coloring in any n-vertex 3-colorable graph G, it suffices to have an algorithm which makes progress towards a $f(n)$-coloring in such a graph.*

We follow Blum and Karger [7] to guarantee progress towards an $\tilde{O}\left(\left(\frac{n}{d_{\min}}\right)^{3/5}\right)$-coloring (where $d_{\min}$ is the average degree). As in [6], for any vertex $v \in V$ and any two sets $S, T \subseteq V$, we define

- $d_T(v) \stackrel{\text{def}}{=} |\Gamma(v) \cap T|$
- $D(S) \stackrel{\text{def}}{=} \sum_{v \in S} d(v)$
- $D_T(S) \stackrel{\text{def}}{=} \sum_{v \in S} d_T(v)$
- $d_{\text{avg}}(S) \stackrel{\text{def}}{=} \frac{1}{|S|} D(S)$

Assuming G is 3-colorable, let (R, Y, B) be some partition of V into independent sets (a legal 3-coloring), where R is the color set with the most incident edges, i.e. we assume $D(R) \geq \frac{1}{3} D(V)$. Finally, we use the following partitions of V: for any set $S \subseteq V$, $\delta > 0$, and $i, j = 0, 1, 2, \ldots, \log_{1+\delta} n$, we define

- $I_i^\delta \stackrel{\text{def}}{=} \{v \in V \mid (1+\delta)^i \leq d(v) < (1+\delta)^{i+1}\}$
- $J_j^\delta(S) \stackrel{\text{def}}{=} \{v \in V \mid (1+\delta)^j \leq d_S(v) < (1+\delta)^{j+1}\}$

As in [7], we use the following three theorems of [6] (in slightly simplified form):

Theorem 3.7.6 ([6], Theorem 7). *Given a 3-colorable n-vertex graph $G = (V, E)$, there is some $v \in R$ and some $i \in \{0, 1, \ldots, \log_{1+\delta} n\}$ s.t.*

1. $|S| \geq \frac{\delta^2}{\log_{1+\delta} n} d_{\text{avg}}(R)$
2. $D_R(S) \geq \frac{1}{2}(1 - 3\delta) D(S)$

Theorem 3.7.7 ([6], Theorem 8). *Given a 3-colorable n-vertex graph $G = (V, E)$, and $\lambda \in [0, 1]$: For any set $S \subseteq V$ such that $D_R(S) \geq \lambda D(S)$, there is some $j < \log_{1+\delta} n$ such that the set $T = J_j^\delta(S)$ satisfies the following:*

1. $D_T(S) \geq \delta D_R(S) / \log_{1+\delta} n$

2. $|T \cap R| / |T| \geq (1 - 2\delta)\lambda$

Theorem 3.7.8 ([6], Theorem 13). *Given sets of vertices S and T in an n-vertex 3-colorable graph G, and $k > 0$, in order to make progress towards an $O(k)$-coloring of G, it suffices to have the following conditions hold:*

1. *S is 2-colored under some legal 3-coloring of G,*

2. *$D_T(S) = \Omega(|S| (n \log^2 n)/k^2)$, and*

3. *the following bound holds:*

$$[D_T(S)]^3 = \Omega\left(\left[|S| + \max_{v \in S} d_T(v)\right] \cdot \left[|S| |T|^2 (n \log n)/k^2 + |S|^2 |T| n^2/k^4\right]\right).$$

We now combine these theorems to prove a generalization of Theorem 6 in [7]. The proof follows essentially the same lines as in [7].

Theorem 3.7.9. *For any 3-colorable graph $G = (V, E)$ with minimum degree $d_{\min}$, there is a poly-time algorithm to make progress towards an $\tilde{O}\left((n/d_{\min})^{3/5}\right)$-coloring.*

Proof. First, note that we can assume

$$\left(\frac{n}{d_{\min}}\right)^{3/5} < \frac{d_{\min}}{\log^6 n} \tag{3.18}$$

Otherwise, we can make progress of type 2 towards an $\tilde{O}(d_{\min})$-coloring (i.e. a $\tilde{O}((n/d_{\min})^{3/5})$-coloring) just by taking any vertex of degree $d_{\min}$ (by itself an independent set), and its neighborhood.

Let $\delta = \frac{1}{5 \log n}$. Theorems 3.7.6 and 3.7.7 imply that for some vertex $v \in V$ and some indices $i, j \in \{0, \ldots, \log_{1+\delta} n\}$, the sets $S = \Gamma(v) \cap I_i^\delta$ and $T = J_j^\delta(S)$ satisfy

$$|S| = \Omega\left(\frac{d_{\min}}{\log^4 n}\right), \tag{3.19}$$

$$|T \cap R| \,/\, |T| \geq \frac{1}{2}(1-2\delta)(1-3\delta) \geq \frac{1}{2}\left(1 - \frac{1}{\log n}\right) \tag{3.20}$$

and

$$D_T(S) \geq \frac{\delta}{\log_{1+\delta} n} D_R(S) \geq \frac{(1-3\delta)\delta}{2\log_{1+\delta} n} D(S) = \Omega\left(\frac{d_{\min}|S|}{\log^3 n}\right). \tag{3.21}$$

(We make progress by trying all choices of v, i and j.)

Since by (3.20), $|T|$ has an independent set of size $\frac{1}{2}(1 - \frac{1}{\log n})\,|T|$, we can use the vertex-cover approximation algorithm of Karakostas [22] to find an independent set of size $\Omega(|T|\,/\sqrt{\log |T|})$. Hence, we have made progress if $T = \tilde{\Omega}(n^{2/5} d_{\min}^{3/5})$. Now, suppose this is not the case, that is,

$$|T| \leq \frac{n^{2/5} d_{\min}^{3/5}}{\log^5 n}. \tag{3.22}$$

We want to show that progress towards an $O(k)$ coloring can be made using Theorem 3.7.8 with $k = \log^3 n (n/d_{\min})^{3/5}$. Condition 1 of the theorem holds since $S \subseteq \Gamma(v)$. Substituting for k, we rewrite condition 2 as

$$D_T(S) = \Omega\left(|S| \cdot \frac{d_{\min}}{\log^4 n}\left(\frac{d_{\min}}{n}\right)^{1/5}\right).$$

Since $d_{\min} < n$, it suffices to check that $D_T(S) = \Omega\left(|S| \cdot d_{\min}/\log^4 n\right)$, which follows from (3.21).

To verify condition 3, let us first note that, for the value of k specified above,

$$
\begin{aligned}
|T| \log n &\le \frac{n^{2/5} d_{\min}^{3/5}}{\log^4 n} && \text{by (3.22)} \\
&= O\left(|S|\, n \left(\frac{n}{d_{\min}} \right)^{2/5} \right) && \text{by (3.19)} \\
&= O\left(|S| \cdot \frac{\log^6 n}{k^2} \left(\frac{n}{d_{\min}} \right)^{8/5} \right) \\
&= O(|S|\, n/k^2). && \text{by (3.18)}
\end{aligned}
$$

Hence, to verify condition (3) it suffices to check that

$$
[D_T(S)]^3 = \Omega\left(\left[|S| + \max_{v \in S} d_T(v) \right] |S|^2\, |T|\, n^2/k^4 \right).
$$

Now, by our choice of S, all vertices in S have nearly the same degree, thus

$$
\begin{aligned}
&\left[|S| + \max_{v \in S} d_T(v) \right] |S|^2\, |T|\, n^2/k^4 \\
&\qquad = \left[|S|^2 + \max_{v \in S} d_T(v)\, |S| \right] |S|\, |T|\, n^2/k^4 \\
&\qquad \le \left[|S|^2 + \max_{v \in S} d(v)\, |S| \right] |S|\, |T|\, n^2/k^4 \\
&\qquad \le \left[|S|^2 + D(S)(1+\delta) \right] |S|\, |T|\, n^2/k^4 \\
&\qquad \le \left[|S|^2 + O(D_T(S) \log^3 n) \right] |S|\, |T|\, n^2/k^4 && \text{by (3.21)} \\
&\qquad = \left[|S|^2 + O(D_T(S) \log^3 n) \right] |S|\, |T| \cdot \frac{d_{\min}^{12/5}}{n^{2/5} \log^{12} n} \\
&\qquad \le \left[|S|^2 + O(D_T(S) \log^3 n) \right] |S|\, \frac{d_{\min}^3}{\log^{17} n} && \text{by (3.22)} \\
&\qquad = O\left(D_T(S) \left[|S|^2 + D_T(S) \log^3 n \right] \frac{d_{\min}^2}{\log^{14} n} \right). && \text{by (3.21)}
\end{aligned}
$$

If $D_T(S)\log^3 n \le |S|^2$, then from the above calculation, we get

$$\left[|S| + \max_{v\in S} d_T(v)\right] |S|^2 |T| n^2/k^4 = O\left(D_T(S)\,|S|^2 \frac{d_{\min}^2}{\log^{14} n}\right)$$
$$= O\left([D_T(S)]^3/\log^8 n\right). \qquad \text{by (3.21)}$$

Otherwise, we have

$$D_T(S)\log^3 n \ge |S|^2, \tag{3.23}$$

which gives

$$\left[|S| + \max_{v\in S} d_T(v)\right] |S|^2 |T| n^2/k^4 = O\left([D_T(S)]^2 \frac{d_{\min}^2}{\log^{11} n}\right)$$
$$= O\left([D_T(S)]^2 \frac{|S|^2}{\log^3 n}\right) \qquad \text{by (3.19)}$$
$$= O\left([D_T(S)]^3\right). \qquad \text{by (3.23)}$$

□

3.7.2 Combining Combinatorial Coloring Tools and SDP Rounding

We use the following graph partitioning lemma:

Lemma 3.7.10. *Given an undirected n-vertex graph $G = (V, E)$ and a parameter $\Delta \ge 0$, one can find, in polynomial time, a vertex-induced subgraph G' of size $\ge \frac{n}{4}$, such that either*

- *the maximum degree in G' is at most 2Δ, or*
- *the minimum degree in G' is at least $\frac{\Delta}{2}$*

Proof. Let $V_1 = V$, and $V_2 = \emptyset$. If there is any vertex with degree $\leq \Delta/2$ in the subgraph of G induced on V_1, move this vertex from V_1 to V_2. Repeat this operation until no such vertices are left. If $|V_1| \geq n/2$, let G' be the subgraph of G induced on V_1. This graph has minimum degree $\geq \Delta/2$. Otherwise, we have $|V_2| \geq n/2$. By construction of V_2, there are at most $|V_2|\,\Delta/2$ edges (of the original graph) incident to vertices in V_2. Therefore, the vertices in V_2 have total degree

$$\sum_{v \in V_2} d(v) \leq |V_2|\,\Delta.$$

Of these, at most $|V_2|/2$ vertices can have degree $\geq 2\Delta$. Hence there are at least $|V_2|/2 \geq n/4$ vertices in V_2 with maximum degree $\leq 2\Delta$. Letting G' be the subgraph of G induced on these vertices, we are done. □

We can now give a generalization of [6] and [7]:

Proof of Theorem 3.7.2. By Lemma 3.7.5, it suffices to show that we can make progress towards the desired coloring. Let Δ_0 be the value of Δ which minimizes $f(n/4, 2\Delta) + (n/\Delta)^{3/5}$ (if it is not computable, we can try all values of Δ). Apply the algorithm in Lemma 3.7.10 with $\Delta = \Delta_0$, to find the corresponding subgraph G'. If the maximum degree in G' is at most $2\Delta_0$, apply algorithm $\mathcal{A}$ to make progress towards an $O(f(n/4, 2\Delta_0))$-coloring. Otherwise, the minimum degree in G' must be at least $\Delta_0/2$, in which case, by Theorem 3.7.9, we can make progress towards an $\tilde{O}((n/\Delta_0)^{3/5})$-coloring. □

Chapter 4

Coloring 3-Colorable Graphs Using the Lasserre Hierarchy

In this chapter we present a further improvement for coloring 3-colorable graphs which involves tighter SDP relaxations and a new rounding algorithm. This algorithm gives an $O\left(n^{0.2072}\right)$ coloring. While it is not necessary to understand the algorithm and analysis in the previous chapter in order to understand the present chapter, we will rely heavily on notation and terminology defined in Section 3.1. Moreover, algorithm KMS$'$, which is used as part of the rounding algorithm presented in this chapter, appears in Section 3.2, and the analysis of the algorithm described here will make crucial use of lemmas proved in Sections 3.6 and 3.7.

4.1 Hierarchical SDP Relaxations for 3-Coloring

We first give an SDP relaxation for 3-coloring based on the Lasserre hierarchy described in Section 2.1.2. Since the relaxations discussed earlier were for Independent Set, we first reduce the 3-coloring problem to an Independent Set problem. This is done as follows. Given a graph $G = (V, E)$, construct graph $G' = (V', E')$ where $V' = V \times \{R, B, Y\}$ and

$$E' = \{((i, C), (j, C)) \mid (i, j) \in E \text{ and } C \in \{R, B, Y\}\} \\ \cup \{((i, C_1), (i, C_2)) \mid i \in V \text{ and } C_1 \neq C_2\}.$$

Note that G' has an independent set of size $|V|$ in G' if and only if G is 3-colorable of G (since in an independent set in G', every vertex $i \in V$ can appear in at most 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 we have $v_\emptyset = v_{(i,R)} + v_{(i,B)} + v_{(i,Y)}$ for any $i \in V$. Moreover, the constraints of IS$_k(G')$ are symmetric w.r.t. $\{R, B, Y\}$. Thus, defining $\pi(I) = \{(i, \pi(C)) \mid (i, C) \in I\}$ for every permutation $\pi \in \mathrm{Sym}(\{R, B, Y\})$ and $I \subseteq V'$ (Sym(X) is the group of permutations on X), 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 matrix $\pi(M)$ has entries $\pi(M)_{I,J} = M_{\pi(I),\pi(J)}$ for all subsets $I, J \subseteq V$ of cardinality at most k. This suggests the following SDP relaxation for 3-coloring:

3COL$_k(G)$

$$\{v_I \mid I \subseteq V'\} \text{ satisfy IS}_k(G') \tag{4.1}$$

$$\forall i \in V \qquad v_\emptyset = v_{(i,R)} + v_{(i,B)} + v_{(i,Y)} \tag{4.2}$$

$$\forall \pi,\ \forall I, J \subseteq V' \text{ s.t. } |I|, |J| \leq k \qquad v_I \cdot v_J = v_{\pi(I)} \cdot v_{\pi(J)} \tag{4.3}$$

We now show that the relaxation 3COL$_1(G)$ is equivalent to strict vector 3-coloring, one of the standard SDP relaxation for 3-coloring, which was defined in Section 3.1. This will be useful in the following sections, as we will present an SDP rounding algorithm which uses the algorithm KMS$'$ defined in Chapter 3. For every vertex $i \in V$, by constraints (2.11), (4.2) and (4.3), 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 as

$$v_{(i,R)} = \tfrac{1}{3}v_\emptyset + \tfrac{\sqrt{2}}{3}u_i, \tag{4.4}$$

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

$$\forall (i,j) \in E \quad u_i \cdot u_j = -\tfrac{1}{2}. \tag{4.5}$$

Indeed, this follows immediately from (4.4), since for edges $(i,j) \in E$ we have

$$0 = v_{(i,R)} \cdot v_{(j,R)} = \frac{1}{9}v_\emptyset^2 + \frac{2}{9}u_i \cdot u_j.$$

It is not hard to see that one can similarly construct a solution to $3\text{COL}_1(G)$ given any vector 3-coloring $\{u_i\}$.

4.2 A New Rounding Algorithm

To describe our algorithm we need the following notation. Given a solution $\{v_I\}$ of $3\text{COL}_3(G)$, and vertices $i,k \in V$ s.t. $v_{(i,R),(k,R)} \neq 0$, let w_{ik} be the unit vector which satisfies

$$\begin{aligned} v_{(i,R),(k,R)} &= \left(\frac{\left\|v_{(i,R),(k,R)}\right\|}{\left\|v_{(i,R)}\right\|}\right)^2 v_{(i,R)} + \left\|v_{(i,R),(k,R)}\right\| \sqrt{1 - \left(\frac{\left\|v_{(i,R),(k,R)}\right\|}{\left\|v_{(i,R)}\right\|}\right)^2} w_{ik} \\ &= 3\left\|v_{(i,R),(k,R)}\right\|^2 v_{(i,R)} + \left\|v_{(i,R),(k,R)}\right\| \sqrt{1 - 3\left\|v_{(i,R),(k,R)}\right\|^2} w_{ik}. \end{aligned} \tag{4.6}$$

By (2.11), $\left\|v_{(i,R),(k,R)}\right\|^2 = v_{(i,R),(k,R)} \cdot v_{(i,R)}$, hence such a vector exists, and $w_{ik} \perp v_{(i,R)}$.

Our algorithm is as follows:

KMS$^*(G)$

1. Solve the SDP $3\text{COL}_3(G)$ to get vector solution $\{v_I\}$.
2. Run $\text{KMS}'(G, \{u_i\})$, and let $I \subseteq V$ be the independent set returned.
3. For all i, let $V_i = \left\{k \;\middle|\; \left\|v_{(i,R),(k,R)}\right\|^2 > \frac{1}{6}\left(1 - \frac{1}{\sqrt{\ln n}}\right)\right\}$.
4. Run $\text{KMS}(V_i, \{w_{ik} \mid k \in V_i\}, (\ln n)^{1/4})$, to get independent set W_i.
5. Output the largest set among I and the various W_i (for all $i \in V$).

Figure 4.1: Algorithm **KMS***

Recall that by Lemma 2.2.1, if the threshold parameter $t > 0$ is exactly c-inefficient, then $N(t) = \tilde{\Theta}(\Delta^{-\frac{1}{3+3c}})$. Thus, our objective is to find the largest possible $c = c(\Delta)$ for which KMS* is guaranteed to return an independent set of size $\Omega(N(t)n)$ for a c-inefficient threshold t. Using this terminology, we give the following explicit guarantee for the performance of KMS*.

Theorem 4.2.1. *For every $\tau > \frac{6}{11}$ there exists $c_2(\tau) > 0$ such that for $0 < c < c_2(\tau)$, and any n vertex graph G with maximum degree $\leq n^\tau$, if the threshold t is (at most) c-inefficient for G, then KMS$^*(G)$ returns an independent set of size $\Omega(N(t)n)$.*

Furthermore, $c_2(\tau)$ satisfies

$$c_2(\tau) \stackrel{\text{def}}{=} \min\left\{\frac{1}{2}, \tilde{c}_2(\tau)\right\},$$

where

$$\tilde{c}_2(\tau) \stackrel{\text{def}}{=} \sup \left\{ c \,\middle|\, \min_{0 \le \alpha \le \frac{c}{1+c}} \left(\frac{1}{3} - \frac{1+c}{\tau} + \frac{1}{1-\alpha^2} + \frac{1}{2} \left(\sqrt{1-\alpha} - \sqrt{b(1+\alpha)} \right)^2 \right) > 0 \right\}.$$

Corollary 4.2.2. *For any n-vertex 3-colorable graph G with maximum degree $\le \Delta = n^{0.6546}$, KMS*(G) returns an independent set of size $\Omega(\Delta^{-0.3166} n)$.*

Combining this result with the Blum coloring tools (see Theorem 3.7.2), immediately yields the following result:

Theorem 4.2.3. *For 3-colorable graphs, one can find an $O(n^{0.2072})$ coloring in polynomial time.*

4.3 Overview of Analysis of KMS*

In this section we give an informal description of the analysis of KMS*. Recall that in KMS$'$ we have, by Lemma 3.1.4, for all nodes $i \in V$,

$$\Pr[i \text{ is eliminated from } V_\zeta(t) \mid i \in V_\zeta(t)] \le \Pr[\exists j : \zeta \cdot u'_{ij} \ge \sqrt{3} t]. \tag{4.7}$$

Our goal is to show that, in step 2 of KMS*, either the probability on the right is small for many vectors (and thus, in expectation an $\Omega(N(t))$-fraction of them will be in I), or we can extract a large independent set from the 2-neighborhood $\Gamma(\Gamma(i))$ of some vertex i in step 4. 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,

$$u_i \cdot u_j = \Pr_{\text{col}}[\text{col}(i) = \text{col}(j)] - \frac{1}{2}\Pr_{\text{col}}[\text{col}(i) \neq \text{col}(j)]. \tag{4.8}$$

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 t is chosen such that $N(t) \approx \Delta^{-1/3}$, and KMS$'$ in step 2 fails for this value of t (in fact, we make the stronger assumption that the right-hand-side of inequality (4.7) is large for *all* vertices $i \in V$).

We would first like to show that joint neighborhoods (intersections of two neighborhoods) are clustered. Consider some vertex $i \in V$, a neighbor $j \in \Gamma(i)$ of i, and some neighbor $k \in \Gamma(j)$ of j. By our assumption about statistically significant behavior, if we condition on the choice of color $\text{col}(j)$, any neighbor of j is assigned with equal probability one of the two remaining colors, one of which must be $\text{col}(i)$. Thus, $\Pr[\text{col}(i) = \text{col}(k)] \approx \frac{1}{2}$. 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 $\text{col}(j) = \text{col}(j')$ whenever $\text{col}(i) \neq \text{col}(k)$ (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 $\text{col}(i) = \text{col}(k) = C$ for some color C, we have $\Pr[\text{col}(j) = \text{col}(j') \mid \text{col}(i) = \text{col}(k) = C] \geq \frac{1}{2} - o(1)$ for many pairs $j, j' \in \Gamma(i) \cap \Gamma(k)$ (since in a large set of events, most pairs of events can be only weakly anti-correlated – see the discussion preceeding Lemma 2.1.1). Summarizing, for such pairs we have

$$\begin{aligned}
\Pr[\text{col}(j) = \text{col}(j')] &= \Pr[\text{col}(i) \neq \text{col}(k)] \\
&\quad + \Pr[\text{col}(i) = \text{col}(k)] \cdot \Pr[\text{col}(j) = \text{col}(j') \mid \text{col}(i) = \text{col}(k)] \\
&\geq \tfrac{1}{2} + \tfrac{1}{2} \cdot \tfrac{1}{2} - o(1) = \tfrac{3}{4} - o(1)
\end{aligned}$$

Now, by (4.8), this implies that $u_j \cdot u_{j'} \geq \frac{5}{8} - o(1)$. This, in turn, implies $u'_{ij} \cdot u'_{ij'} \geq \frac{1}{2} - o(1)$, so the vectors $\{u'_{ij} \mid j \in \Gamma(i) \cap \Gamma(k)\}$ form a $\frac{1}{2}$-cluster (this intuition is formalized in Lemma 2.1.1). The cardinality of such clusters must be small, since otherwise, by the bound in Lemma 2.2.3, they would have a disproportionately small contribution to the probability in (4.7). This is made precise in Lemma 4.4.5, which in this case implies that for $i, k \in V$ as above, $|\Gamma(i) \cap \Gamma(k)| \leq \sqrt{\Delta}$.

This now implies that the number of vertices at distance 2 from i is large. Indeed,

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

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 $\tilde{\Omega}(|\Gamma(\Gamma(i))|) = \tilde{\Omega}(\Delta^{3/2})$ (in fact, as we shall see, the subgraph induced on a significant portion of this set is almost vector 2-colorable, allowing us to simply use KMS in step 4). This gives the following trade-off: For t s.t. $N(t) \approx \Delta^{-1/3}$, either step 2 produces an independent set of size $N(t)n \approx \Delta^{-1/3}n$, or step 4 produces an independent set of size $\tilde{\Omega}(\Delta^{3/2})$.

Slightly relaxing the above argument (by decreasing the threshold t in KMS$'$, hence increasing the size of the independent set produced in step (2)), 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 t 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.

4.4 Analysis of KMS*

In this section we prove Theorem 4.2.1. The goal of the analysis is to show that if KMS* 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, since V_i is nearly vector 2-colorable.

Lemma 4.4.1. *For any vertex $i \in V$, let the vertex set V_i be as in algorithm KMS*. Then KMS$(V_i, \{w_{ik} \mid k \in V_i\}, 4(\ln n)^{1/4})$ returns an independent set of size $\Omega(|V_i|\, N(4(\ln n)^{1/4})) = \tilde{\Omega}\left(|V_i|\, n^{-8/\sqrt{\ln n}}\right)$.*

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 (4.6),

$$w_{ik} \cdot w_{ik'} = -\frac{3\left\|v_{(i,R),(k,R)}\right\|\left\|v_{(i,R),(k',R)}\right\|}{\sqrt{\left(1 - 3\left\|v_{(i,R),(k,R)}\right\|^2\right)\left(1 - 3\left\|v_{(i,R),(k',R)}\right\|^2\right)}} < -1 + \frac{2}{\sqrt{\ln n}}.$$

In particular, $\|w_{ik} + w_{ik'}\|^2 < 4/\sqrt{\ln n}$. Thus, for $t = 4(\ln n)^{1/4}$, the probability that both $k, k' \in V_\zeta(t)$ is at most

$$\begin{aligned}
\Pr_\zeta[\zeta \cdot (w_{ik} + w_{ik'}) \geq 2t] &= N(2t/\left\|w_{ik} + w_{ik'}\right\|) \\
&< N\left(4\sqrt{\ln n}\right) \\
&= O\left(n^{-8}\right). && \text{by Lemma 2.2.1}
\end{aligned}$$

Note that, again by Lemma 2.2.1, $N(t) = \tilde{\Omega}\left(n^{-8/\sqrt{\ln n}}\right)$. In particular, the expected number of edges contained in $V_\zeta(t)$ is at most $o(N(t))$, whereas the expected number of vertices is $\Omega(|V_i|\, N(t))$. □

The following theorem, together with Lemma 4.4.1, directly implies Theorem 4.2.1.

Theorem 4.4.2. *For every $\tau > \frac{6}{11}$ and $0 < c < c_2(\tau)$, 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 threshold t which is at most c-inefficient for G, either*

1. *The independent set I found in step (2) of KMS*(G) has expected size $\Omega(N(t)n)$, or*

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

The rest of this section is devoted to proving Theorem 4.4.2. Recall the definition of (s, δ)-covers defined in Section 3.1. We further refine this definition as follows:

Definition 4.4.3. A set of unit vectors X is said to be a *uniformly c-inefficient* (s, δ)-cover, if $|X| \geq \delta N(s)^{-(1+c)}$, and every subset $S \subseteq X$ is a $(s, N(s)^{1+c}|S|)$-cover.

Note that a uniformly c-inefficient cover is indeed a c-inefficient cover. In the above definition, we have

$$1 \geq \Pr[\exists x \in X : \zeta \cdot x \geq s] \geq N(s)^{1+c}|X| \tag{4.9}$$

$$\geq N(s)^{1+c}\delta N(s)^{-(1+c)} = \delta, \tag{4.10}$$

which shows that X is an (s, δ)-cover. The c-inefficiency condition follows from (4.9). Using this definition, we will show that every cover which has bounded inefficiency, contains a large core which has bounded *uniform* inefficiency.

Lemma 4.4.4. *Let X be a c-inefficient (s, δ)-cover. Then*

1. *For some $0 \leq b \leq c + O(\ln(1/\delta)/s^2)$, there exists a subset $X' \subseteq X$ which is a uniformly b-inefficient $(s, \Omega(\delta/s^2))$-cover.*

2. *If, in addition, X is a ρ-cluster and $\delta = \Omega\left(\frac{1}{\text{poly}(s)}\right)$, then $b \geq \frac{\rho}{1-\rho} - \tilde{O}(\frac{1}{s})$.*

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 $(s, \mu(S))$-cover, i.e. $\Pr_\zeta[\exists x \in S : \zeta \cdot x \geq s] \geq \mu(S)$. A natural choice is given by $\mu(x) \stackrel{\text{def}}{=} \Pr_\zeta[\zeta \cdot x \geq s \text{ and } \zeta \cdot x = \max_{x' \in X} \zeta \cdot x']$. Let $X_+ = \{x \mid \mu(x) > \delta N(s)^{1+c}/2\}$, and $X_- = X \setminus X_+$. Then, by the efficiency and cover properties of X, we have

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

Thus, $\mu(X_+) \geq \delta/2$, and, by Lemma 2.2.1 and definition of X_+, for every $x \in X_+$, $\mu(x) = N(s)^{1+b_x}$ for some $b_x \in [0, c + O(\ln(1/\delta)/s^2)]$. Divide this range into s^2 subintervals I_i of length $(c + O(\ln(1/\delta)/s^2))/s^2$, and divide X_+ into corresponding bins $X_i = \{x \in X_+ \mid b_x \in I_i\}$. Thus, some such bin must have measure $\mu(X_i) \geq \delta/(2s^2)$. Thus we can show part (1), as this X_i is uniformly b-inefficient for $b = \max I_i$. Indeed, by definition of I_i, for all $X' \subseteq X_i$ we have $\mu(X') \geq |X'| \, N(s)^{1+b}$. The required lower bound on $|X_i|$ follows from the following observation:

$$\begin{aligned} \frac{\delta}{2s^2} \leq \mu(X_i) &\leq |X_i| \, N(s)^{1+\min I_i} \\ &\leq |X_i| \, N(s)^{1+b-c/s^2} \\ &= O\left(|X_i| \, N(s)^{1+b}\right). && \text{by Lemma 2.2.1} \end{aligned}$$

As noted earlier, this implies in particular that X_i is a b-inefficient $(s, \Omega(\delta/s^2))$-cover. For part (2), let r be such that $N(rs) = o(\delta/s^2)$. By Lemma 2.2.1 there is some $r = O(\sqrt{\log(s^2/\delta)}/s) = O(\sqrt{\log s}/s)$ satisfying this property. Therefore, by

Corollary 2.2.4, we have

$$\frac{\delta}{2s^2} \le \mu(X_i) \le \text{poly}(s)\,|X_i|\,N(s)^{1+\rho/(1-\rho)-O(r)} + o(\delta/s^2)$$
$$\le \text{poly}(s)N(s)^{-(1+b)+1+\rho/(1-\rho)-O(r)} + o(\delta/s^2). \quad \text{by } b\text{-inefficiency of } X_i$$

And so the desired lower bound on b follows, since by the above inequality, we have

$$N(s)^{-b+\rho/(1-\rho)-O(\sqrt{\log s}/s)} \ge \frac{\delta}{2s^2}(1-o(1)) \ge 1/\text{poly}(s) = N(s)^{O(\log s/s^2)}.$$

□

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

Lemma 4.4.5. *Let X be a uniformly b-inefficient (s,δ)-cover, then for all $\rho \ge b/(1+b) + 3\ln s/s^2$ any ρ-cluster in X has cardinality at most*

$$O(\text{poly}(s)N(s)^{-(\sqrt{1-\rho}+\sqrt{b\rho})^2}).$$

Proof. Let the subset $K \subset X$ be a ρ-cluster of cardinality $N(s)^{-\beta}$, and let $r = \sqrt{\rho} - \sqrt{b(1-\rho)+\eta}$ for $\eta = (3+3b)\ln s/s^2$. Then by Lemma 2.2.3 (with the above choice of r), we have

$$\begin{aligned}
\Pr_\zeta[\exists x \in K : \zeta \cdot x \geq s] &\leq O\left(e^{-(1+b+\eta/(1-\rho))s^2/2}|X|\right) + 2N(rs) \\
&\leq O\left(s^{-(1+b)/2}N(s)^{1+b}N(s)^{-\beta}\right) + 2N(rs) \quad \text{by Lemma 2.2.1} \\
&= o\left(N(s)^{-\beta+1+b}\right) + 2N(rs) \\
&\leq o\left(N(s)^{-\beta+1+b}\right) \\
&\quad + \text{poly}(s)N(s)^{\left(\sqrt{\rho}-\sqrt{b(1-\rho)}\right)^2}. \quad \text{by Lemma 2.2.1}
\end{aligned}$$

On the other hand, since X is uniformly b-inefficient, we have $\Pr_\zeta[\exists x \in K : \zeta \cdot x \geq s] \geq N(s)^{-\beta+1+b}$. Hence, we have

$$\begin{aligned}
|K| = N(s)^{-\beta} &\leq \text{poly}(s)N(s)^{\left(\sqrt{\rho}-\sqrt{b(1-\rho)}\right)^2-(1+b)} \\
&= O(\text{poly}(s)N(s)^{-(\sqrt{1-\rho}+\sqrt{b\rho})^2}).
\end{aligned}$$

□

We are now ready to prove Theorem 4.4.2.

Proof of Theorem 4.4.2. In algorithm KMS′ (in step 2 of KMS*) consider the independent set $V'_\zeta(t)$ produced for a c-inefficient threshold t. Since KMS′ tries all possible values of t, we have $\mathbb{E}[|I|] \geq \mathbb{E}\left[\left|V'_\zeta(t)\right|\right]$. If for at least $n/2$ vertices $i \in V$ we have $\Pr_\zeta[i \in V'_\zeta(t)] \geq \frac{1}{2}N(t)$, then by linearity of expectation, we have $\mathbb{E}[|I|] = \Omega(N(t)n)$. Suppose not. Now prune as in Lemma 3.6.11, and for simplicity, assume $G = (V, E)$ is the remaining graph. Now, fixing some $i \in V$, we have that $\{u'_{ij} \mid j \in \Gamma(i)\}$ and the sets $\{u'_{jk} \mid k \in \Gamma(j)\}$ for every $j \in \Gamma(i)$ are all uniform $\left(\sqrt{3}t, \frac{1}{8} - O\left(\frac{1}{\log t}\right)\right)$-covers which are at most c-inefficient. Moreover, there exists some constant $C > 0$ such that letting $W_j = \left\{k \in \Gamma(j) \;\middle|\; -\frac{C}{t^2} \leq u'_{ji} \cdot u'_{jk} \leq \frac{c}{1+c} + \frac{C}{\log t}\right\}$ for every $j \in \Gamma(i)$, the sets $\{u'_{jk} \mid k \in W_j\}$ are $\left(\sqrt{3}t, \Omega\left(1/t^3\right)\right)$-covers. Note that for all $j \in \Gamma(i)$, $W_j \subseteq V_i$. Therefore, it suffices to give a lower bound on $\left|\bigcup_{j\in\Gamma(i)} W_j\right|$.

Now, subdivide the interval $[-\frac{C}{t^2}, \frac{c}{1+c}+\frac{C}{\log t}]$ into $O(t)$ subintervals I_l of length $1/t$. For every $j \in \Gamma(i)$ there is some $l = l(j)$ such that the set $\{u'_{jk} \mid u'_{ji} \cdot u'_{jk} \in I_{l(j)}\}$ is a $(\sqrt{3}t, \Omega(1/t^4))$-cover. Moreover, there is some l_0 such that, defining $U_i = \{j \in \Gamma(i) \mid l(j) = l_0\}$, the set $\{u'_{ij} \mid j \in U_i\}$ is a $(\sqrt{3}t, \Omega(1/t))$-cover. Let α be such that $I_{l_0} = [\alpha, \alpha+1/t)$. By Lemma 4.4.4, there is some subset $U'_i \subseteq U_i$ s.t. the set $\{u'_{ij} \mid j \in U'_i\}$ is a uniformly b-inefficient $(\sqrt{3}t, \Omega(1/t^3))$-cover for some $0 \leq b \leq c+o(1)$. Similarly, for every $j \in U'_i$, there is some set $W'_j \subseteq \{k \in \Gamma(j) \mid u'_{ji} \cdot u'_{jk} \in I_{l(0)}\}$ for which the set $\{u'_{jk} \mid k \in W_j\}$ is a uniformly a_j-inefficient $(\sqrt{3}t, \Omega(1/t^6))$-cover, where, by part 2 of Lemma 4.4.4, $a_j \geq \alpha^2/(1-\alpha^2) - o(1)$ (since the sets $\{u'_{jk} \mid u'_{ji} \cdot u'_{jk} \in I_{l_0}\}$ are α^2-clusters for all $j \in \Gamma(i)$). Let us summarize the situation:

1. U'_i is a uniformly b-inefficient $(\sqrt{3}t, \Omega(1/t^3))$-cover for some $0 \leq b \leq c+o(1)$.

2. $\forall j \in U'_i$, the set $\{u'_{jk} \mid W'_j\}$ is a uniformly a_j-inefficient $(\sqrt{3}t, \Omega(1/t^5))$-cover for some $a_j \geq \alpha^2/(1-\alpha^2) - o(1)$.

3. $\forall j \in U'_i, k \in W'_j$:

$$\begin{aligned} v_{(i,R)} \cdot v_{(k,R)} &= \frac{1}{9} + \frac{2}{9} u_i \cdot u_k \\ &= \frac{1}{9} + \frac{2}{9}\left(\frac{1}{4} + \frac{3}{4} u'_{ji} \cdot u'_{jk}\right) \in [(1+\alpha)/6, (1+\alpha)/6 + o(1)]. \end{aligned}$$

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

For the sake of simplicity, we will strengthen property 3, and assume that in fact for all $j \in U'_i, k \in W'_j$, $\left\|v_{(i,R),(k,R)}\right\|^2 = v_{(i,R)} \cdot v_{(k,R)} = \frac{1+\alpha}{6}$. This is w.l.o.g. as the $o(1)$ error terms will have a negligble effect. By constraint (4.3), 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) $\frac{1}{3} = \left\|v_{(i,B)}\right\|^2 = \sum_{C\in R,B,Y} \left\|v_{(i,B),(k,C)}\right\|^2$, we have (again by (4.3)),

$$\left\|v_{(i,B),(k,Y)}\right\|^2 = \left\|v_{(i,Y),(k,B)}\right\|^2 = \frac{1}{2}\left(\frac{1}{3} - \frac{1+\alpha}{6}\right) = \frac{1-\alpha}{12}.$$

Similarly, for $j \in \Gamma(i) \cap \Gamma(k)$ and $(C_1, C_2) \in \{(B,Y),(Y,B)\}$, we have by constraints (4.1) and (4.3),

$$v_{(i,C_1),(k,C_1)} \cdot v_{(j,R)} = \frac{1}{2}\left\|v_{(i,C_1),(k,C_1)}\right\|^2 = \frac{1+\alpha}{12},$$

and

$$v_{(i,C_1),(k,C_2)} \cdot v_{(j,R)} = \left\|v_{(i,C_1),(k,C_2)}\right\|^2 = \frac{1-\alpha}{12}.$$

Finally, we note that for all $(i,j) \in E$,

$$\begin{aligned}
v_{(j,R)} &= \frac{1}{3}v_\emptyset + \frac{\sqrt{2}}{3}u_j \\
&= \frac{1}{3}v_\emptyset - \frac{1}{3\sqrt{2}}u_i + \frac{1}{\sqrt{6}}u'_{ij} \\
&= \frac{1}{2}v_\emptyset - \frac{1}{2}v_{(i,R)} + \frac{1}{\sqrt{6}}u'_{ij} \\
&= \frac{1}{2}\left(v_{(i,B)} + v_{(i,Y)}\right) + \frac{1}{\sqrt{6}}u'_{ij}. && \text{by constraint (4.2)}
\end{aligned}$$

We now fix some $k \in \bigcup_{j\in U'_i} W'_j$, and apply Lemma 2.1.1, where for all $C_1 \in \{B,Y\}$, we let $p_{C_1} = \left\|v_{(i,C_1)}\right\|^2 = \frac{1}{3}$, $p_{C_1}q_{C_1} = v_{(i,C_1)} \cdot v_{(j,R)} = \frac{1}{6}$, and for all $C_1 \in \{B,Y\}$ and $C_2 \in \{R,B,Y\}$ we let

$$p_{C_1,C_2} = \left\|v_{(i,C_1),(k,C_2)}\right\|^2 = \begin{cases} \frac{1+\alpha}{6} & C_1 = C_2 \\ \frac{1-\alpha}{12} & C_1 \neq C_2 \end{cases}$$

and

$$p_{C_1,C_2}q_{C_1,C_2} = v_{(i,C_1),(k,C_2)} \cdot v_{(j,R)} = \begin{cases} \frac{1+\alpha}{12} & C_1 = C_2 \\ \frac{1-\alpha}{12} & C_2 \in \{B,Y\} \setminus C_1 \\ 0 & C_2 = R. \end{cases}$$

By the lemma, there is some unit vector

$$x_0' \in \operatorname{Span}\left(\{v_I \mid I \subseteq (\{i\} \times \{B,Y\}) \cup (\{k\} \times \{R,B,Y\})\}\right)$$

such that for all $j \in U_i'$ s.t. $W_j' \ni k$, we have

$$\begin{aligned} x_0' \cdot \frac{1}{\sqrt{6}} u_{ij}' &= \sqrt{\sum_{C_1=B,Y} \sum_{C_2=R,B,Y} p_{C_1,C_2} q_{C_1,C_2}^2 - \sum_{C_1=B,Y} p_{C_1} q_{C_1}^2} \\ &= \sqrt{2 \cdot \frac{1+\alpha}{6} \cdot \left(\frac{1}{2}\right)^2 + 2 \cdot \frac{1-\alpha}{12} - 2 \cdot \frac{1}{3} \cdot \left(\frac{1}{2}\right)^2} = \sqrt{\frac{1-\alpha}{12}}. \end{aligned}$$

Thus, for all k, the set $\{u_{ij}' \in U_i' \mid W_j' \ni k\}$ is in fact a $(1-\alpha)/2$-cluster, and so by property 1 above and Lemma 4.4.5, we have $\left|\{j \in U_i' \mid W_j' \ni k\}\right| \leq N(\sqrt{3}t)^{-\left(\sqrt{(1+\alpha)/2}+\sqrt{b(1-\alpha)/2}\right)^2 - o(1)}$. Hence,

$$\begin{aligned} \sum_{j \in U_i'} \left|W_j'\right| &= \left|\{(j,k) \mid j \in U_i' \text{ and } k \in W_j'\}\right| \\ &= \sum_{k \in \bigcup_{j \in U_i'} W_j'} \left|\{j \in U_i' \mid W_j' \ni k\}\right| \\ &\leq \left|\bigcup_{j \in U_i'} W_j'\right| \cdot N(\sqrt{3}t)^{-\frac{1}{2}\left(\sqrt{1+\alpha}+\sqrt{b(1-\alpha)}\right)^2 - o(1)}. \end{aligned}$$

Yet, by properties 1 and 2, $\sum_{j \in U_i'} \left|W_j'\right| \geq N(\sqrt{3}t)^{-(1+b)-(1+\alpha^2/(1-\alpha^2))+o(1)}$.

Thus we have shown

$$\begin{aligned}\Big|\bigcup_{j\in U_i'} W_j'\Big| &\geq N(\sqrt{3}t)^{-\left(2+\alpha^2/(1-\alpha^2)+b-\frac{1}{2}\left(\sqrt{1+\alpha}+\sqrt{b(1-\alpha)}\right)^2\right)+o(1)} \\ &= N(\sqrt{3}t)^{-\left(1/(1-\alpha^2)+\frac{1}{2}\left(\sqrt{1-\alpha}-\sqrt{b(1+\alpha)}\right)^2\right)}.\end{aligned} \tag{4.11}$$

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 $(1-\alpha)/(1+\alpha) \geq c - o(1)$. Therefore, from (4.11) we get, for some constant $\varepsilon' > 0$

$$\begin{aligned}\Big|\bigcup_{j\in\Gamma(i)} W_j\Big| \geq \Big|\bigcup_{j\in U_i'} W_j'\Big| & \\ &\geq N(\sqrt{3}t)^{-\left(1/(1-\alpha^2)+\frac{1}{2}\left(\sqrt{1-\alpha}-\sqrt{c(1+\alpha)}\right)\right)+o(1)} \\ &\geq N(\sqrt{3}t)^{\frac{1}{3}-\frac{1+c}{\tau}-\varepsilon'+o(1)} && \text{since } c < c_2(\tau) \\ &\geq N(\sqrt{3}t)^{\frac{1}{3}-\varepsilon'+o(1)}\Delta^{1/\tau} && \text{by } c\text{-inefficiency of } t \\ &= N(\sqrt{3}t)^{\frac{1}{3}-\varepsilon'+o(1)}n \\ &= \omega(N(t)n). && \text{by Lemma 2.2.1}\end{aligned}$$

□

Chapter 5

Maximum Independent Set in 3-Uniform Hypergraphs

In this chapter, we describe relaxations for Maximum Independent Set in 3-uniform hypergraphs which arise from SDP hierarchies. We present two algorithms which, for every $\gamma > 0$, in any n-vertex 3-uniform hypergraph containing an independent set of size γn, find an independent set of size $n^{\Omega(\gamma^2)}$. Each of these rounding algorithms works for vector solutions which satisfy the $\Theta(1/\gamma^2)$-level of some SDP hierarchy. For the hierarchy used in the first algorithm, we also present an integrality gap which rules out this performance guarantee at any level up to $\frac{1}{\gamma} + 1$.

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

The relaxation proposed in [28] 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 $(i, j, k) \in E$ it follows that

$x_i + x_j + x_k \in \{0,1,2\}$ (and hence $|x_i + x_j + x_k - 1| \leq 1$). Thus, we have the following relaxation (where vector v_i represents x_i, and $v_\emptyset$ represents 1:

MAX-KNS(H)

$$\text{Maximize } \textstyle\sum_i \|v_i\|^2 \text{ s.t.} \qquad v_\emptyset^2 = 1 \tag{5.1}$$

$$\forall i \in V \qquad v_\emptyset \cdot v_i = v_i \cdot v_i \tag{5.2}$$

$$\forall (i,j,l) \in E \qquad \|v_i + v_j + v_k - v_\emptyset\|^2 \leq 1 \tag{5.3}$$

The existence of local distributions on independent sets corresponding to solutions to $\mathrm{IS}_l^{\mathrm{mix}}(H)$ and $\mathrm{IS}_l^{\mathrm{Las}}(H)$ (for $l \geq 3$) implies that all valid constraints involving at most l vertices (other than integrality gaps) hold true for these relaxations. In particular, constraints (5.2) and (5.3) above are implied by such SDP relaxations, and all results pertaining to the relaxation proposed in [28] hold true for the tighter relaxations we shall consider.

5.2 A Simple Integrality Gap

As observed in [28], 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 $\text{MAX-IS}^{\mathrm{mix}}$ to yield improved approximations, by demonstrating a simple extention of the above integrality gap:

Theorem 5.2.1. *For every integer $l \geq 3$ and any 3-uniform hypergraph H, we have $\text{MAX-IS}_l^{\mathrm{mix}} \geq \frac{1}{l-1}n$.*

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 $v_i = \frac{1}{l-1}v_\emptyset + \sqrt{\frac{1}{l-1} - \frac{1}{(l-1)^2}}u_i$, and $y_{\{i\}} = \frac{1}{l-1}$. Let $y_\emptyset = 1$

and for every pair of distinct vertices $i, j \in V$ let $y_{\{i,j\}} = \frac{1}{(l-1)^2}$. For all sets $I \subseteq V$ s.t. $3 \leq |I| \leq l$, let $y_I = 0$.

It is immediate that constraint (2.13) and the Sherali-Adams constraint (2.7) are satisfied. Since $y_I = 0$ for all sets I of size 3, Sherali-Adams constraint (2.9) is also satisfied. To verify Sherali-Adams constraints (2.8), it suffices to show, for any set $S \subseteq [n]$ of size l, a corresponding distribution on random $0-1$ 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 $\frac{l}{2(l-1)}$, set $x_i^* = x_j^* = 1$ and for all other $l \in S$, set $x_l^* = 0$. Otherwise, set all $x_l^* = 0$. □

5.3 The Algorithm of Krivelevich, Nathaniel and Sudakov

We first review the algorithm and analysis given in [28]. Let us introduce the following notation: For all $h \in \{0, 1, \ldots, \lfloor \log n \rfloor\}$, let $T_h \stackrel{\text{def}}{=} \{i \in V \mid h/\log n \leq \|v_i\|^2 < (h+1)/\log n\}$. Also, since $\|v_i\|^2 = v_\emptyset \cdot v_i$, we can write

$$v_i = (v_\emptyset \cdot v_i)v_\emptyset + \sqrt{v_\emptyset \cdot v_i(1 - v_\emptyset \cdot v_i)}u_i, \tag{5.4}$$

where u_i is a unit vector orthogonal to $v_\emptyset$. They show the following two lemmas, slightly rephrased here:

Lemma 5.3.1. *If the optimum of* KNS(H) *is* $\geq \gamma n$*, there exists an index* $h \geq \gamma \log n - 1$ *s.t.* $|T_h| = \Omega(n/\log^2 n)$*.*

Lemma 5.3.2. *For index* $h = \beta \log n$ *and hyperedge* $(i, j, k) \in E$ *s.t.* $i, j, k \in T_h$*,* *constraint* (5.3) *implies*

$$\|u_i + u_j + u_k\|^2 \leq (6 - 9\beta)/(1 - \beta) + O(1/\log n). \tag{5.5}$$

Note that constraint (5.5) 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 [28], which is applied to the subhypergraph induced on T_h, where h is as in Lemma 5.3.1.

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.

Figure 5.1: Algorithm **KNS-Round**

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 5.3.2, $\zeta \cdot \frac{u_i+u_j+u_k}{\|u_i+u_j+u_k\|} \geq (\sqrt{(3-3\gamma)/(2-3\gamma)} - O(1/\log n))t$. By Lemma 2.2.1, 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 – roughly, so that $N(t) = \tilde{\Theta}\left((n/\,|E|)^{2-3\gamma}\right)$ (assume $\gamma \leq 2/3$) – then yields the guarantee given in [28]:

Theorem 5.3.3. *Given a 3-uniform hypergraph H on n vertices and m hyperedges* *containing an independent set of size γn, 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 $\gamma \le \frac{1}{2}$. Using techniques similar to those used in Chapter 4, one can show that when the vectors satisfy $\mathrm{IS}_3^{\mathrm{Las}}(H)$, the same rounding algorithm does give a non-trivial guarantee (n^ε) for $\gamma \ge \frac{1}{2} - \varepsilon$ (for some fixed $\varepsilon > 0$). However, it is unclear whether this approach can work for arbitrarily small $\gamma > 0$.

Let us introduce the following notation for hyperedges e along with the corresponding vectors $\{u_i \mid i \in e\}$:

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

In particular, when the hyperedge e is of size 3, we have

$$\alpha(e) \stackrel{\mathrm{def}}{=} \tfrac{1}{6} \textstyle\sum_{i\in e} \sum_{j \in e\setminus\{i\}} u_i \cdot u_j.$$

Using this notation, we note the following Lemma which was implicitly used in the above analysis, and which follows immediately from Lemma 2.2.1.

Lemma 5.3.4. *In algorithm KNS-Round, the probability that a hyperedge e is fully contained in $V_\zeta(t)$ is at most $\tilde{O}\left(N(t)^{3/(1+2\alpha(e))}\right)$.*

5.4 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 of vertices i, j with large (positive) inner-product $u_i \cdot u_j$. As we shall see in Section 5.5, 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. This property in turn allows us to successfully use KNS-Round on this subhypergraph, thus extracting a large independent set.

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 $S_1 = \{k \in V \mid (i_0, j_0, k) \in E\}$ 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 step recursively due to the existence of bad pairs $(i_0, j_0), \ldots, (i_s, j_s)$ (thus focusing on the subhypergraphs induced on a chain of sets $S_1 \supseteq S_2 \supseteq \ldots \supseteq S_{s+1}$), 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. The algorithm takes as input an n-vertex hypergraph H for which $\text{MAX-IS}^{\text{mix}}_l(H) \geq \gamma n$, where $l = \Omega(1/\gamma^2)$ and $\{v_i\}$ is the corresponding SDP solution, and outputs an independent set in H. We will use the notation $\{u_i\}$ as in (5.4).

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$(H, \{w_i \mid i \in V\}, t)$.

Figure 5.2: Algorithm **H-Round**

Theorem 5.4.1. *For any constant $\gamma > 0$, given an n-vertex 3-uniform hypergraph $H = (V, E)$, and $\{v_i\}$ satisfying $\mathrm{IS}^{\mathrm{mix}}_{4/\gamma^2}(H)$ and $|\,\|v_i\|^2 - \gamma| \leq 1/\log n$ ($\forall i \in V$), 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 5.3.1, we get:

Corollary 5.4.2. *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 γn, finds an independent set of size $\tilde{\Omega}(n^{\gamma^2/32})$ in H.*

Before we prove Theorem 5.4.1, let check us that we make few recursive calls in Step 2, and that $|V|$ is still large when Step 3 is reached.

Proposition 5.4.3. *For $\gamma > 0$, $H = (V, E)$, and $\{v_i\}$as in Thereom 5.4.1:*

1. *Algorithm* H-Round *makes at most $2/\gamma^2$ recursive calls in Step 2.*
2. *In the final recursive call to* H-Round, *$|V| \geq \sqrt{n}$.*

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 $s' \le \min\{s, 2/\gamma^2\}$,

$$\sum_{l=1}^{s'} v_{i_l} \cdot v_{j_l} \le 1. \tag{5.6}$$

Indeed, let $T = \bigcup\{i_l, j_l \mid 1 \le l \le s'\}$. Since vectors $\{v_i\}$ satisfy $\mathrm{IS}^{\mathrm{mix}}_{4/\gamma^2}(H)$, and $|T| \le 2s' \le 4/\gamma^2$, 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

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

Similarly, if $s' \le \min\{s, 2/\gamma^2 - 1\}$, then for any vertex $k \in \bigcap_{l \le s'} \Gamma(i_l, j_l)$ 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 $s' = \lceil 2/\gamma^2 - 1 \rceil$ and $k = i_{s'+1}$ above, we 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 $\ge n^{\prod(1 - v_{i_l} \cdot v_{j_l}/2)}$, and that by (5.6) 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 5.4.1.

Proof of Theorem 5.4.1. For simplicity, let us assume that $\|v_i\|^2 = \gamma$ for all $i \in V$. Violating this assumption can 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 polynomial slack to absorb such errors. Thus, for any $i, j \in V$, we have

$$v_i \cdot v_j = \gamma^2 + (\gamma - \gamma^2) u_i \cdot u_j. \tag{5.7}$$

For brevity, we write $v_i \cdot v_j = \theta_{ij}\gamma$ for all $i, j \in V$ (note that all $\theta_{ij} \in [0, 1]$). We reintroduce the notation $\alpha(e)$ introduced earlier, in the context of the vectors $\{w_i\}$:

$$\alpha(e) = \frac{1}{6}\sum_{i \in e}\sum_{j \in e \setminus \{i\}} w_i \cdot w_j.$$

The upper-bound on the running time follows immediately from part 1 of Proposition 5.4.3. By part 2 of Proposition 5.4.3, 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 $\Omega(N(t)n) = \Omega(n^{\gamma^2/16})$.

Let us examine the performance of KNS-Round in this instance. Recall that for every $i \in V$, $\Pr[i \in V_\zeta(t)] = N(t)$. Thus, by linearity of expectation, the expected cardinality of $V_\zeta(t)$ 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 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 $\theta_{ij} \geq \gamma/2$. 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 $\Gamma_+(i, j) = \{k \in \Gamma(i, j) \mid \theta_{ik}, \theta_{jk} \leq \theta_{ij}\}$. By (5.7), for all $i, j \in V$ and $k \in \Gamma_+(i, j)$ we have

$$\alpha(i, j, k) \leq w_i \cdot w_j = \frac{\gamma}{24}(u_i \cdot u_j) = \frac{\gamma(\theta_{ij} - \gamma)}{24(1 - \gamma)} \leq \frac{\theta_{ij}\gamma}{24}. \tag{5.8}$$

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

$$|\Gamma_+(i,j)| \leq |\Gamma(i,j)| \leq n^{1-\theta_{ij}\gamma/2}. \tag{5.9}$$

By linearity of expectation, we have

$$\begin{aligned}
\mathbb{E}[|\{e \in E_+ \mid e \subseteq V_\zeta(t)\}|] &= \sum_{e \in E_+} \Pr[e \subseteq V_\zeta(t)] \\
&\leq \sum_{e \in E_+} \tilde{O}\left(N(t)^{3/(1+2\alpha(e))}\right) && \text{by Lemma 5.3.4} \\
&\leq \sum_{\substack{i,j \in V \\ \theta_{ij} \geq \gamma/2}} \sum_{k \in \Gamma_+(i,j)} \tilde{O}\left(N(t)^{3/(1+\frac{1}{12}\theta_{ij}\gamma)}\right). && \text{by (5.8)}
\end{aligned}$$

By (5.9), this gives

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

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 (5.7), and by definition of $\{w_i\}$, we have

$$\alpha(e) \leq -\frac{\gamma^2}{48(1-\gamma)} \tag{5.10}$$

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

$$\begin{aligned}
\mathbb{E}[|\{e \in E_- \mid e \subseteq V_\zeta(t)\}|] &\leq \sum_{e \in E_-} \Pr[e \subseteq V_\zeta(t)] \\
&\leq \sum_{e \in E_-} \tilde{O}\left(N(t)^{3/(1+2\alpha(e))}\right) && \text{by Lemma 5.3.4} \\
&= N(t) \sum_{e \in E_-} \tilde{O}\left(N(t)^{(2-2\alpha(e))/(1+2\alpha(e))}\right) \\
&\leq N(t) n^3 \tilde{O}\left(N(t)^{(2-2\gamma+\frac{1}{24}\gamma^2)/(1-\gamma-\frac{1}{24}\gamma^2)}\right). && \text{by (5.10)}
\end{aligned}$$

By our choice of t, this gives

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

This completes the proof. □

5.5 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 $\text{MAX-IS}_l^{\text{Las}}(H) \geq \gamma n$, where $l = \Omega(1/\gamma^2)$ and $\{v_I \mid |I| \leq l\}$ 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 $h = \gamma' \log n - 1$ be as in Lemma 5.3.1 (where $\gamma' \geq \gamma$). If $\gamma' > 2/3 + 2/\log n$, output T_h.
2. Otherwise, set $H = H|_{T_h}$, and $\gamma = \gamma'$.
3. If for some $i, j \in T_h$ 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$(H|_{\Gamma(i,j)}, \{v_I|_{x_i^*=0 \vee x_j^*=0} \mid I \subseteq \Gamma(i,j), |I| \leq l-2\}, \gamma/(1-\rho_{ij}))$.
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$(H, \{w_i \mid i \in V\}, t)$.

Figure 5.3: Algorithm **H-Round**$^{\text{Las}}$

For this algorithm, we have the following guarantee:

Theorem 5.5.1. *For any constant $\gamma > 0$, given an n-vertex 3-uniform hypergraph $H = (V, E)$ for which $\text{MAX-IS}_{8/(3\gamma^2)}^{\text{Las}}(H) \geq \gamma n$ 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 5.4.1. 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 (5.3) cannot be satisfied (assuming (5.2) holds) when $\|v_i\|^2, \|v_j\|^2, \|v_k\|^2 > 2/3$.

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 (as in the proof of Proposition 5.4.3), 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

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

The proof of this fact is identical to the proof of Theorem 5.4.1.

Chapter 6

Conclusion

We have presented two algorithms for coloring 3-colorable graphs, and one for Maximum Independent Set in 3-uniform hypergraphs. At present, the latter two algorithms give the best known approximation guarantees for their respective problems. The algorithms and proof techniques discussed make crucial use of non-local analysis, and in some cases non-local SDP relaxations.

For graph coloring, there is a possibility for obtaining an improved guarantee by proving certain geometric conjectures [3] and applying them in the context of the analysis in Chapter 3. However, this approach seems to require a more nuanced understanding of measure concentration which may be out of reach at this time. Moreover, given the integrality gaps in [16], this direction cannot yield legal colorings using fewer than $n^{0.156}$ colors unless we can make crucial use of tighter SDP relaxations.

Indeed, we have used tighter relaxations in Chapter 4 to obtain an algorithm with a better performance guarantee. Generalizing this approach appears to be a promising direction which merits further investigation.

For Maximum Independent Set in 3-uniform hypergraphs, Theorem 5.4.1, together with the integrality gap of Theorem 5.2.1, 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 5.2.1 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 the more intriguing of the two, a result of either sort would provide crucial insights into the importance of LP and SDP hierarchies for approximation algorithms. Recently Schoenebeck [36] 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.

From the recent work of Raghavendra [33], we know that improved approximations at unboundedly many constant levels of an SDP hierarchy are not possible for binary CSPs which admit a constant-factor approximation, unless the Unique Games Conjecture [25] is false. While disproving the Unique Games Conjecture (for example, by obtaining improved algorithms for Unique Games based on tighter SDP relaxations) is an exciting prospect, there remains much scope for improvement for problems such as the ones we have discussed here, where we do not believe that constant-factor approximations are achievable.

Bibliography

[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 Annual ACM Symposium of Theory of Computing, pp. 215–224, 2006.

[4] S. Arora, J. R. Lee, and A. Naor. *Euclidean distortion and the sparsest cut.* In Proceedings of the 37th Annual ACM Symposium on Theory of Computing, pp. 553–562, 2005.

[5] 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.

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

[7] 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.

[8] C. Borell. The Brunn-Minkowski inequality in Gauss space. *Invent. Math.*, 30:207–216, 1975.

[9] M. Charikar, K. Makarychev, and Y. Makarychev. *Integrality gaps for Sherali-Adams relaxations.* Manuscript.

[10] 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.

[11] 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.

[12] E. Chlamtac and G. Singh. *Improved approximation guarantees through higher levels of SDP hierarchies.* In Proceedings of the 11th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, pp. 49–62, 2008.

[13] 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.

[14] I. Dinur, E. Mossel, and O. Regev. *Conditional hardness for approximate coloring.* In Proceedings of 38th Annual ACM Symposium on Theory of Computing, pp. 344–353, 2006.

[15] U. Feige and J. Kilian *Zero knowledge and the chromatic number. Journal of Computer and System Sciences* 57(2):187–199, 1998.

[16] U. Feige, M. Langberg, and G. Schechtman. Graphs with tiny vector chromatic numbers and huge chromatic numbers. *SIAM J. Computing*, 33(6):1338–1368, 2004.

[17] W. Feller. *An Introduction to Probability Theory and Its Applications,* volume 1. John Wiley & Sons, 3rd edition, 1968.

[18] 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.

[19] 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.

[20] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization.* Springer Verlag, 1988.

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

[22] G. Karakostas A better approximation ratio for the Vertex Cover problem. In Proceedings of 32nd International Colloquium on Automata, Languages and Programming, pp. 1043–1050, 2005.

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

[24] 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.

[25] S. Khot. *On the power of unique 2-prover 1-round games.* In Proceedings of the 34th Annual ACM Symposium on Theory of Computing, pp. 767–775, 2002.

[26] 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.

[27] S. Khot and O. Regev. *Vertex cover might be hard to approximate to within $2 - \varepsilon$.* Journal of Computer and System Sciences, 74(3):335-349, 2008.

[28] 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.

[29] 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.

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

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

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

[33] P. Raghavendra. *Optimal algorithms and inapproximability results For every CSP?* In Proceedings of 40th Annual ACM Symposium on Theory of Computing, pp. 245–254, 2008.

[34] G. Schoenebeck, L. Trevisan, and M. Tulsiani. *Tight integrality gaps for Lovász-Schrijver LP relaxations of Vertex Cover and Max Cut.* In Proceedings of 39th Annual ACM Symposium on Theory of Computing, pp. 302–310, 2007.

[35] 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.

[36] G. Schoenebeck. *Linear level Lasserre lower bounds for certain k-CSPs.* In Proceedings of 49th IEEE Symposium on Foundations of Computer Science, pp. 593–602, 2008.

[37] V. Sudakov and B. Tsirel'son. Extremal properties of half-spaces for spherically invariant measures. Soviet Math, 9:9–18, 1978. Translated from Zap. Nauch. Sem. L.O.M.I 41:14–24, 1964.

[38] 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.

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

[40] *Linear degree extractors and the inapproximability of max clique and chromatic number.* Theory of Computing 3(1):103–128, 2007.

[41] U. Zwick. *Computer assisted proof of optimal approximability results.* In Proceedings of 13th ACM Symposium on Discrete Algorithms, pp. 496–505, 2002.