Quadratic Forms on Graphs and Their Applications

Konstantin Makarychev

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

January 2008

Abstract

We study the following quadratic optimization problem.

MAX QP: *Given a real matrix a_{ij}, maximize the quadratic form*

$$\sum_{ij} a_{ij} \cdot x_i x_j,$$

where variables x_i take values ± 1.

We show that the integrality gap of the natural SDP relaxation depends on the structure of the support of the matrix A. We define a new graph parameter, the Grothendieck constant of a graph $G = (V, E)$, to be the worst integrality gap among matrices A with support restricted to the edges of G (i.e. we require that if $(i, j) \notin E$, then $a_{ij} = 0$).

We give upper and lower estimates for the Grothendieck constant of the graph G: We show that it is less than $O(\log \vartheta(\bar{G}))$, where $\vartheta(\bar{G})$ is the Lovász theta function of the complement of G, which is always smaller than the chromatic number of G. This yields an efficient constant factor approximation algorithm for the above maximization problem for a wide range of graphs G.

We prove that the Grothendieck constant is always at least $C \log w(G)$, where $w(G)$ is the clique number of G. In particular it follows that the maximum possible integrality gap for the complete graph on n vertices is $C \log n$.

We present approximation algorithms for the MAX k-CSP and Advantage over Random for Maximum Acyclic Subgraph problems. These algorithms solve the MAX QP problem at an intermediate step and then use the obtained solution to solve more complex combinatorial problems.

Acknowledgments

First and foremost, I would like to thank my advisor, Moses Charikar. I am very grateful to Moses for introducing me to one of the most beautiful areas of computer science, approximation algorithms. I want to thank Moses for his guidance and his encouragement. This thesis would not have been possible without his help.

I would like to thank the members of my PhD committee, Boaz Barak, Bernard Chazelle, Assaf Naor, and Robert Tarjan for their valuable comments and suggestions. I am very grateful to my coauthors, Noga Alon, Yury Makarychev, Assaf Naor and, of course, Moses Charikar. I have learned a lot working with them; and I hope to continue collaborating with them in the future. I want to thank my colleagues at Microsoft Research and IBM Research whom I worked with as a summer intern in 2005 and 2006, and especially my mentors Assaf Naor and Maxim Sviridenko.

I am very grateful to Princeton University and the Department of Computer Science. I spent four wonderful years here. I want to thank Melissa Lawson and Mitra Kelly for handling all administrative tasks.

I want to thank my teachers at Moscow High School *57*: Lev D. Altshuler, Boris M. Davidovich, Sergey Dorichenko, Arkadiy Skopenkov, Yevgeniy A. Vyrodov, and Ivan V. Yashchenko; and my undergraduate advisors Alexander Shen and Nikolai K. Vereshchagin.

I gratefully acknowledge a Gordon Wu fellowship, IBM Graduate fellowship and Moses Charikar's NSF ITR grant CCR-0205594, NSF CAREER award CCR-0237113, MSPA-MCS award 0528414 for supporting my PhD research.

Finally, I would like to thank my family: my brother, Yury Makarychev, who not only was my main coauthor, but was always my best friend, and my parents, Marina Makarycheva and Sergey Makarychev, who supported me throughout my life. This thesis is dedicated to them.

Konstantin Makarychev

Princeton University

November, 2007

To my parents,

Marina and Sergey

Contents

List of Figures

Chapter 1

Preface

In this dissertation, we study the following quadratic optimization problem:

MAX QP: *Given a real $n \times n$ matrix $A = (a_{ij})$ maximize the quadratic form*

$$\langle x, Ax \rangle \equiv \sum_{ij} a_{ij} \cdot x_i x_j, \tag{1.1}$$

where variables x_i take values ± 1.

The problem naturally arises in the design of approximation algorithms for several combinatorial optimization problems. In these problems the objective function is or can be approximated by a second degree polynomial. Though MAX QP is $\mathcal{NP}$-hard, it can be approximately solved in polynomial time with a reasonable (constant or logarithmic) approximation guarantee.

The problem has been studied in the literature for different families of matrices A. Charikar and Wirth [18], Megretski [48] and Nemirovski, Roos and Terlaky [49] gave $\Omega(\log n)$-approximation algorithms for arbitrary matrices A with zeros on the diagonal[1]. Alon and Naor [7] presented a constant factor approximation

[1] We explain why this condition is necessary in Chapter 2.

algorithm for the *bilinear case* of MAX QP (see Chapter 2 for details), which they used to obtain a constant factor approximation for the cut norm. Finally, Nesterov [50] found a constant factor approximation algorithm for positive semidefinite matrices.

We prove that in general the $\log n$ approximation factor is optimal for algorithms based on semidefinite programming (the technique used in all the algorithms above). However, we show that in many cases — when the matrix A has a special form — we can get a much better approximation ratio. Let G be the graph on vertices $\{1, \dots, n\}$ corresponding to the support of matrix A: that is, $(i, j) \in E$, if $a_{ij} \neq 0$ and $(i, j) \notin E$, if $a_{ij} = 0$. We give an algorithm that finds an $O(\log \vartheta(\bar{G}))$ approximation, where $\vartheta(\bar{G})$ is the Lovász theta function. Let us note that the graph G often has a natural interpretation in terms of the original combinatorial optimization problem (that is reduced to MAX QP) and its support has a special structure (which may be known in advance). Therefore, in many applications, our algorithm performs considerably better than the general algorithm (e.g. see discussion of the Correlation Clustering Problem in the next section).

In the second part of the dissertation, we give new approximation algorithms for the Maximum Constraint Satisfaction Problem and the Maximum Acyclic Subgraph problem; both algorithms use quadratic integer optimization.

1.1 Algorithmic Applications

We now describe some applications of MAX QP.

Spin glass model. The first application arises in the study of the *spin glass model* in mathematical physics. In this model, there is a system of atoms each of which has a spin that can point either up or down. Some of the pairs of atoms have non-negligible interactions. The system can be described by a graph $G = (V, E)$, with a (not necessarily positive) weight (a_{ij}) for each edge $(i, j) \in E$. The atoms are the vertices, the spin of the atom i is $x_i \in \{-1, 1\}$, where 1 represents the "up" position, the weights (a_{ij}) represent the interactions between i and j, and the energy of the system is given by its *Hamiltonian* $H = -\sum_{(i,j)\in E} a_{ij}x_ix_j$. A *ground state* is a state that minimizes the energy, and thus the problem of finding a ground state is precisely that of finding the maximum of the integer program (1.1).

It is known that if the graph G is planar, one can find a ground set in polynomial time using matching algorithms (see Barahona [13]), but in general this problem is $\mathcal{NP}$-hard (and in fact even hard to approximate). Our technique here thus supplies an efficient way to find a configuration that approximates the minimum energy, and in many cases this approximation is up to a constant factor.

Correlation clustering. Typical clustering problems involve the partitioning of a data set into classes which are small in some quantitative (typically metric) sense. In contrast, in correlation clustering, first considered by Bansal, Blum and Chawla [12] (and referred to as *clustering with qualitative information* in Charikar, Guruswami and Wirth [15]), we are given a judgment graph $G = (V, E)$ and for every $(i, j) \in E$ a real number a_{ij} which is interpreted as a judgment of the similarity of i and j. In the simplest case $a_{ij} \in \{1, -1\}$, where if $a_{ij} = 1$ then i and j are said to be similar, and if $a_{ij} = -1$ then i and j are said to be dissimilar. Given a partition of V into clusters, a pair is called an agreement if it is a similar pair within one cluster or a dissimilar pair across two distinct clusters. Analogously, a disagreement is a similar pair across two different clusters or a dissimilar pair in one cluster. In the maximum correlation problem the goal is to partition V so that the correlation is maximized, where the correlation of a partition is the difference of the number of agreements and the number of disagreements. In the case of arbitrary weights, for a partition P of V into pairwise disjoint clusters, the value of the partition is the sum of all a_{ij} for i, j that lie in the same cluster, minus the sum of all entries a_{ij} for i and j that lie in distinct clusters.

Charikar and Wirth [18] showed that the general problem can be reduced to the problem of partitioning of V into two clusters. This problem, in turn, is equivalent to MAX QP: the correlation of a partition P equals

$$\sum_{(i,j)\in E} a_{ij} x_i x_j,$$

where $x_i = 1$, if the vertex i is in the first cluster; and $x_i = -1$, if i is in the second cluster.

Charikar and Wirth [18] developed $O(\log n)$ approximation algorithm for the maximum correlation clustering problem when the judgment graph is the complete graph on n vertices. Our results improve the approximation guarantee to $O(\log(\vartheta(\overline{G})))$. In particular, this is $O(1)$ for any bounded degree graph or any graph with a bounded chromatic number or genus — in all these cases and some additional ones considered in Section 2.8 no constant factor approximation algorithm was previously known.

Maximum Constraint Satisfaction Problem. In the MAX k-CSP problem we are given a set of boolean variables and constraints, where each constraint depends on k variables, the goal is to find an assignment that maximizes the number of satisfied constraints. The motivation for studying this problem comes from complexity theory (see Chapter 3). We give $\Omega(2^k/k)$ approximation algorithm for MAX k-CSP problem. Our result implies that the ratio between the completeness and soundness parameters of any k-query PCP for SAT is at most $2^k/k$. This bound matches (up to a constant factor) the lower bound on this ratio obtained by Samorodnitsky and Trevisan [56] assuming the Unique Games Conjecture.

Our algorithm for MAX k-CSP uses integer quadratic programming. However, the connection between the MAX k-CSP problem and the MAX QP problem is significantly more complex than in the previous two examples. We describe the algorithm in Chapter 3.

Maximum Acyclic Subgraph. The Maximum Acyclic Subgraph problem is a classical linear arrangement problem: Given a directed graph on n vertices, arrange the vertices on a line, so as to maximize the number of edges going forward. It is easy to obtain a 2 approximation for this problem: if we randomly permute all vertices, then half of all edges on average will go forward. It is a long standing open question, however, whether it is possible to find a $2 - \Omega(1)$ approximation.

We develop a new algorithm for this problem. It has the following performance guarantee: if in the optimal solution $(1/2 + \delta)$ fraction of all edges go forward, the algorithm returns a solution where $(1/2 + \Omega(\delta/\log n))$ fraction of all edges go forward. Its approximation ratio is still $2 - o(1)$, but its approximation guarantee is better than the approximation guarantee of known algorithms for a wide range of parameters.

Our algorithm is based on a new approach: we first partition the graph into several pieces (in a special way), then permute the vertices in each of them randomly and output the concatenation of the pieces. In order to find the partitioning, the algorithm solves a certain MAX QP problem and obtains probabilities with which each vertex should belong to each part. We describe the algorithm and quadratic relaxation in detail in Chapter 4.

1.2 Prior publications

Most of the dissertation was previously published in refereed conferences and journals.

- **Chapter 1** is based on the paper

  N. Alon, K. Makarychev, Y. Makarychev, A. Naor. *Quadratic Forms on Graphs.* In Inventiones Mathematicae, vol. 163, no. 3, pp. 499–522, March 2006. Conference version in the 37^{th} ACM STOC, pp. 486–493, 2005.

  However, some of the proofs we present in the dissertation were slightly simplified.

- **Chapter 2** is based on the paper

  M. Charikar, K. Makarychev, and Y. Makarychev. *Near-Optimal Algorithms for Maximum Constraint Satisfaction Problems.* In Proceedings of the 18th ACM-SIAM SODA, pp. 62–68, 2007

- **Chapter 3** is based on the paper

  M. Charikar, K. Makarychev, and Y. Makarychev. *On the Advantage over Random for Maximum Acyclic Subgraph.* Proceedings of 48^{th} IEEE FOCS, pp. 625–633, 2007.

Chapter 2

Quadratic Forms on Graphs

2.1 Introduction

In this chapter we study approximation algorithms for the following optimization problem.

MAX QP: *Given a real $n \times n$ matrix $A = (a_{ij})$ maximize the quadratic form*

$$\sum_{i,j=1}^{n} a_{ij} \cdot x_i x_j, \tag{2.1}$$

where variables x_i take values ± 1.

A natural approach to solving the problem is using Semidefinite Programming (SDP): we replace all variables x_i with unit vectors z_i in the n-dimensional space. The obtained relaxation

$$\max \sum_{i,j} a_{ij} \cdot \langle z_i, z_j \rangle \tag{2.2}$$

subject to $z_i \in S^{n-1}$ for all i, can be solved efficiently (i.e., in polynomial time) using Semidefinite Programming. However, the new problem is not equivalent to the original one. So the main questions is how well the solution to the SDP (SDP value) approximates the original solution (OPT value). The integrality gap GAP_A is the ratio between the SDP value and the OPT value:

$$GAP_A = \frac{\max_{z_i \in S^{n-1}} \sum_{i,j} a_{ij} \cdot \langle z_i, z_j \rangle}{\max_{x_i \in \{\pm 1\}} \sum_{i,j} a_{ij} \cdot x_i x_j}.$$

Note that since in the new problem the constraints on the variables are relaxed (variables x_i are unit vectors in one dimensional space; variables z_i are unit vectors in the n dimensional space) $SDP \geq OPT$ and $GAP_A \geq 1$.

This relaxation has been studied before for several families of matrices A. Rietz [55] and Nesterov [50] showed that the integrality gap is at most $\pi/2$ for positive semidefinite matrices. Charikar and Wirth [18], Megretski [48] and Nemirovski, Roos and Terlaky [49] proved $O(\log n)$ bound on the integrality gap of arbitrary matrices with zeros on the diagonal (where n is the size of the matrix). Kashin and Szarek [36] showed that the integrality gap is at least $\Omega(\sqrt{\log n})$ for such matrices. Alon and Naor [7] observed that Grothendieck's inequality [24] implies that the integrality gap is a constant in the *bilinear version* of MAX QP i.e. when the matrix A has the form

$$A = \begin{pmatrix} 0 & B \\ B^T & 0 \end{pmatrix}. \tag{2.3}$$

Grothendieck's inequality [24] states that for every $n \times m$ matrix (b_{ij}) and every choice of unit vectors $u_1, \ldots, u_n, v_1, \ldots, v_m \in S^{n+m-1}$ there exist a choice of signs $x_1, \ldots, x_n, y_1, \ldots, y_m \in \{-1, +1\}$ for which

$$\sum_{i=1}^{n} \sum_{j=1}^{m} b_{ij} \langle u_i, v_j \rangle \leq K_G \sum_{i=1}^{n} \sum_{j=1}^{m} b_{ij} x_i y_j. \tag{2.4}$$

Here K_G is an absolute constant, which is called the Grothendieck constant. The best value of K_G (which is unknown[1]) is exactly equal to the worst integrality gap GAP_A for matrices of the form (2.3).

In this chapter we generalize the Grothendieck inequality to matrices with a fixed support. We suppose that (a_{ij}) are weights on edges of a fixed graph G and investigate the dependence between the gap GAP_A and the graph G.

Let us now give a formal definition.

Definition 2.1. *Let $G = (V, E)$ be a graph on the vertices $\{1, \ldots, n\}$. Denote by $\mathcal{A}_G$ the family of all $n \times n$ matrices $A = (a_{ij})$ such that $a_{ij} = 0$, if $(i, j) \notin E$.*

Throughout the chapter we assume that the graph G is on the vertices $\{1, \ldots, n\}$.

Definition 2.2. *The Grothendieck constant of the graph G, denoted by $K(G)$, is the maximum integrality gap GAP_A among matrices A in $\mathcal{A}_G$.*

Note that bipartite graphs correspond to matrices of the form (2.3); and complete graphs (without loops) correspond to matrices with zeros on the diagonal. Hence, $K(G) \le K_G$ for bipartite graphs; and $K(K_n) = O(\log n)$ for the complete graph K_n on n vertices. We show that there is a natural way to interpolate between these two cases: namely, for every loop-free graph G,

$$K(G) = O(\log \vartheta(\overline{G})),$$

where $\vartheta(\overline{G})$ is the *Lovász theta function* of the complement of G, denoted $\overline{G}$. The theta function of a graph on the vertices $\{1, \ldots, n\}$, defined in [45], is the minimum

[1] The best known upper bound due to Krivine [41] is $K_G \le \pi/[2\log(1+\sqrt{2})] = 1.782...$

of

$$\max_{1\le i\le n} \frac{1}{\langle z_i, e\rangle^2},$$

where the minimum is taken over all choices of unit vectors z_i and e such that z_i and z_j are orthogonal for every pair of non-adjacent vertices i and j. Our proof is algorithmic, and provides an efficient randomized algorithm that approximates the maximum possible value of a given quadratic form.

We also show that the integrality gap of $O(\log n)$ for arbitrary $n \times n$ matrices with zeros on the diagonal is in fact optimal. This implies that for the complete graph K_n,

$$K(K_n) \ge \Omega(\log n);$$

and for arbitrary graph G,

$$K(G) \ge \Omega(\log \omega(G)),$$

where $\omega(G)$ is the clique number of G. Our construction is based on a refinement of a recent result of Kashin and Szarek [36], who established an $\Omega(\sqrt{\log n})$ lower bound.

2.2 Our results

To summarize the main results of this chapter are the following theorems.

Theorem 2.3. *There exists an absolute positive constant C such that for every (loop-free) graph G,*

$$K(G) \le C \log \vartheta(\bar{G}),$$

where $\vartheta(\bar{G})$ is the Lovász theta function of the complement of G.

Theorem 2.4. *There exists an absolute positive constant C such that for every (loop-free) graph G,*

$$K(G) \geq C \log \omega(G),$$

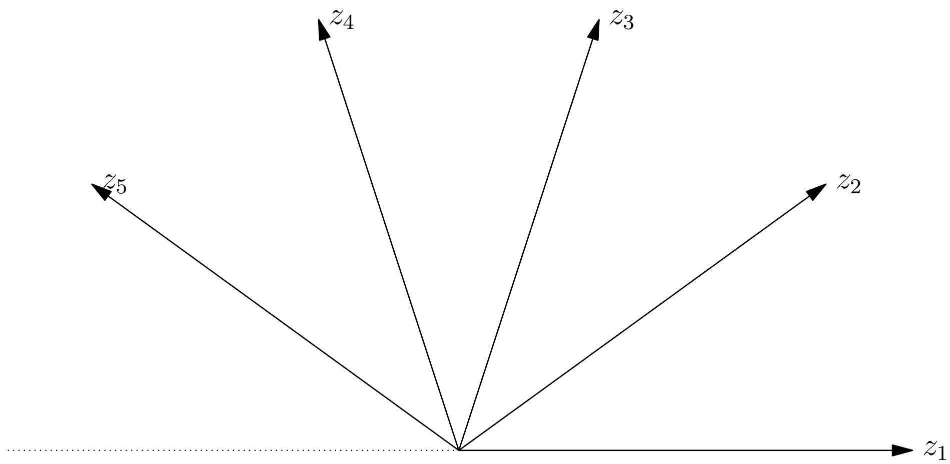
where $\omega(G)$ is the clique number of the graph G.

Remark 2.5. *Note that we are interested only in the case of simple graphs (i.e. graphs without loops). If G has loops, then $K(G)$ is equal to infinity, unless G is a forest. That is, there exists a matrix A in $\mathcal{A}_G$ such that (2.2) is positive, but (2.1) is negative or zero. Consider an example. The graph G contains a cycle $1 \to 2 \to \cdots \to m \to 1$ and a loop $l \to l$. Then the following quadratic form in $\mathcal{A}_G$ is non-negative for $x_1, \ldots, x_n \in \{-1, 1\}$:*

$$\sum_{i=1}^{m-1} x_i x_{i+1} - x_1 x_m - (m-2)(x_\ell)^2,$$

since $(x_1 x_2 + \cdots + x_{m-1} x_m) - x_1 x_m \leq m - 2$ and $(m-2)(x_\ell)^2 = m - 2$. However, if $z_1, \ldots, z_m$ are vectors on the plain such that the angle between z_i and z_{i+1} is π/m for all i; the angle between z_1 and z_m is $\pi - \pi/m$, then

$$\sum_{i=1}^{m} \langle z_i, z_{i+1} \rangle - \langle x_1, x_m \rangle - (m-2)\|z_l\|^2 = m \cos(\pi/m) - (m-2) > 0.$$



Finally, note that if G is a forest, then trivially $K(G) = 1$.

2.3 Basic Facts

Observation 2.6. *For every $n \times n$ matrix $A = (a_{ij})$ with non-negative entries on the diagonal*

$$\max_{x_i \in [-1,1]} \sum_{i,j} a_{ij} \cdot x_i x_j = \max_{x_i \in \{-1,+1\}} \sum_{i,j} a_{ij} \cdot x_i x_j;$$
$$\max_{\|z_i\|_2 \le 1} \sum_{i,j} a_{ij} \cdot \langle z_i, z_j \rangle = \max_{\|z_i\|_2 = 1} \sum_{i,j} a_{ij} \cdot \langle z_i, z_j \rangle.$$

Proof. Since the diagonal entries are non-negative, the quadratic form is convex in each variable and thus its maximum is attained at the boundary. $\square$

An immediate consequence of this observation is that if the diagonal entries of A are non-positive, then the integrality gap GAP_A is equal to

$$GAP'_A = \frac{\max_{\|z_i\|_2 \le 1} \sum_{i,j} a_{ij} \cdot \langle z_i, z_j \rangle}{\max_{|x_i| \le 1} \sum_{i,j} a_{ij} \cdot \langle x_i, x_j \rangle}.$$

It will be often more convenient for us, to search for an approximate solution among $x_i \in [-1,1]$. Then we can sequentially round each x_i to -1 or 1 without decreasing the value of the objective function. We shall also consider a slightly modified SDP relaxation which is equivalent to (2.2) if the diagonal entries of A are non-negative:

$$\max_{\|z_i\|_2 \le 1} \sum_{i,j} a_{ij} \cdot \langle z_i, z_j \rangle. \tag{2.2$'$}$$

2.4 Dual Problem

In this section, we introduce a parameter of the graph G *dual* to $K(G)$ and prove its basic properties. We need the following standard definitions. Denote by μ the Lebesgue measure on $\mathbb{R}$. Let f be a measurable function from $[0,1]$ to $\mathbb{R}$. Then $\|f\|_\infty$ is the (essential) supremum of $|f|$ (in other words, $\mu\{t : |f(t)| > \|f\|_\infty\} = 0$, but for every positive ε, $\mu\{t : |f(t)| \ge \|f\|_\infty - \varepsilon\} > 0$.). The class $L_\infty[0,1]$ is the class of all (essentially) bounded functions on the segment $[0,1]$ i.e. functions with finite $\|\cdot\|_\infty$ norm. Finally, the inner product between two functions f and g is defined as

$$\langle f, g \rangle \equiv \int_0^1 f(t)g(t)dt;$$

and the $\|\cdot\|_2$ norm of f is

$$\|f\|_2 = \sqrt{\langle f, f \rangle} \equiv \sqrt{\int_0^1 f(t)^2 dt}.$$

Definition 2.7 (The Gram representation constant of G). *Let $G = (V, E)$ be a graph on n vertices. Denote by $R(G)$ the minimum $R \in \mathbb{R}$ such that for every sequence of unit vectors $z_1, \dots, z_n \in S^{n-1}$ there exists a sequence of functions $f_1, \dots, f_n$ in*

$L_\infty[0,1]$ *such that*

- *for every* $i \in V$ *we have* $\|f_i\|_\infty \le R$*; and*

- *for every* $(i,j) \in E$*,*

$$\langle z_i, z_j \rangle = \langle f_i, f_j \rangle \equiv \int_0^1 f_i(t) f_j(t) dt.$$

Our goal is to show that $R(G)^2 = K(G)$ for graphs without loops. For clarity of presentation we will split the proof in two parts. First using a probabilistic argument we will prove that $K(G) \le R(G)^2$. Then we will define the Gram representation constant in terms of Gram matrices and show that certain sets of Gram matrices are convex. This will allow us to use a duality argument to prove that $K(G) \ge R(G)^2$.

Theorem 2.8. *For every (loop-free) graph* G*,*

$$K(G) = R(G)^2.$$

Lemma 2.9. *For every (loop-free) graph* G*,* $K(G) \le R(G)^2$*.*

Proof. Fix a $n \times n$ matrix A in $\mathcal{A}_G$. Let $z_1, \dots, z_n$ be a sequence of unit vectors that maximizes the quadratic form

$$\sum_{(i,j)} a_{ij} \cdot \langle z_i, z_j \rangle = \sum_{(i,j) \in E} a_{ij} \cdot \langle z_i, z_j \rangle.$$

By the definition of $R(G)$ there exists a sequence of functions such that $\langle f_i, f_j \rangle =$

$\langle z_i, z_j \rangle$ for all $(i,j) \in E$ and $|f_i(t)| \le R(G)$ for all i and $t \in [0,1]$. Observe that

$$\sum_{(i,j)\in E} a_{ij} \cdot \langle z_i, z_j \rangle = \sum_{(i,j)\in E} a_{ij} \cdot \langle f_i, f_j \rangle = \int_0^1 \sum_{(i,j)\in E} a_{ij} \cdot f_i(t) f_j(t) dt.$$

Hence there exists $t_0 \in [0,1]$ for which

$$\sum_{(i,j)\in E} a_{ij} \cdot f_i(t_0) f_j(t_0) \ge \sum_{(i,j)\in E} a_{ij} \cdot \langle z_i, z_j \rangle.$$

Let $x_i = f_i(t_0)/R(G)$, then $x_i \in [-1,1]$ and

$$R(G)^2 \sum_{i,j} a_{ij} \cdot x_i x_j = \sum_{i,j} a_{ij} \cdot f_i(t_0) f_j(t_0) \ge \sum_{i,j} a_{ij} \cdot \langle z_i, z_j \rangle.$$

This inequality (together with Observation 2.6) implies that $K(G) \le R(G)^2$. $\square$

For every sequence of functions $f_1, \ldots, f_n$ in $L_\infty[0,1]$ define its Gram matrix restricted to the graph $G = (V, E)$ as follows

$$\mathcal{G}_G(f_1, \ldots, f_n)_{ij} = \begin{cases} \langle f_i, f_j \rangle, & \text{if } (i,j) \in E; \\ 0, & \text{otherwise.} \end{cases}$$

Let $\mathcal{F}_G(R)$ be the set of Gram matrices (restricted to G) of functions $f_1, \ldots f_n$ with $\|\cdot\|_\infty$ norm bounded by R:

$$\mathcal{F}_G(R) = \left\{ \mathcal{G}_G(f_1, \ldots, f_n) : \|f_i\|_\infty \le R \text{ for all } i \right\}.$$

Similarly define $\mathcal{G}_G(z_1, \ldots, z_n)$ for arbitrary vectors $z_1, \ldots, z_n$:

$$\mathcal{G}_G(z_1, \ldots, z_n)_{ij} = \begin{cases} \langle z_i, z_j \rangle, & \text{if } (i,j) \in E; \\ 0, & \text{otherwise;} \end{cases}$$

and let $\mathcal{Z}_G$ be the set of Gram matrices of all sets of unit vectors $z_1, \ldots, z_n$:

$$\mathcal{Z}_G = \{\mathcal{G}_G(z_1, \ldots, z_n) : \|z_i\|_2 = 1\}.$$

Observation 2.10. *The Gram representation constant $R(G)$ is the minimum R for which the set $\mathcal{F}_G(R)$ contains the set $\mathcal{Z}_G$.*

Proof. The set $\mathcal{Z}_G$ is a subset of $\mathcal{F}_G(R)$, if and only if for all unit vectors $z_1, \ldots, z_n$ there exist functions $f_1, \ldots, f_n$ with $\|f_i\|_\infty \leq R$ such that

$$\mathcal{G}_G(z_1, \ldots, z_n) = \mathcal{G}_G(z_1, \ldots, z_n)$$

which is equivalent to

$$\langle z_i, z_j \rangle = \langle f_i, f_j \rangle \text{ for all } (i,j) \in E.$$

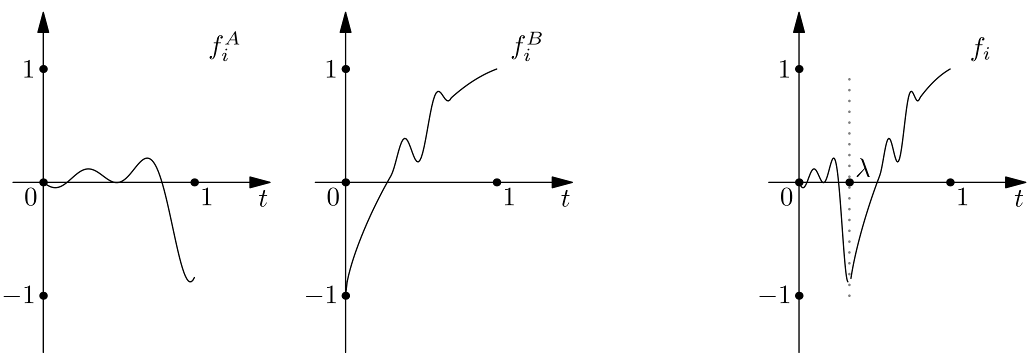
□

Lemma 2.11. *The sets $\mathcal{F}_G(R)$ and $\mathcal{Z}_G$ are convex.*

Proof. I. First we show that the set $\mathcal{F}_G(R)$ is convex. Fix $\lambda \in (0, 1)$ and consider arbitrary Gram matrices

$$\begin{aligned} g^A &= \mathcal{G}_G(f_1^A, \ldots, f_n^A) \in \mathcal{F}_G(R); \\ g^B &= \mathcal{G}_G(f_1^B, \ldots, f_n^B) \in \mathcal{F}_G(R). \end{aligned}$$

We need to prove that $\lambda g^A + (1-\lambda) g^B \in \mathcal{F}_G(R)$.



Define

$$f_i(t) = \begin{cases} f_i^A(t/\lambda), & \text{if } t \in [0, \lambda]; \\ f_i^B\left(\frac{t-\lambda}{1-\lambda}\right), & \text{if } t \in [\lambda, 1]. \end{cases}$$

Then

$$\begin{aligned} \langle f_i, f_j \rangle &= \int_0^\lambda f_i^A(t/\lambda) f_j^A(t/\lambda) dt + \int_\lambda^1 f_i^B\left(\frac{t-\lambda}{1-\lambda}\right) f_j^B\left(\frac{t-\lambda}{1-\lambda}\right) dt \\ &= \lambda \langle f_i^A, f_j^A \rangle + (1-\lambda) \langle f_i^B, f_j^B \rangle \end{aligned}$$

and we are done.

II. The proof that the set $\mathcal{Z}_G$ is convex is similar. Consider two matrices

$$\begin{aligned} g^A &= \mathcal{G}_G(z_1^A, \ldots, z_n^A) \in \mathcal{Z}_G; \\ g^B &= \mathcal{G}_G(z_1^B, \ldots, z_n^B) \in \mathcal{Z}_G. \end{aligned}$$

Define $z_i = (\sqrt{\lambda}\, z_i^A) \oplus (\sqrt{1-\lambda}\, z_i^B)$. Then z_i are unit vectors and

$$\begin{aligned}\langle z_i, z_j\rangle &= \langle \sqrt{\lambda}\, z_i^A, \sqrt{\lambda}\, z_j^A\rangle + \langle \sqrt{1-\lambda}\, z_i^B, \sqrt{1-\lambda}\, z_j^B\rangle \\ &= \lambda\, \langle z_i^A, z_j^A\rangle + (1-\lambda)\, \langle z_i^B, z_j^B\rangle.\end{aligned}$$

□

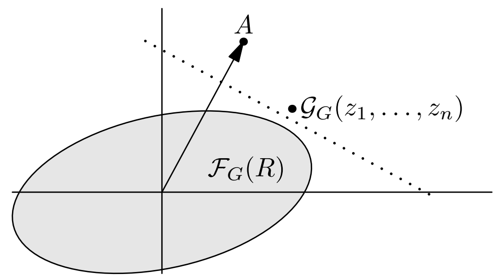
Lemma 2.12. *Let G be a graph without loops. Then $K(G) \geq R(G)^2$.*

Proof. We prove that $\mathcal{Z}_G \subset \mathcal{F}_G(\sqrt{K(G)})$ and thus $R(G)^2 \leq K(G)$ (see Observation 2.10). Suppose to the contrary, that there exists a set of unit vectors $z_1, \ldots, z_n$ such that the Gram matrix $\mathcal{G}_G(z_1, \ldots, z_n)$ is not in the set $\mathcal{F}_G(\sqrt{K(G)})$. Since the set $\mathcal{F}_G(\sqrt{K(G)})$ is convex there exists a hyperplane separating it from $\mathcal{G}_G(z_1, \ldots, z_n)$. In other words, there exists a set of weights a_{ij} such that for every $g \in \mathcal{F}_G(\sqrt{K(G)})$,

$$\sum_{i,j} a_{ij} \cdot g_{ij} < \sum_{i,j} a_{ij} \cdot \langle z_i, z_j\rangle I((i,j) \in E); \tag{2.5}$$

where I denotes the indicator function. If $(i,j) \notin E$, then $g_{ij} = I((i,j) \in E) = 0$, hence we can assume that a_{ij} is also 0, for $(i,j) \notin E$. Then $A \in \mathcal{A}_G$ and

$$\sum_{i,j} a_{ij} \cdot g_{ij} < \sum_{i,j} a_{ij} \cdot \langle z_i, z_j\rangle. \tag{2.6}$$



On the other hand, by the definition of $K(G)$ there exists a sequence $x_1, \ldots, x_n \in \{\pm 1\}$ such that

$$\sum_{i,j} a_{ij} \cdot \left(x_i \sqrt{K(G)}\right) \left(x_j \sqrt{K(G)}\right) \geq \sum_{i,j} a_{ij} \cdot \langle z_i, z_j \rangle$$

and hence for constant functions $f_i(t) = x_i \sqrt{K(G)}$

$$\sum_{i,j} a_{ij} \cdot \langle f_i, f_j \rangle \geq \sum_{i,j} a_{ij} \cdot \langle z_i, z_j \rangle.$$

This contradicts the inequality (2.6) and finishes the proof. $\square$

2.5 Upper Bounds

In this section, we present an algorithm for MAX QP and analyze its approximation ratio for different families of matrices.

Let MIN_A and MAX_A be the minimum and maximum of the quadratic form $\sum_{i,j} a_{ij} \langle z_i, z_j \rangle$ over vectors $z_1, \ldots, z_n$ in the unit ball:

$$\begin{aligned} MAX_A &= \max_{\|z_i\|_2 \leq 1} \sum_{i,j} a_{ij} \langle z_i, z_j \rangle; \\ MIN_A &= \min_{\|z_i\|_2 \leq 1} \sum_{i,j} a_{ij} \langle z_i, z_j \rangle; \end{aligned}$$

and let

$$\alpha_A = -\frac{MIN_A}{MAX_A}.$$

We point out that MIN_A is non-positive, and therefore α_A is non-negative. Also notice that MIN_A and MAX_A can be computed efficiently using Semidefinite Programming. We need the following simple observation.

Lemma 2.13. *Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be (correlated) random variables with mean 0.*

1. *If* $\mathrm{Var}[X_i] \leq 1$ *for all i, then*

$$MIN_A \leq \sum_{i,j} a_{ij} \langle X_i, X_j \rangle \leq MAX_A.$$

2. *If* $\mathrm{Var}[X_i] \leq \alpha^2$ *and* $\mathrm{Var}[Y_i] \leq \beta^2$ *for all i, then*

$$\sum_{i,j} a_{ij} \langle X_i, Y_j \rangle \leq \alpha\beta \cdot (MAX_A - MIN_A).$$

Proof. I. Consider the linear space of random variables spanned by $X_1, \ldots, X_n$ with inner product between two random variables equal to their covariance. This space is isometric to Euclidean space (as every linear space over $\mathbb{R}$ with positive definite inner product). So the first inequality follows from the definition of MIN_A and MAX_A.

II. Let $\widehat{X}_i = \alpha^{-1}X_i/2$ and $\widehat{Y}_i = \beta^{-1}Y_i/2$. Write

$$\begin{aligned}\sum_{i,j} a_{ij}\langle X_i, Y_j\rangle &= 4\alpha\beta \cdot \sum_{i,j} a_{ij}\langle \widehat{X}_i, \widehat{Y}_j\rangle \\ &= \alpha\beta \cdot \left(\sum_{i,j} a_{ij}\langle \widehat{X}_i + \widehat{Y}_i, \widehat{X}_j + \widehat{Y}_j\rangle - \sum_{i,j} a_{ij}\langle \widehat{X}_i - \widehat{Y}_i, \widehat{X}_j - \widehat{Y}_j\rangle\right).\end{aligned}$$

Notice that $\mathrm{Var}\left[\widehat{X}_i + \widehat{Y}_i\right] \leq 1$ and $\mathrm{Var}\left[\widehat{X}_i - \widehat{Y}_i\right] \leq 1$ for all i, hence the right hand side is bounded by $\alpha\beta \cdot (MAX_A - MIN_A)$. □

Lemma 2.14. *I. For every $n \times n$ matrix $A = (a_{ij})$,*

$$GAP'_A \leq C_1 + C_2 \log\left(1 + |\alpha_A|\right),$$

for some absolute constants C_1 and C_2. Moreover, there exists a randomized polynomial time algorithm that given a matrix A computes $x_1, \ldots x_n \in [-1, 1]$ such that

$$\sum_{i,j} a_{ij} \cdot x_i x_j \geq \frac{MAX_A}{C_1 + C_2 \log\left(1 + |\alpha_A|\right)}.$$

II. For every $n \times n$ matrix $A = (a_{ij})$ with non-negative diagonal entries,

$$GAP_A \leq C_1 + C_2 \log\left(1 + |\alpha_A|\right).$$

Proof. I. Let $z_1, \ldots, z_n \in S^{n-1}$ be a solution of the SDP relaxation (2.2):

$$\sum_{i,j} a_{ij}\langle z_i, z_j\rangle = MAX_A.$$

Our goal is to construct random variables X_i taking values in $[-1, 1]$ such that

$$\sum_{i,j} a_{ij}\langle X_i, X_j\rangle = \mathbb{E}\left[\sum_{i,j} a_{ij} X_i X_j\right] \geq \frac{MAX_A}{C_1 + C_2 \log(1 + |\alpha_A|)}.$$

Then with constant probability (by the Markov inequality)

$$MAX_A \leq (C_1' + C_2' \log(1 + |\alpha_A|)) \cdot \sum_{i,j} a_{ij} X_i X_j,$$

which will conclude the proof.

Pick a random Gaussian vector g in the n-dimensional space with coordinates i.i.d. as $\mathcal{N}(0, 1)$. Project g on each vector z_i and denote the lengths of the projections by P_i. In other words, $P_i = \langle g, z_i\rangle$. Note that each P_i is a normal random variable with mean 0, variance $\|z_i\|_2^2$ and

$$\langle P_i, P_j\rangle \equiv \mathbb{E}[P_i P_j] = \langle z_i, z_j\rangle.$$

This follows from the equality $\mathbb{E}[\langle u, g\rangle, \langle v, g\rangle] = \langle u, v\rangle$ that holds[2] for all vectors u and v. Hence,

$$MAX_A = \sum_{i,j} a_{ij} \cdot \langle X_i, X_j\rangle.$$

Fix $M = 4 + 4\sqrt{\log(1 + |\alpha_A|)}$ and consider truncated random variables P_i^M:

$$P_i^M = \begin{cases} P_i, & \text{if } P_i \in [-M, M]; \\ 0, & \text{otherwise.} \end{cases}$$

[2]This equality holds for all basis vectors u and v just by the definition of g. We can extend it to all u and v in $\mathbb{R}^n$, because both functions $\mathbb{E}[\langle u, g\rangle, \langle v, g\rangle]$ and $\langle u, v\rangle$ are linear in u and v.

Let $T_i = P_i - P_i^M$. Write

$$\begin{aligned}\sum_{i,j} a_{ij} \cdot \langle P_i^M, P_j^M \rangle &= \sum_{i,j} a_{ij} \cdot \langle P_i - T_i, P_j - T_j \rangle \\ &= MAX_A - 2\sum_{i,j} a_{ij} \cdot \langle P_i, T_j \rangle + \sum_{i,j} a_{ij} \cdot \langle T_i, T_j \rangle.\end{aligned} \tag{2.7}$$

We need to estimate the terms

$$-\sum_{i,j} a_{ij} \cdot \langle P_i, T_j \rangle \text{ and } \sum_{i,j} a_{ij} \cdot \langle T_i, T_j \rangle. \tag{2.8}$$

The variance of T_i is equal to

$$\mathrm{Var}\,[T_i] = \sqrt{\frac{2}{\pi}} \int_M^\infty t^2 e^{-t^2/2} dt \leq \sqrt{\frac{3}{\pi}} M e^{-M^2/2}.$$

The inequality above holds since (1) when M tends to infinity, both sides of the inequality tend to 0; (2) the derivative of the left hand side ($-\sqrt{2/\pi} M^2 e^{-M^2/2}$) is greater than the derivative of the right hand side ($\sqrt{3/\pi}(1-M^2)e^{-M^2/2}$) for $M \geq 3$. Plugging M in we get

$$\mathrm{Var}\,[T_i] \leq e^{-M^2/4} \leq \frac{1}{50} \times \left(\frac{MAX_A + |MIN_A|}{MAX_A} \right)^2.$$

Hence, by Lemma 2.13, expressions (2.8) are greater than $-MAX_A/7$. Let us now return to inequality (2.7):

$$\sum_{i,j} a_{ij} \cdot \langle P_i^M, P_j^M \rangle \geq MAX_A - \frac{3}{7} MAX_A \geq \frac{1}{2} MAX_A.$$

1. Solve the Semidefinite Program (2.2). Denote by z_i the solution and let MAX_A be the SDP value.

2. Compute MIN_A and denote $\alpha_A = -MIN_A/MAX_A$.
   Set $M = 4 + 4\sqrt{\log(1 + |\alpha_A|)}$.

3. Pick a random Gaussian vector g in the n-dimensional space with coordinates i.i.d. as $\mathcal{N}(0, 1)$.

4. For each i,

   - Let $P_i = \langle g, z_i \rangle$.
   - Let
   $$P_i^M = \begin{cases} P_i, & \text{if } P_i \in [-M, M]; \\ 0, & \text{otherwise.} \end{cases}$$
   - Let $x_i = P_i^M/M$.

5. Return $x_1, \dots, x_n$.

Figure 2.1: Approximation Algorithm for MAX QP

We are almost done. Let $X_i = P_i^M/M$. Then (for appropriate C_1 and C_2)

$$\sum_{i,j} a_{ij} \cdot \langle X_i, X_j \rangle \geq \frac{MAX_A}{2M^2} \geq \frac{MAX_A}{C_1 + C_2 \log(1 + |\alpha|)}.$$

The proof we just described immediately gives an algorithm for finding x_i (see Figure 2.1).

II. If the diagonal entries of A are non-negative, then due to Observation 2.6, $GAP_A = GAP'_A$. □

Remark 2.15. *Rietz [55] and Nesterov [50] showed that there exists a $\pi/2$ approximation algorithm for* MAX QP *for positive semidefinite matrices A. Note that our algorithm also gives a constant factor approximation (with suboptimal constant), since if A is positive semidefinite, then $\alpha_A = 0$ and the diagonal entries of A are non-negative.*

The next theorem gives a good upper bound for matrices whose entries are highly non-uniform.

Theorem 2.16. *There exists a positive constant C such that for every $n \times n$ matrix $A = (a_{ij})$:*

$$\max_{\|z_i\|_2 \le 1} \sum_{i,j=1}^{n} a_{ij} \langle z_i, z_j \rangle \le C \log \left(\frac{\sum_{i,j=1}^{n} |a_{ij}|}{\operatorname{tr} A + \sqrt{2 \sum_{i \ne j} a_{ij}^2}} \right) \cdot \max_{|x_i| \le 1} \sum_{i,j=1}^{n} a_{ij} \cdot x_i x_j.$$

Here $\operatorname{tr} A = a_{11} + \cdots + a_{nn}$ *denotes the trace of the matrix A.*

Proof. Without loss of generality assume that the matrix A is symmetric. We need to show that

$$|\alpha_A| \equiv \frac{|MIN_A|}{MAX_A} \le \frac{\sum_{i,j=1}^{n} |a_{ij}|}{\operatorname{tr} A + \sqrt{2 \sum_{i \ne j} a_{ij}^2}}.$$

Clearly,

$$|MIN_A| \equiv |\min \sum_{i,j=1}^{n} a_{ij} \langle z_i, z_j \rangle| \le \sum_{i,j=1}^{n} |a_{ij}|.$$

We shall prove that

$$MAX_A \equiv \max_{\|z_i\|=1} \sum_{i,j=1}^{n} a_{ij} \langle z_i, z_j \rangle \ge \max_{x_i \in \{\pm 1\}} \sum_{i,j=1}^{n} a_{ij} x_i x_j \ge \operatorname{tr} A + \sqrt{2 \sum_{i \ne j} a_{ij}^2}.$$

Let $X_1, \ldots, X_n$ be i.i.d. Bernoulli random variable taking values $\{-1, 1\}$ with probability 1/2. Consider the random variable

$$\sum_{i<j} a_{ij} \cdot X_i X_j.$$

Notice that the terms in this sum are pairwise independent random variables and the expectation of each of them is zero. Therefore,

$$\mathrm{Var}\left[\sum_{i<j} a_{ij} \cdot X_i X_j\right] = \sum_{i<j} \mathrm{Var}\left[a_{ij} \cdot X_i X_j\right] = \sum_{i<j} a_{ij}^2,$$

and

$$\mathbb{E}\left[\left(\sum_{i\neq j} a_{ij} \cdot X_i X_j\right)^2\right] = \mathrm{Var}\left[\sum_{i\neq j} a_{ij} \cdot X_i X_j\right] = 4\sum_{i<j} a_{ij}^2.$$

Hence there exist $x_1, \ldots, x_n \in \{\pm 1\}$ such that

$$\left(\sum_{i\neq j} a_{ij} \cdot x_i x_j\right)^2 \geq 4\sum_{i<j} a_{ij}^2.$$

We get

$$\sum_{i,j=1}^{n} a_{ij} \cdot x_i x_j = \mathrm{tr}\, A + \sum_{i\neq j} a_{ij} \cdot x_i x_j \geq \mathrm{tr}\, A + \sqrt{2\sum_{i\neq j} a_{ij}^2}.$$

$\square$

We now define strict vector coloring and then bound α_A in terms of the strict vector chromatic number.

Definition 2.17. *A strict vector k-coloring of a graph on vertices $\{1, \ldots, n\}$ is a collection of unit vectors $v_1, \ldots, v_n$ such that for every two adjacent vertices i and j*

$$\langle v_i, v_j \rangle = -\frac{1}{k-1}.$$

A graph is strictly vector k-colorable if it has a strict vector k-coloring. The strict vector chromatic number of a graph is the smallest real k for which the graph is strictly vector k-colorable. The strict vector chromatic number is denoted by $\chi_{vec}(G)$.

The notion of strict vector colorability was introduced by Karger, Motwani, and Sudan [35]. They showed that the strict vector chromatic number of a graph G is equal to the Lovász theta function of the complement of G, $\vartheta(\overline{G})$. Note that the strict vector chromatic number of graph G is always bounded by the chromatic number of G:

$$\chi_{vec}(G) \leq \chi(G).$$

We need the following new characterization of the strict vector chromatic number.

Theorem 2.18. *For every graph $G = (V, E)$ without loops*

$$\chi_{vec}(G) - 1 = \max_{A \in \mathcal{A}_G} \alpha_A \equiv \max_{A \in \mathcal{A}_G} \frac{-MIN_A}{MAX_A}.$$

We split the proof into two lemmas.

Lemma 2.19. *For every graph G and matrix $A = (a_{ij})$ in $\mathcal{A}_G$,*

$$\chi_{vec}(G) - 1 \geq \frac{-MIN_A}{MAX_A}.$$

Proof. Let $v_1, \ldots, v_n$ be the optimal strict vector coloring and let $z_1, \ldots z_n$ be arbitrary vectors in the unit ball. Define $w_i = v_i \otimes z_i$. Notice that the lengths of the vectors w_i do not exceed 1, hence

$$\begin{aligned} MAX_A &\geq \sum_{i,j} a_{ij}\langle w_i, w_j\rangle = \sum_{i,j} a_{ij}\langle v_i, v_j\rangle \cdot \langle z_i, z_j\rangle \\ &= -\frac{1}{\chi_{vec}(G)-1}\left(\sum_{i,j} a_{ij}\langle z_i, z_j\rangle\right). \end{aligned}$$

Therefore, $MAX_A \geq -MIN_A/(\chi_{vec}(G)-1)$. ☐

Lemma 2.20. *For every (loop-free) graph G,*

$$\chi_{vec}(G) - 1 \leq \max_{A \in \mathcal{A}_G} \alpha_A.$$

Proof. The proof of the lemma uses a duality argument and is similar to the proof of Lemma 2.12. Denote $\alpha = \max_{A\in\mathcal{A}_G} \alpha_A$, and assume to the contrary that $\alpha < \chi_{vec}(G) - 1$. This means that the graph G is not strictly $(\alpha+1)$-colorable. In other words, the matrix

$$g_{ij} \equiv \begin{cases} -\alpha^{-1}, & \text{if } (i,j) \in E; \\ 0, & \text{otherwise}; \end{cases}$$

is not a Gram matrix restricted to G of any set of unit vectors i.e. $(g_{ij}) \notin \mathcal{Z}_G$. Hence there is a hyperplane separating the matrix (g_{ij}) and the set of matrices $\mathcal{Z}_G$: there exists a matrix (a_{ij}) in $\mathcal{A}_G$ such that

$$\sum_{(i,j)\in E} a_{ij} g_{ij} > \sum_{(i,j)\in E} a_{ij} \cdot \langle z_i, z_j\rangle$$

for every $z_1, \ldots, z_n \in S^{n-1}$. Particularly,

$$\sum_{(i,j)\in E} a_{ij} g_{ij} > MAX_A.$$

Substituting $g_{ij} = -\alpha^{-1}$ (for $(i,j) \in E$) we get

$$-\alpha^{-1} \sum_{(i,j)\in E} a_{ij} > MAX_A.$$

Hence

$$\alpha < \frac{-\sum_{(i,j)\in E} a_{ij}}{MAX_A} = \frac{-\sum_{(i,j)\in E} a_{ij} \times 1 \cdot 1}{MAX_A} \leq \frac{-MIN_A}{MAX_A}.$$

This inequality contradicts the definition of α and we are done. □

We now prove Theorem 2.3.

Theorem 2.3. *There exists a constant C such that for every (loop-free) graph G,*

$$K(G) \leq C \log \vartheta(\bar{G}),$$

where $\vartheta(\bar{G})$ is the Lovász theta function of the complement of G.

Proof. By Theorem 2.16 (see also Observation 2.6),

$$K(G) \leq \max_{A\in\mathcal{A}_G} \left(C_1 + C_2 \log \left(1 + \frac{|MIN_A|}{MAX_A} \right) \right).$$

By Theorem 2.18,

$$\max_{A\in\mathcal{A}_G} \frac{-MIN_A}{MAX_A} = \chi_{vec}(G) = \vartheta(\bar{G}).$$

Hence,

$$K(G) \leq \vartheta(\bar{G}).$$

$\square$

Remark 2.21. *Since the strict vector chromatic number is less than or equal to the chromatic number $\chi(G)$, the result above implies that for any loop-free graph G,*

$$K(G) = O\left(\log \chi(G)\right).$$

2.6 Tight Lower Bound for Complete Graphs

In this section we prove that the Gram representation constant $R(K_n)$ of the complete graph K_n is $\Omega(\sqrt{\log n})$, and thus the Grothendieck constant $K(K_n)$ of the complete graph is $\Omega(\log n)$. We construct an explicit example of vectors z_i for which every set of functions f_i satisfying $\langle f_i, f_j \rangle = \langle z_i, z_j \rangle$ for $i \neq j$ contains a function with $\|\cdot\|_\infty$ norm greater than $\Omega(\sqrt{\log n})$. Our proof is a refinement of the argument of Kashin and Szarek [36] and is also based on a 1-net on the unit sphere.

Definition 2.22. *A set of points S on the unit sphere S^{d-1} forms an ε-net if the distance from every point z on the sphere to the set S is at most ε:*

$$d_2(z, S) \equiv \min\left\{\|z - z'\|_2 : z' \in S\right\} \leq \varepsilon.$$

We prove a well-known bound on the size of a 1-net on S^{d-1}.

Lemma 2.23. *For every d there exists a 1-net S on the sphere S^{d-1} of size at most 3^d.*

Proof. Consider a maximal subset of points S of the sphere S^{d-1} such that the distance between every two points in S is at least 1. Clearly, S forms a 1-net:

indeed, if there existed a point z in S^{d-1} such that $d_2(z, S) > 1$ we could add z to S, which would contradict the maximality of the set S.

To estimate the size of S notice that balls of radius $1/2$ around points in S do not intersect with each other and all of them lie in the ball of radius $3/2$ around the origin. Therefore, the number of points in S is at most the ratio between the volumes of balls of radius $3/2$ and $1/2$ in the d dimensional space, that is 3^d. □

We need the following simple property of 1-nets.

Lemma 2.24. *Suppose vectors $z_1, \dots, z_n$ form a 1-net on the unit sphere S^{d-1}. Then for every vector v in $\mathbb{R}^d$,*

$$\max_m \langle z_m, v \rangle \geq \frac{\|v\|_2}{2}.$$

Proof. If $v = 0$, then both sides are equal to 0. Suppose $v \neq 0$. Let $\hat{v} = v/\|v\|_2$. The vector $\hat{v}$ lies on the unit sphere and hence the exists z_m such that $\|\hat{v} - z_m\| \leq 1$. The angle between $\hat{v}$ and z_m is at most $\pi/3$. Thus, $\langle z_m, v \rangle \geq \|v\|_2/2$. □

In the definition of $R(K_n)$ we require that $\langle f_i, f_j \rangle = \langle z_i, z_j \rangle$ for $i \neq j$, but $\langle f_i, f_i \rangle$ can be arbitrary. It turns out, however, that we can always find functions f_i with $\| \cdot \|_2$ norm very close to 1. The following lemma gives a quantitative bound on the norm. A similar statement was used (implicitly) by Kashin and Szarek [36].

Lemma 2.25. *For every set of unit vectors $z_1, \dots, z_n \in S^{n-1}$ there exists a sequence of functions $f_1, \dots, f_n \in L_\infty[0, 1]$ such that for all $i \neq j$,*

1.

$$\langle f_i, f_j \rangle = \langle z_i, z_j \rangle;$$

2.

$$\|f_i\|_\infty \leq R(K_{n^3});$$

3.

$$1 \leq \|f\|_2^2 \leq 1 + \frac{R(K_{n^3})^2}{n^2}.$$

Proof. For each vector z_i, consider n^2 its copies

$$z_i^k = z_i,$$

where $i \in \{1, \dots, n\}$ and $k \in \{1, \dots, n^2\}$. By the definition of $R(K_{n^3})$ there exist functions f_i^k such that for all $(i,k) \neq (j,l)$

$$\langle f_i^k, f_j^l \rangle = \langle z_i^k, z_j^l \rangle$$

and $\|f_i^k\|_\infty \leq R(K_{n^3})$. Define f_i to be the average of f_i^k:

$$f_i = \frac{f_i^1 + \cdots + f_i^{n^2}}{n^2}.$$

Verify that $f_1, \dots, f_n$ satisfy conditions 1–3.

1. For $i \neq j$,

$$\begin{aligned}
\langle f_i, f_j \rangle &= \left\langle \frac{f_i^1 + \cdots + f_i^{n^2}}{n^2}, \frac{f_j^1 + \cdots + f_j^{n^2}}{n^2} \right\rangle \\
&= \frac{\sum_{k,l=1}^{n^2} \langle f_i^k, f_j^l \rangle}{n^4} = \frac{\sum_{k,l=1}^{n^2} \langle z_i, z_j \rangle}{n^4} = \langle z_i, z_j \rangle.
\end{aligned}$$

2. We have
$$\|f_i\|_\infty \le \frac{\|f_i^1\|_\infty + \cdots + \|f_i^{n^2}\|_\infty}{n^2} \le R(K_{n^3}).$$

3. Notice that $\langle f_i^k, f_i^l\rangle = \langle z_i, z_i\rangle = 1$ for $k \neq l$, therefore
$$\|f_i\|_2^2 = \frac{1}{n^4}\sum_{k=1}^{n^2}\|f_i^k\|_2^2 + \frac{1}{n^4}\sum_{\substack{k,l=1\\k\neq l}}^{n^2}\langle f_i^k, f_i^l\rangle = \frac{1}{n^4}\sum_{k=1}^{n^2}\|f_i^k\|_2^2 + (1-\frac{1}{n^2}).$$

Now,
$$\begin{aligned}\frac{1}{n^4}\sum_{k=1}^{n^2}\|f_i^k\|_2^2 &= \frac{1}{n^4(n^2-1)}\sum_{\substack{k,l=1\\k\neq l}}^{n^2}\frac{\|f_i^k\|_2^2+\|f_i^l\|_2^2}{2} \ge \frac{1}{n^4(n^2-1)}\sum_{\substack{k,l=1\\k\neq l}}^{n^2}\langle f_i^k, f_i^l\rangle \\ &= \frac{1}{n^4(n^2-1)}\sum_{\substack{k,l=1\\k\neq l}}^{n^2}\langle f_i, f_i\rangle = \frac{1}{n^2}.\end{aligned}$$

On the other hand,
$$\frac{1}{n^4}\sum_{k=1}^{n^2}\|f_i^k\|_2^2 \le \frac{1}{n^4}\sum_{k=1}^{n^2}\|f_i^k\|_\infty^2 \le \frac{1}{n^4}\sum_{k=1}^{n^2}R(K_{n^3})^2 = \frac{R(K_{n^3})^2}{n^2}.$$

Hence
$$1 \le \|f_i\|_2^2 \le 1 + \frac{R(K_{n^3})^2}{n^2}.$$

$\square$

Theorem 2.26. *There exists an absolute positive constant C such that for every n,*

$$R(K_n) \geq C\sqrt{\log n}.$$

Proof. Fix a positive d and set $n = 3^d$. Our construction consists of $N = d + n$ vectors: vectors $e_1, \ldots, e_d$ that form an orthonormal basis[3] in the d-dimensional space, and vectors $z_1, \ldots, z_n$ that form a 1-net on the sphere S^{d-1}. Let $\psi(e_1), \ldots, \psi(e_d)$ and $\psi(z_1), \ldots, \psi(z_n)$ be functions that satisfy conditions 1–3 of Lemma 2.25. Denote $R = R(K_{N^3})$ and $\varepsilon = R^2/N^2$. Then for every $i \in \{1, \ldots, d\}$ and $j \in \{1, \ldots, n\}$, $1 \leq \|\psi(e_i)\|_2^2 \leq 1 + \varepsilon$ and $1 \leq \|\psi(z_j)\|_2^2 \leq 1 + \varepsilon$.

We extend linearly the mapping ψ from the basis $e_1, \ldots, e_d$ to the whole space $\mathbb{R}^d$ i.e. for every z in $\mathbb{R}^d$ define

$$\varphi(z) = \sum_{i=1}^{d} \langle z, e_i \rangle \psi(e_i).$$

Note that for a fixed z, $\varphi(z)$ is a function from $[0, 1]$ to $\mathbb{R}$. We show that $\varphi(z_m)$ is close to $\psi(z_m)$ for all m. For all unit vectors z, we have

$$\begin{aligned}
\|\varphi(z)\|_2^2 &= \left\langle \sum_{i=1}^{d} \langle z, e_i \rangle \psi(e_i), \sum_{i=1}^{d} \langle z, e_i \rangle \psi(e_i) \right\rangle \\
&= \sum_{i=1}^{d} (\langle z, e_i \rangle)^2 \|\psi(e_i)\|_2^2 \\
&\leq \max_i \|\psi(e_i)\|_2^2 \leq 1 + \varepsilon.
\end{aligned}$$

[3] Note that we added the vectors $e_1, \ldots, e_d$ to our construction for simplicity of presentation. In fact, it suffices to consider only vectors $z_1, \ldots, z_n$. See Alon, Makarychev, Makarychev and Naor [6] for details.

Then for every z_m,

$$\begin{aligned}\|\varphi(z_m)-\psi(z_m)\|_2^2 &= \|\varphi(z_m)\|_2^2+\|\psi(z_m)\|_2^2-2\langle\varphi(z_m),\psi(z_m)\rangle \\ &\le 2+2\varepsilon-2\left\langle\sum_{i=1}^{d}\langle z_m,e_i\rangle\psi(e_i),\psi(z_m)\right\rangle \\ &= 2+2\varepsilon-2\sum_{i=1}^{d}(\langle z_m,e_i\rangle)^2=2+2\varepsilon-2\|z_m\|_2^2 \\ &= 2\varepsilon.\end{aligned}$$

Consider the function $\Phi:[0,1]\to\mathbb{R}^d$ defined as follows:

$$\Phi(t)=\sum_{i=1}^{d}\psi(e_i)(t)e_i.$$

Our goal is to estimate

$$\|\Phi\|_2^2\equiv\int_0^1\|\Phi(t)\|_2^2dt.$$

Notice that the linear operator φ can now be expressed as

$$\varphi(z)(t)=\langle\Phi(t),z\rangle.$$

Hence for every z_m,

$$\int_0^1|\langle\Phi(t),z_m\rangle-\psi(z_m)(t)|^2dt=\|\varphi(z_m)-\psi(z_m)\|_2^2\le 2\varepsilon.$$

Since $\|\psi(z_m)\|_\infty\le R$, we have

$$\mu\left\{t:|\langle\Phi(t),z_m\rangle|>R+1\right\}\le\mu\left\{t:|\langle\Phi(t),z_m\rangle-\psi(z_m)(t)|\ge 1\right\}\le 2\varepsilon.$$

Now by Lemma 2.24,

$$\mu\{t : \|\Phi(t)\|_2 > 2(R+1)\} \leq \sum_{m=1}^{n} \mu\{t : \langle \Phi(t), z_m \rangle > R+1\} \leq 2\varepsilon n.$$

The length of $\Phi(t)$ is less than $(R+1)$ for t in a set of measure $1 - 2\varepsilon n$. Moreover, since the coordinates of $\Phi(t)$ are less than R in absolute value, the length of $\Phi(t)$ is bounded by $R\sqrt{d}$ for the rest of t. Therefore,

$$\|\Phi\|_2^2 \equiv \int_0^1 \|\Phi(t)\|_2^2 dt \leq 1 \times (R+1)^2 + 2\varepsilon n \times dR^2 \leq (R+1)^2 + \frac{2dR^4}{N}.$$

On the other hand,

$$\|\Phi\|_2^2 \equiv \int_0^1 \|\Phi(t)\|_2^2 dt = \sum_{i=1}^{d} \int_0^1 \|\psi(e_i)(t)\|_2^2 dt = \sum_{i=1}^{d} \|\psi(e_i)\|_2^2 \geq d.$$

We get

$$(R+1)^2 + \frac{2dR^4}{3^d} \geq d.$$

Which implies

$$R \equiv R(K_{(3^d+d)^3}) \geq \Omega(\sqrt{d}).$$

Since $R(K_n)$ is a non-decreasing function of n, we get that $R(K_n) \geq \Omega(\sqrt{\log n})$ for every n. $\square$

Corollary 2.27. *For every integer n,*

$$K(K_n) \geq C \log n,$$

where $C > 0$ is a universal constant.

Proof. By Theorem 2.26,

$$R(K_n) \geq C_1 \sqrt{\log n}.$$

Then by Theorem 2.8,

$$K(K_n) = R(K_n)^2 \geq C_1^2 \log n.$$

$\square$

The corollary immediately implies Theorem 2.4, since $K(G)$ is a monotone function with respect to taking subgraphs.

Theorem 2.4 *There exists an absolute positive constant C such that for every graph G,*

$$K(G) \geq C \log \omega(G),$$

where $\omega(G)$ is the clique number of the graph G.

2.7 Homogenous Grothendieck Inequality

The classical Grothendieck inequality has the following equivalent homogenous formulation: For every $n \times m$ matrix (b_{ij}),

$$\sup_{x_i, y_j \in \ell_2} \frac{\sum_{i=1}^n \sum_{j=1}^m b_{ij} \langle x_i, y_j \rangle}{\max_{i,j}(\|x_i\|_2 \cdot \|y_j\|_2)} \leq K_G \cdot \sup_{x_i, y_j \in \mathbb{R}} \frac{\sum_{i=1}^n \sum_{j=1}^m b_{ij} x_i y_j}{\max_{i,j}(|x_i| \cdot |y_j|)},$$

where K_G is Grothendieck's constant. There is a natural homogenous extension of Grothendieck's inequality to the case of arbitrary graphs.

Theorem 2.28. *For every loop-free graph $G = (V, E)$ and matrix A in $\mathcal{A}_G$,*

$$\max_{z_i \in \mathbb{R}^n} \frac{\sum_{i,j} a_{ij} \cdot \langle z_i, z_j \rangle}{\max_{(i,j)\in E}(\|z_i\|_2 \cdot \|z_j\|_2)} \leq K(G) \cdot \max_{x_i \in \mathbb{R}} \frac{\sum_{i,j} a_{ij} \cdot x_i x_j}{\max_{(i,j)\in E}(|x_i| \cdot |x_j|)}.$$

Proof. We shall prove a stronger statement: *For every set of vectors $z_1, \ldots, z_n$ there exist numbers $x_1, \ldots, x_n$ such that $|x_i| = \|z_i\|_2$ for all i and*

$$\sum_{i,j} a_{ij} \langle z_i, z_j \rangle \leq K(G) \sum_{i,j} a_{ij} \cdot x_i x_j.$$

Particularly, $|x_i| \cdot |x_j| = \|z_i\|_2 \cdot \|z_j\|_2$ for every $(i, j) \in E$, which implies the inequality we want to prove.

Consider the matrix

$$\hat{a}_{ij} = \|z_i\|_2 \cdot \|z_j\|_2 \times a_{ij}$$

and let $\hat{z}_i = z_i / \|z_i\|_2$ (if $z_i = 0$, then $\hat{z}_i$ is an arbitrary unit vector.). Write

$$\sum_{i,j} a_{ij} \langle z_i, z_j \rangle = \sum_{i,j} a_{ij} \cdot \|z_i\|_2 \cdot \|z_i\|_2 \times \left\langle \frac{z_i}{\|z_i\|_2}, \frac{z_j}{\|z_j\|_2} \right\rangle = \sum_{i,j} \hat{a}_{ij} \times \langle \hat{z}_i, \hat{z}_j \rangle.$$

Since $\hat{z}_i$ are unit vectors, there exist numbers $\hat{x}_1, \ldots, \hat{x}_n \in \{\pm 1\}$ such that

$$\sum_{i,j} \hat{a}_{ij} \cdot \langle \hat{z}_i, \hat{z}_j \rangle \leq K(G) \sum_{i,j} \hat{a}_{ij} \cdot \hat{x}_i \hat{x}_j = K(G) \sum_{i,j} a_{ij} \cdot (\|z_i\|_2 \hat{x}_i) \cdot (\|z_j\|_2 \hat{x}_j).$$

Thus we proved that the sequence $x_i = \|z_i\|_2 \cdot \hat{x}_i$ satisfies the required conditions. $\square$

2.8 Restricted Families of Graphs

By Theorems 2.3 and 2.4, if G is a graph containing a clique of size whose logarithm is proportional to the logarithm of the theta function of $\overline{G}$, then $K(G) = \Theta(\log \omega(G))$. In particular, this holds for any perfect graph G, as in this case $\omega(G) = \chi(G) = \vartheta(\overline{G})$. Note that it may be desirable to optimize quadratic forms as in (2.1) for various classes of perfect graphs, like comparability graphs or chordal graphs. Since the chromatic number of perfect graphs can be determined in polynomial time, it follows that in such cases we can determine the value of $K(G)$ up to a constant factor, and in case it is smaller than n^{ε} where n is the number of vertices, we can obtain a guaranteed improved approximation of (2.1).

Moreover, by the remark above and the fact that $\vartheta(\overline{G}) \leq \chi(G)$ for every G, $K(G) = \Theta(\log(\chi(G)))$ for any graph G whose chromatic number is bounded by a polynomial of the size of its largest clique. There are various known classes of such graphs, including complements of intersection graphs of the edges of hypergraphs with certain properties. See, for example, Alon [3] and Ding, Seymour and Winkler [19]. Another interesting family of classes of graphs for which this holds is obtained as follows. Let $k \geq 2$ be an integer, and let $\mathcal{G}_k$ denote the family of all graphs which contain no induced star on k vertices. In particular, $\mathcal{G}_2$ is the family of all unions of pairwise vertex disjoint cliques, and $\mathcal{G}_3$ is the family of all claw-free graphs. An easy application of Ramsey Theorem implies that the maximum degree Δ of any graph $G \in \mathcal{G}_k$ whose largest clique is of size ω, is at most ω^{k-1}. This implies that the chromatic number χ of any such graph is at most $\omega^{k-1}+1$, showing that (for fixed k) its Grothendieck constant is $\Theta(\log \Delta) = \Theta(\log \omega) = \Theta(\log \chi)$. Since the maximum degree (as well as the fact that $G \in \mathcal{G}_k$ for some fixed k) can be computed efficiently, this is another case in which we can compute the value of $K(G)$ up to a constant factor.

A graph G is d-degenerate if any subgraph of it contains a vertex of degree at most d. It is easy and well known that any such graph is $(d+1)$-colorable, and that there is a linear time algorithm that finds, for a given graph G, the smallest number d such that G is d-degenerate. In particular, graphs of genus g are $O(\sqrt{g})$-degenerate, implying that their Grothendieck constant is $O(\log g)$.

Other classes of graphs for which the clique number is proportional (with a universal constant) to the chromatic number (and hence also to the theta function of the complement), are intersection graphs of a family of homothetic copies of a fixed convex set in the plane (see Kim, Kostochka and Nakprasit [40]). A few additional examples appear in the next subsection.

2.9 New Grothendieck-type Inequalities

Theorem 2.3 enables us to generate new Grothendieck type Inequalities, and Theorem 2.4 can be used to show that in some cases these are essentially tight. We list here several examples that seem interesting.

Let $m > 2k > 1$ be integers. The *Kneser graph* $S(m,k)$ is the graph whose $n = \binom{m}{k}$ vertices are all k-subsets of an m-element set, where two vertices are adjacent if the corresponding subsets are disjoint. The clique number of $S(m,k)$ is clearly $\lfloor m/k \rfloor$, and as shown by Lovász [44], its chromatic number is $m - 2k + 2$. Its theta function, computed by Lovász [45], is $\binom{m-1}{k-1}$, and as this graph is vertex transitive, the product of its theta function and that of its complement is the number of its vertices (see [45]). It thus follows that the theta-function of the complement of $S(m,k)$ is m/k. We conclude that the Grothendieck constant of $S(m,k)$ is $\Theta(\log(m/k))$. This gives the following Grothendieck-type inequality.

Proposition 2.29. *There exists an absolute constant C such that the following holds. Let $m > 2k > 1$ be positive integers. Put $M = \{1,2,\ldots,m\}$, $n = \binom{m}{k}$, and let $A(I,J)$ be a real number for each pair of disjoint k-subsets I, J of M. Then for every vectors z_I of lenght 1, there are signs $x_I \in \{-1,1\}$ such that*

$$\sum_{\substack{I,J\subset M\\|I|=|J|=k\\I\cap J=\emptyset}} A(I,J)\langle z_I, z_J\rangle \le C\log\left(\frac{m}{k}\right) \sum_{\substack{I,J\subset M\\|I|=|J|=k\\I\cap J=\emptyset}} A(I,J)x_Ix_J.$$

Moreover, the above inequality is tight, up to the constant factor C, for all admissible values of m and k.

Let D_m denote the line graph of the directed complete graph on m vertices. This is the graph whose vertices are all ordered pairs (i,j) with $i,j \in M = \{1,2,\ldots,m\}$, $i \neq j$, in which (i,j) is adjacent to (j,k) for all for all admissible $i,j,k \in M$. It is known that the chromatic number of this graph is $(1+o(1))\log_2 m$, and it is not difficult to see that its clique number is 2 and its fractional chromatic number is at most 4, (see, for example, [8]). As shown in Lovász [45], the theta function of the complement of any graph is bounded by its fractional chromatic number. This gives the following inequality.

Proposition 2.30. *Let m be a natural number and let $A(i,j,k)$ be a sequence of real numbers, where $1 \le i,j,k \le m$, $i \neq j, j \neq k$. Then for all vectors $z_{ij}, (1 \le$*

$i, j \leq m, i \neq j)$ *of length* 1*, there exist numbers* $x_{ij} \in \{-1, 1\}$ *such that*

$$\sum_{\substack{i,j,k \in M \\ i \neq j \neq k}} A(i,j,k)\langle z_{ij}, z_{jk}\rangle \leq C \sum_{\substack{i,j,k \in M \\ i \neq j \neq k}} A(i,j,k) x_{ij} x_{jk},$$

where C *is an absolute constant.*

Let $n = 2^m$, and let M be as before. Let G be the comparability graph of all subsets of M, ordered by inclusion. Then G is a perfect graph, and its clique number is $m + 1 = \log n + 1$. As G is perfect, this is also its chromatic number (and the theta function of its complement), providing the following inequality.

Proposition 2.31. *Let* m *be a natural number, let* $n = 2^m$*, and let* $A(I, J)$ *be a real number for each* $I \subsetneq J \subseteq M$*. Then for every* n *vectors* $z_I, (I \subseteq M)$ *of length* 1*, there are numbers* $x_I \in \{-1, 1\}$ *such that*

$$\sum_{I \subsetneq J \subseteq M} A(I,J)\langle z_I, z_J\rangle \leq C \log\log n \sum_{I \subsetneq J \subseteq M} A(I,J) x_I x_J,$$

where C *is an absolute positive constant. This is tight, up to the constant* C*.*

Chapter 3

Approximation Algorithm for MAX k-CSP

3.1 Introduction

In this chapter we study the maximum constraint satisfaction problem with k variables in each constraint (MAX k-CSP): *Given a set of boolean variables and constraints, where each constraint depends on at most k variables*[1]*, our goal is to find an assignment so as to maximize the number of satisfied constraints.*

The approximation factor for MAX k-CSP is of interest in complexity theory since it is closely tied to the relationship between the completeness and soundness of k-bit Probabilistically Checkable Proofs (PCPs). Let us give an informal definition of a PCP[2]. A language L has a k-query PCP system with soundness α and completeness β, if given a string y and a proof/witness $b_1 b_2 \dots b_m$ the verifier can

[1] i.e. each constraint is an arbitrary predicate of k variables

[2] For a formal definition and more details on PCPs, we refer the reader to the book of Arora and Barak [10].

decide (in probabilistic polynomial time in $|y|$) whether y belongs to L by querying (non-adaptively) only k bits from $b_1 b_2 \dots b_m$ and using only $O(\log |y|)$ random bits; we require that

- If $y \in L$, then there exists a proof $b_1 b_2 \dots b_m \in \{0,1\}^m$ that the verifier accepts with probability at least β.
- If $y \notin L$, then the verifier rejects every proof $b_1 b_2 \dots b_m \in \{0,1\}^m$ with probability at least $1 - \alpha$.

Suppose that a language L has a k-query PCP system. Then given a string y we can in polynomial time construct a k-CSP instance such that if $y \in L$, then the number of satisfied constraints in the optimal solution is at least β; if $y \notin L$, then the number of satisfied constraints in the optimal solution is at most α. The variables in this CSP are bits $b_1, \dots, b_m$. For each possible query of the verifier we have a constraint on the variables being queried[3]: The constraint is true, if the verifier accepts the proof; and false otherwise (the decision of the verifier depends only on k bits from $b_1 \dots b_m$, hence each constraint in the CSP depends only on k variables). Notice that the probability that the verifier accepts a proof $\tilde{b}_1 \dots \tilde{b}_m$ is exactly equal to the fraction of satisfied constraints in the solution $b_1 = \tilde{b}_1, \dots, b_m = \tilde{b}_m$. Therefore, if there exists a k-query PCP system with soundness α and completeness β for an $\mathcal{NP}$ - complete language (say, SAT), then it is $\mathcal{NP}$-hard to distinguish between k-CSP instances where at most α fraction of all constraints is satisfiable and instances where at least β fraction of all constraints is satisfiable. This implies that the ratio β/α is not greater than the approximation factor one can achieve for the MAX k-CSP problem.

[3]Note that we can efficiently enumerate all queries, since the verifier uses only $O(\log |y|)$ random bits.

A trivial algorithm for k-CSP is to pick a random assignment. It satisfies each constraint with probability at least $1/2^k$ (except those constraints which cannot be satisfied). Therefore, its approximation ratio is $1/2^k$. Trevisan [59] improved on this slightly by giving an algorithm with approximation ratio $2/2^k$. Until recently, this was the best approximation ratio for the problem. Recently, Hast [27] proposed an algorithm with an asymptotically better approximation guarantee $\Omega\left(k/(2^k \log k)\right)$. Also, Samorodnitsky and Trevisan [56] proved that it is hard to approximate MAX k-CSP within $2k/2^k$ for every $k \geq 3$, and within $(k+1)/2^k$ for infinitely many k assuming the Unique Games Conjecture of Khot [37]. We close the gap between the upper and lower bounds for k-CSP by giving an algorithm with approximation ratio $\Omega\left(k/2^k\right)$. By the results of [56], our algorithm is asymptotically optimal within a factor of approximately $1/0.44 \approx 2.27$ (assuming the Unique Games Conjecture).

In our algorithm, we use the approach of Hast [27]: we first obtain a "preliminary" solution $x_1, \ldots, x_n \in \{-1, 1\}$ and then independently flip the values of x_i using a slightly biased distribution (i.e. we keep the old value of x_i with probability slightly larger than $1/2$). In this paper, we improve and simplify the first step in this scheme. Namely, we present a new method of finding $x_1, \ldots, x_n$, based on solving a certain semidefinite program (SDP) and then rounding the solution to ± 1 using the result of Rietz [55] and Nesterov [50]. Note, that Hast obtains $x_1, \ldots, x_n$ by maximizing a quadratic form (which differs from our SDP) over the domain $\{-1, 1\}$ using the algorithm of Charikar and Wirth [18]. The second step of our algorithm is essentially the same as in Hast's algorithm.

Our result is also applicable to MAX k-CSP with a larger domain[4]: it gives a $\Omega\left(k \log d/d^k\right)$ approximation for instances with domain size d.

[4]To apply the result to an instance with a larger domain, we just encode each domain value with $\log d$ bits.

Let us point out that the case of $k = 2$ is very different from $k \geq 3$. We gave an approximation algorithm for this case in Charikar, Makarychev, Makarychev [16], but we do not include it in this dissertation.

3.2 Reduction to Max k-AllEqual

We use Hast's reduction of the MAX k-CSP problem to the MAX k-ALLEQUAL problem.

Definition 3.1 (Max k-AllEqual Problem). *Given a set S of clauses of the form $l_1 \equiv l_2 \equiv \cdots \equiv l_k$, where each literal l_i is either a boolean variable x_j or its negation $\bar{x}_j$. The goal is to find an assignment to the variables x_i so as to maximize the number of satisfied clauses.*

The reduction works as follows. First, we write each constraint $f(x_{i_1}, x_{i_2}, \ldots, x_{i_k})$ as a CNF formula. Then we consider each clause in the CNF formula as a separate constraint; we get an instance of the MAX k-CSP problem, where each clause is a conjunction. The new problem is equivalent to the original problem: each assignment satisfies exactly the same number of clauses in the new problem as in the original problem. Finally, we replace each conjunction $l_1 \wedge l_2 \wedge \ldots \wedge l_k$ with the constraint $l_1 \equiv l_2 \equiv \cdots \equiv l_k$. Clearly, the value of this instance of Max k-AllEqual is at least the value of the original problem. Moreover, it is at most two times greater then the value of the original problem: if an assignment $\{x_i\}$ satisfies a constraint in the new problem, then either the assignment $\{x_i\}$ or the assignment $\{\bar{x}_i\}$ satisfies the corresponding constraint in the original problem. Therefore, a ρ approximation guarantee for Max k-AllEqual translates to a $\rho/2$ approximation guarantee for the MAX k-CSP.

Note that this reduction may increase the number of constraints by a factor of $O(2^k)$. However, our approximation algorithm gives a nontrivial approximation only when $k/2^k \geq 1/m$ where m is the number of constraints, that is, when $2^k \leq O(m \log m)$ is polynomial in m.

Below we consider only the Max k-AllEqual problem.

3.3 SDP Relaxation

For brevity, we denote $\bar{x}_i$ by x_{-i}. We think of each clause C as a set of indices: the clause C defines the constraint "(for all $i \in C$, x_i is true) or (for all $i \in C$, x_i is false)". Without loss of generality we assume that there are no unsatisfiable clauses in S, i.e. there are no clauses that have both literals x_i and $\bar{x}_i$.

We consider the following SDP relaxation of the Max k-AllEqual problem:

$$\text{maximize } \frac{1}{k^2} \sum_{C \in S} \left\| \sum_{i \in C} z_i \right\|^2$$

subject to

$$\begin{aligned} \|z_i\|^2 &= 1 && \text{for all } i \in \{\pm 1, \ldots, \pm n\} \\ z_i &= -z_{-i} && \text{for all } i \in \{\pm 1, \ldots, \pm n\} \end{aligned}$$

This is indeed a relaxation of the problem: in the intended solution $z_i = z_0$ if x_i is true, and $z_i = -z_0$ if x_i is false (where z_0 is a fixed unit vector). Then each satisfied clause contributes 1 to the SDP value. Hence the value of the SDP is greater than or equal to the value of the Max k-AllEqual problem. We use the following theorem of Rietz [55] and Nesterov [50].

Theorem 3.2 (Rietz [55], Nesterov [50])**.** *There exists an efficient algorithm that given a positive semidefinite matrix $A = (a_{ij})$, and a set of unit vectors z_i, assigns ± 1 to variables x_i, such that*

$$\sum_{i,j} a_{ij}\, x_i x_j \geq \frac{2}{\pi} \sum_{i,j} a_{ij}\, \langle z_i, z_j \rangle. \tag{3.1}$$

Remark 3.3. *Rietz proved that for every positive semidefinite matrix A and unit vectors z_i there exist $x_i \in \{\pm 1\}$ such that inequality (3.1) holds. Nesterov presented a polynomial time algorithm that finds such values of x_i.*

Observe that the quadratic form

$$\frac{1}{k^2} \sum_{C \in S} \left(\sum_{i \in C} x_i \right)^2$$

is positive semidefinite. Therefore we can use the algorithm from Theorem 3.2. Given vectors z_i as in the SDP relaxation, it yields numbers x_i such that

$$\frac{1}{k^2} \sum_{C \in S} \left(\sum_{i \in C} x_i \right)^2 \geq \frac{2}{\pi} \frac{1}{k^2} \sum_{C \in S} \left\| \sum_{i \in C} z_i \right\|^2$$

$$x_i \in \{\pm 1\}$$

$$x_i = -x_{-i}$$

(Formally, z_{-i} is a shortcut for $-z_i$; x_{-i} is a shortcut for $-x_i$).

In what follows, we assume that $k \geq 3$ — for $k = 2$ we can use the MAX CUT

1. Solve the semidefinite relaxation for Max k-AllEqual. Get vectors z_i.
2. Apply Theorem 3.2 to vectors z_i as described above. Get values x_i.
3. Let $\delta = \sqrt{\frac{2}{k}}$.
4. For each $i \geq 1$ assign (independently)

$$b_i = \begin{cases} true, & \text{with probability } \frac{1+\delta x_i}{2}; \\ false, & \text{with probability } \frac{1-\delta x_i}{2}. \end{cases}$$

Figure 3.1: Approximation Algorithm for Max k-AllEqual

algorithm by Goemans and Williamson [23] to get a better approximation[5].

The approximation algorithm is shown in Figure 3.1.

3.4 Analysis

Theorem 3.4. *The approximation algorithm finds an assignment satisfying at least $ck/2^k \cdot OPT$ clauses (where $c > 0.88$ is an absolute constant), given that OPT clauses are satisfied in the optimal solution.*

Proof. Denote $Z_C = \frac{1}{k}\sum_{i\in C} x_i$. Then Theorem 3.2 guarantees that

$$\sum_{C\in S} Z_C^2 = \frac{1}{k^2}\sum_{C\in S}\left(\sum_{i\in C} x_i\right)^2 \geq \frac{2}{\pi}\frac{1}{k^2}\sum_{C\in S}\left\|\sum_{i\in C} z_i\right\|^2 = \frac{2}{\pi} SDP \geq \frac{2}{\pi} OPT,$$

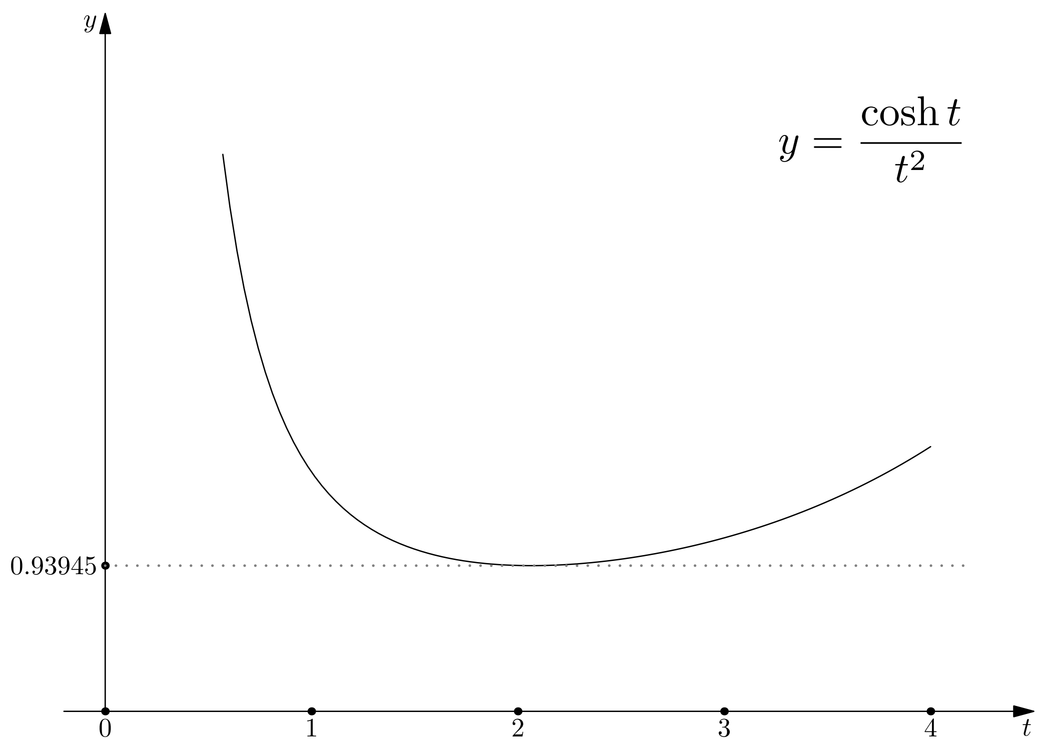
where SDP is the SDP value.

[5]Our algorithm works for $k = 2$ with a slight modification: δ should be less than 1.

Note that the number of x_i equal to 1 is $\frac{1+Z_C}{2}k$, the number of x_i equal to -1 is $\frac{1-Z_C}{2}k$. The probability that a constraint C is satisfied equals

$$\begin{aligned}
\Pr(C \text{ is satisfied}) &= \Pr\left(\forall i \in C\ b_i = 1\right) + \Pr\left(\forall i \in C\ b_i = -1\right) \\
&= \prod_{i\in C} \frac{1+\delta x_i}{2} + \prod_{i\in C} \frac{1-\delta x_i}{2} \\
&= \frac{1}{2^k}\big((1+\delta)^{(1+Z_C)k/2} \cdot (1-\delta)^{(1-Z_C)k/2} \\
&\quad + (1-\delta)^{(1+Z_C)k/2} \cdot (1+\delta)^{(1-Z_C)k/2}\big) \\
&= \frac{(1-\delta^2)^{k/2}}{2^k} \left(\left(\frac{1+\delta}{1-\delta}\right)^{Z_C k/2} + \left(\frac{1-\delta}{1+\delta}\right)^{Z_C k/2} \right) \\
&= \frac{1}{2^k}\,(1-\delta^2)^{k/2} \cdot 2\cosh\left(\frac{1}{2}\ln\frac{1+\delta}{1-\delta} \cdot Z_C k\right).
\end{aligned}$$

Here, $\cosh t \equiv (e^t + e^{-t})/2$. Let α be the minimum of the function $\cosh t/t^2$. Numerical computations show that $\alpha > 0.93945$.



We have,

$$\cosh\left(\frac{1}{2}\ln\frac{1+\delta}{1-\delta}\cdot Z_C k\right) \geq \alpha\left(\frac{1}{2}\ln\frac{1+\delta}{1-\delta}\cdot Z_C k\right)^2 \geq \ \alpha\left(\delta\cdot Z_C k\right)^2 = 2\alpha Z_C^2 k.$$

Recall that $\delta = \sqrt{2/k}$ and $k \geq 3$. Hence

$$(1-\delta^2)^{k/2} = \left(1-\frac{2}{k}\right)^{k/2} \geq \left(1-\frac{2}{k}\right)\cdot\frac{1}{e}.$$

Combining these bounds we get,

$$\Pr\left(C \text{ is satisfied}\right) \geq \frac{4\alpha}{e}\cdot\frac{k}{2^k}\cdot\left(1-\frac{2}{k}\right)\cdot Z_C^2.$$

However, a more careful analysis shows that the factor $1-2/k$ is not necessary, and the following bound holds (we give a proof in the next section):

$$2\alpha(1-\delta^2)^{k/2}\left(\frac{1}{2}\ln\frac{1+\delta}{1-\delta}\cdot Z_C k\right)^2 \geq \frac{4\alpha}{e}Z_C^2 k. \tag{3.2}$$

Therefore,

$$\Pr\left(C \text{ is satisfied}\right) \geq \frac{4\alpha}{e}\cdot\frac{k}{2^k}\cdot Z_C^2.$$

So the expected number of satisfied clauses is

$$\sum_{C\in S}\Pr\left(C \text{ is satisfied}\right) \ \geq \ \frac{4\alpha}{e}\frac{k}{2^k}\sum_{C\in S} Z_C^2 \geq \frac{4\alpha}{e}\frac{k}{2^k}\cdot\frac{2}{\pi}OPT.$$

We conclude that the algorithm finds an

$$\frac{8\alpha}{\pi e}\frac{k}{2^k} > 0.88\frac{k}{2^k}$$

approximation with high probability. $\square$

3.5 Proof of Inequality (3.2)

In this section, we will prove inequality (3.2):

$$2\alpha(1-\delta^2)^{k/2}\left(\frac{1}{2}\,\ln\frac{1+\delta}{1-\delta}\cdot Z_C k\right)^2 \geq \frac{4\alpha}{e}Z_C^2 k. \tag{3.2}$$

Proof. Let us first simplify this expression

$$(1-\delta^2)^{k/2}\left(\sqrt{\frac{k}{2}}\cdot\frac{\ln(1+\delta)-\ln(1-\delta)}{2}\right)^2 \geq e^{-1}.$$

Note that this inequality holds for $3 \leq k \leq 7$, which can be verified by direct computation. So assume that $k \geq 8$. Denote $t = 2/k$; and replace k with $2/t$ and δ with $\sqrt{t}$. We get

$$(1-t)^{1/t}\left(\frac{1}{\sqrt{t}}\cdot\frac{\ln(1+\sqrt{t})-\ln(1-\sqrt{t})}{2}\right)^2 \geq e^{-1}.$$

Take the logarithm of both sides:

$$\frac{\ln(1-t)}{t} + 2\ln\left(\frac{1}{\sqrt{t}}\cdot\frac{\ln(1+\sqrt{t})-\ln(1-\sqrt{t})}{2}\right) \geq -1.$$

Observe that

$$\frac{1}{\sqrt{t}}\cdot\frac{\ln(1+\sqrt{t})-\ln(1-\sqrt{t})}{2} = 1+\frac{t}{3}+\frac{t^2}{5}+\frac{t^3}{7}+\cdots \geq 1+\frac{t}{3};$$

and

$$
\begin{aligned}
\frac{\ln(1-t)}{t} &= -1 - \frac{t}{2} - \frac{t^2}{3} - \cdots \geq -1 - \frac{t}{2} - \frac{t^2}{3} \times \sum_{i=0}^{\infty} t^i \\
&\geq -1 - \frac{t}{2} - \frac{4t^2}{9}.
\end{aligned}
$$

In the last inequality we used our assumption that $t \equiv 2/k \leq 1/4$. Now,

$$
\begin{aligned}
\frac{\ln(1-t)}{t} + 2\ln\left(\frac{1}{\sqrt{t}} \cdot \frac{\ln(1+\sqrt{t}) - \ln(1-\sqrt{t})}{2}\right) &\geq \left(-1 - \frac{t}{2} - \frac{4t^2}{9}\right) + 2\ln\left(1 + \frac{t}{3}\right) \\
&\geq \left(-1 - \frac{t}{2} - \frac{4t^2}{9}\right) + 2\left(\frac{t}{3} - \frac{t^2}{18}\right) \\
&\geq -1 + \frac{t}{6} - \frac{5t^2}{9} \geq -1.
\end{aligned}
$$

Here $(t/6 - 5t^2/9)$ is positive, since $t \in (0, 1/4]$. This concludes the proof. $\square$

Chapter 4

Advantage over Random for Maximum Acyclic Subgraph

4.1 Introduction

The focus of this chapter is the MAX ACYCLIC SUBGRAPH problem which is the following:

Definition 4.1. *Given a directed graph $G = (V, E)$, find the largest subset of edges which are acyclic. Equivalently, find an ordering of the vertices so as to maximize the number of edges going forward.*

A simple randomized algorithm achieves a factor 1/2 for this problem: Simply pick a random ordering of the vertices. In fact, one can achieve factor 1/2 by an even simpler algorithm: Pick an arbitrary ordering of the vertices π and its reverse π^R. One of them has at least 1/2 fraction of the edges in the forward direction. Improving the 1/2 approximation for MAX ACYCLIC SUBGRAPH is a long standing open problem. The motivating question for our work was whether it is possible to beat this 1/2 approximation.

In fact, algorithms very slightly better than 1/2 are known: Berger and Shor [14] showed how to get $1/2 + \Omega(1/\sqrt{d_{max}})$, where d_{max} is the maximum vertex degree in the graph. Later Hassin and Rubinstein [26] proposed another algorithm with the same approximation guarantee, but better running time in certain cases. Note that these algorithms achieve such a guarantee for any graph: the guarantee does not depend on the value of the optimal solution.

Let us measure the objective as a fraction of the number of edges. If the optimal value OPT is $1 - \varepsilon$, for a very small ε, we could use the best known approximation algorithm for the complementary problem (Min Feedback Arc Set) to beat the random algorithm. Using the $O(\log n \log\log n)$ algorithm of Seymour [57], for instances where $OPT = 1 - \varepsilon$ and $\varepsilon = O(1/(\log n \log\log n))$, we can indeed beat 1/2 for Max Acyclic Subgraph. This yields an approximation ratio of $1/2 + \Omega(1/(\log n \log\log n))$ for the problem.

To summarize, we can beat random for instances where OPT is very close to 1. For instances where OPT is smaller, we do not know of any techniques which perform better than random. (As mentioned, there are algorithms which have the guarantee $1/2 + \Omega(1/\sqrt{d_{max}})$.)

Recently, a related question of approximating the advantage over random has been studied for several basic optimization problems (see [9, 16, 18, 20, 27, 28, 30, 31]). These studies give a fresh perspective on these optimization problems and motivated the development of new techniques to extract information from mathematical programming relaxations for them.

Definition 4.2. *Let $G = (V, E)$ be a directed graph on n vertices; and let π :*

$\{1, \ldots, n\} \to V$ *be a linear arrangement of its vertices. Then the* advantage or gain[1] over random *of the arrangement* π *is equal to the fraction of edges going forward minus the fraction of edges going backward. We denote the gain over random by* $\text{gain}(G, \pi)$.

If a linear arrangement has value $1/2+\delta$ for Max Acyclic Subgraph, then the gain of this arrangement is 2δ. The question of beating random for Max Acyclic Subgraph can be phrased thus: Given an instance with optimal gain δ, can we guarantee that we produce a solution with gain $f(\delta)$?

Note that the usual notion of approximation only focuses on instances where the optimal gain δ is close to 1 (undoubtedly a very interesting question). We ask what guarantee is possible as a function of δ for all values of $\delta \in (0, 1)$.

Such guarantees were developed for MAX CUT by Charikar and Wirth [18]: Given a MAX CUT instance, for which the optimal solution has gain δ, we can find a cut with gain $\Omega(\delta/\log(1/\delta))$ and $\Omega(\delta/\log n)$; the former approximation guarantee is optimal (if the Unique Games Conjecture is true) as was shown by Khot and O'Donnell [39].

There are some parallels between MAX CUT and Max Acyclic Subgraph, since a random assignment achieves factor 1/2 for both problems. For MAX CUT, this was indeed the best known until the seminal work of Goemans and Williamson [23] using semidefinite programming (SDP). In a sequence of later papers, our understanding of the MAX CUT SDP has vastly improved. Arguably, Max Acyclic Subgraph is a more complex problem than MAX CUT. Linear programming (LP) relaxations for the problem have been studied intensively in the

[1] This concept has been referred to in the literature as both *advantage over random* and *gain*. We will use both terms interchangeably.

mathematical programming community (it is sometimes referred to as the linear ordering problem). For more information we refer the reader to the papers of Newman [51] and Newman and Vempala [53]. However the best known approximation for the problem still remains 1/2. Newman [52] recently studied an SDP relaxation for the problem and gave some evidence to suggest that the SDP might be useful in beating the 1/2 approximation (in particular, the SDP does well on the known gap instances for the LP).

In this work, we give an $O(\log n)$ approximation for the advantage over random for Max Acyclic Subgraph. In other words, given an instance where $OPT = 1/2 + \delta$, we find a solution of value $1/2 + \Omega(\delta/\log n)$. Prior to our work, no non-trivial guarantees were known even for instances where OPT was close to 1, say $1 - 1/\log n$. In contrast, our algorithm gives a non-trivial guarantee even for OPT close to 1/2. As a byproduct, we obtain a $1/2 + \Omega(1/\log n)$ approximation for Max Acyclic Subgraph — very slightly better than the $1/2 + \Omega(1/(\log n \log\log n))$ alluded to earlier that comes from Seymour's algorithm [57]. Note that the known hardness results for Max Acyclic Subgraph [34, 54] imply that the advantage over random version has a constant factor hardness.

Vertex ordering problems like Max Acyclic Subgraph seem more complex than the constraint satisfaction problems that have been recently explored with the lens of approximating the advantage over random. It is somewhat surprising therefore that we obtain results for Max Acyclic Subgraph that match the corresponding guarantee for MAX CUT. Despite the similarity in the statement of the result, the techniques are quite different. The $\log n$ in MAX CUT comes from the tail of the Gaussian distribution, while the $\log n$ in our result comes from the number of different distance scales in a linear arrangement. Roughly speaking, our results show how ordering information from one distance scale in the optimum solution can be exploited algorithmically. Extending these ideas further to exploiting information from multiple distance scales simultaneously is a promising avenue for obtaining a constant better than 1/2 approximation for Max Acyclic Subgraph. This would be an exciting result indeed.

Our main result is as follows.

Theorem 4.3. *There exists a randomized polynomial time algorithm that given a directed graph G finds a linear arrangement π of its vertices with gain over random at least $\Omega(\delta/\log n)$, where δ is the maximum possible gain.*

We show a connection between the advantage over random and the cut norm of the adjacency matrix of the graph G. In Section 4.2, we present a simple algorithm that finds a linear arrangement with advantage over random proportional to the cut norm of the adjacency matrix of the graph. Then, in Section 4.6, we prove using Fourier analysis techniques that the cut norm of the adjacency matrix is within a $\log n$ factor of the optimal gain. We also give an example that shows that our analysis is tight.

4.2 Approximation Algorithm

It will be convenient for us to express different quantities in terms of the adjacency matrix W_G of the directed graph G. For unweighted graphs, we define W_G as follows:

$$W_G(u,v) = \begin{cases} 1, & \text{if } (u,v) \in E; \\ -1, & \text{if } (v,u) \in E; \\ 0, & \text{otherwise.} \end{cases}$$

If both $(u,v) \in E$ and $(v,u) \in E$ then $W_G(u,v) = 0$. For weighted graphs,

$$W_G(u,v) = \mathrm{weight}((u,v)) - \mathrm{weight}((v,u)),$$

where $\mathrm{weight}((u,v)) = 0$ if $(u,v) \notin E$. Below $|E|$ denotes the total weight of all edges.

The gain over random is equal to

$$\mathrm{gain}(G,\pi) = \frac{1}{|E|} \sum_{i<j} W_G(\pi_i, \pi_j).$$

In other words the gain is equal to the sum of the elements in the upper triangle of the matrix W_G, in which rows and columns are arranged according to π, divided by the number of edges.

Let us now describe our approach to solving the problem. First partition the vertices of the graph into three sets A, B and C in a special way. Then randomly permute vertices in each of these sets. Finally, with probability a half output all vertices in the order A, B, C and with probability a half in the order C, A, B.

It is easy to see that all edges from A to B go forward; and all edges from B to A go backward. On the other hand, all other edges go backward or forward with probability exactly a half. Hence, the expected gain is equal to

$$\frac{1}{|E|} \sum_{u \in A;\, v \in B} W_G(u, v). \tag{4.1}$$

We just showed that the maximum gain is greater than or equal to (4.1). It turns out that the converse is also true up to an $O(\log n)$ factor. That is, there always exist disjoint sets A and B for which

$$\frac{1}{|E|} \sum_{u \in A;\, v \in B} W_G(u, v) \geq \frac{\text{gain}(G, \pi)}{O(\log n)},$$

where π is the optimal permutation of the vertices. This statement is the main technical component of the proof and we prove it in Section 4.6.

4.3 Combinatorial Interpretation of Proof

As a prelude to the technical analysis in Section 4.6, we give a simplified overview. The goal of the analysis is to show that if there is an ordering π with gain δ, then there are subsets A and B of such that placing all vertices of A before vertices of B gives gain at least $\delta/\log n$. (In fact, the sets A and B we construct in our analysis are not disjoint, but this is easy to fix.)

Define the length of an edge to be the distance between its end points in the ordering π. We can group edges geometrically by length into $O(\log n)$ groups. If the gain of the ordering π is δ, at least one of these groups must have gain $\delta/\log n$.

We construct the sets A and B by random sampling the positions in the ordering

π. The selection probabilities vary periodically with position. Roughly speaking, if we have period P, this targets the group of edges with edge lengths $\Theta(P)$. Since one of the groups has gain $\delta/\log n$, the sampling process targeted towards that group will generate sets A and B such that the corresponding ordering has gain $\delta/\log n$.

In our proof later, this sampling is incorporated into a certain bilinear form (4.5) we analyze. This expression involves terms $x_k(r)$ and $y_k(r)$ (defined later) that corresponds to selecting sets A and B randomly where the selection probabilities vary periodically with position in the ordering π. For appropriate choice of the period our analysis shows that the bilinear form constructed must have value at least $\delta/\log n$.

This is a somewhat simplistic explanation that ignores several issues. In fact, edges could have both positive and negative contributions to the bilinear form constructed (positive contributions come from edges going from A to B, negative contributions come from edges going from B to A). Our intuitive explanation focused on the contribution from one group of edges, but we need to ensure that the potential negative contributions of the other groups do not overwhelm this. The Fourier machinery we use allows us to properly account for positive and negative contributions.

Now, we do not actually know the optimal ordering π, so we do not really perform this sampling in our algorithm to obtain sets A and B. Instead, we focus on a quantity called the cut norm which we define in the next section. Our existential proof is merely an analysis tool that allows us to prove that the cut norm of the adjacency matrix is large if the gain of *some* ordering is large. The actual algorithm uses Alon and Naor's SDP based approximation for the cut norm. This yields sets A and B from which we obtain an ordering of the vertices.

4.4 Efficient Implementation

We now show how to efficiently find sets A and B that maximize (4.1) within a constant factor and thus obtain an $O(\log n)$ approximation for the maximum gain problem. This problem is closely related to the problem of finding the cut norm of the matrix W_G, which can be approximately solved using the algorithm proposed by Alon and Naor [7].

Definition 4.4. *The cut norm of a matrix $W(u,v)$ is equal to*

$$\|W\|_C = \max_{A,B\subset V} \left| \sum_{u\in A;\, v\in B} W(u,v) \right|.$$

Note that for skew-symmetric matrices

$$\sum_{u\in A;\, v\in B} W(u,v) = -\sum_{u\in A;\, v\in B} W(v,u);$$

and therefore

$$\|W\|_C = \max_{A,B\subset V} \sum_{u\in A;\, v\in B} W(u,v).$$

In this definition it is not required that the sets A and B are disjoint. However, given arbitrary sets A and B, we can always find disjoint sets A' and B' such that

$$\sum_{u\in A';v\in B'} W(u,v) \geq \frac{1}{4} \cdot \sum_{u\in A;\, v\in B} W(u,v). \tag{4.2}$$

In order to do so, we simply partition the vertices of the graph G into two random sets X and Y. Then set $A' = A \cap X$; $B' = B \cap Y$.

Lemma 4.5. *For every skew-symmetric matrix G, the sets A' and B' (as described* *above) are disjoint and satisfy the following equality:*

$$\mathbb{E}\left[\sum_{u\in A',v\in B'} W(u,v)\right] = \frac{1}{4}\sum_{u\in A,v\in B} W(u,v).$$

Proof. The sets A' and B' are disjoint, since the sets X and Y are disjoint. Now, for every distinct vertices $u \in A$ and $v \in B$, the probability that $u \in A'$ and $v \in B'$ equals $1/4$. (Note, that the diagonal entries of W are equal to zero.) □

We now state the result of Alon and Naor [7].

Theorem 4.6 (Alon and Naor [7])**.** *There exists a randomized polynomial time algorithm that given a matrix $W(u,v)$ finds two subsets of indices A and B such that*

$$\left|\sum_{u\in A;\, v\in B} W(u,v)\right| \geq \alpha_{AN} \cdot \|W\|_C,$$

where $\alpha_{AN} \approx 0.56$.

Applying Lemma 4.5, we get the following corollary.

Corollary 4.7. *There exists a randomized polynomial time algorithm that given a skew-symmetric matrix $W(u,v)$ finds two disjoint subsets A and B such that*

$$\sum_{u\in A;\, v\in B} W(u,v) \geq \frac{\alpha_{AN}}{4} \cdot \|W\|_C.$$

This Corollary implies that the algorithm described in the beginning of the section finds a linear arrangement with gain $\alpha_{AN}/4 \cdot \|W_G\|_C/|E|$. Thus, we proved:

Lemma 4.8. *There exists a randomized polynomial time algorithm that given a directed graph G with adjacency matrix $W_G(u,v)$ finds a linear arrangement of the* *vertices with gain*

$$\frac{\alpha_{AN}}{4} \cdot \frac{\|W_G\|_C}{|E|}.$$

4.5 Discrete Fourier Sine and Cosine Transforms

In the next two sections, we use the discrete Fourier *sine* and *cosine* transforms, which are analogs of the discrete Fourier transform, but are less well known. For this reason, we briefly describe them and prove the inversion formulas. The reader familiar with these transforms may freely skip this section.

Let f be a function from $\{1, \ldots, n-1\}$ to $\mathbb{R}$. Then its discrete Fourier sine transform is defined as

$$\hat{f}(t) = \sum_{k=1}^{n-1} \sin\Big(\frac{\pi k t}{n}\Big) f(k).$$

The inversion formula is given by

$$f(k) = \frac{2}{n} \sum_{t=1}^{n-1} \sin\Big(\frac{\pi k t}{n}\Big) \hat{f}(t).$$

To prove this formula we need to verify that the functions

$$k \mapsto \sin(\pi k t/n)$$

form an orthonormal basis, i.e.

$$\frac{2}{n} \sum_{k=1}^{n-1} \sin\Big(\frac{\pi k s}{n}\Big) \sin\Big(\frac{\pi k t}{n}\Big) = \begin{cases} 1, & \text{if } s = t; \\ 0, & \text{if } s \neq t. \end{cases} \tag{4.3}$$

We use the following simple lemma.

Lemma 4.9. *For every $n > 1$ and $0 < t < n$,*

$$\sum_{k=0}^{n-1} e^{\frac{\pi k t i}{n}} = \begin{cases} 1 + i\cot\frac{\pi t}{2n}, & t \text{ is odd;} \\ 0, & t \text{ is even.} \end{cases}$$

Proof. We have

$$\sum_{k=0}^{n-1} e^{\frac{\pi k t i}{n}} = \sum_{k=0}^{n-1} \left(e^{\frac{\pi t i}{n}}\right)^k = \frac{1 - e^{\pi t i}}{1 - e^{\frac{\pi t i}{n}}}.$$

If t is even, then $e^{\pi t i} = 1$ and we are done. Otherwise,

$$\begin{aligned} \frac{1 - e^{\pi t i}}{1 - e^{\frac{\pi t i}{n}}} = \frac{2}{1 - e^{\frac{\pi t i}{n}}} &= \frac{2(1 - e^{-\frac{\pi t i}{n}})}{2(1 - \mathrm{Re}(e^{\frac{\pi t i}{n}}))} \\ &= 1 + i\frac{\sin\frac{\pi t}{n}}{1 - \cos\frac{\pi t}{n}} = 1 + i\cot\frac{\pi t}{2n}. \end{aligned} \tag{4.4}$$

□

Corollary 4.10. *For every $n > 1$ and $0 < t < n$,*

$$\sum_{k=0}^{n-1} \sin\frac{\pi k t}{n} = \begin{cases} \cot\frac{\pi t}{2n}, & t \text{ is odd;} \\ 0, & t \text{ is even.} \end{cases}$$

$$\sum_{k=0}^{n-1} \cos\frac{\pi k t}{n} = \begin{cases} 1, & t \text{ is odd;} \\ 0, & t \text{ is even.} \end{cases}$$

Proof of expression (4.3):

$$\begin{aligned}\frac{2}{n}\sum_{k=1}^{n-1}\sin\Big(\frac{\pi ks}{n}\Big)\sin\Big(\frac{\pi kt}{n}\Big) &= \frac{2}{n}\sum_{k=0}^{n-1}\sin\Big(\frac{\pi ks}{n}\Big)\sin\Big(\frac{\pi kt}{n}\Big)\\ &= \frac{1}{n}\sum_{k=0}^{n-1}\cos\frac{\pi(s-t)k}{n}-\frac{1}{n}\sum_{k=0}^{n-1}\cos\frac{\pi(s+t)k}{n}.\end{aligned}$$

If $s \neq t$, then the sums above are equal by Corollary 4.10. Hence the whole expression is equal to zero. If $s = t$, then

$$\frac{2}{n}\sum_{k=1}^{n-1}\sin\Big(\frac{\pi ks}{n}\Big)\sin\Big(\frac{\pi kt}{n}\Big) = 1-\frac{1}{n}\sum_{k=0}^{n-1}\cos\frac{2\pi\cdot kt}{n} = 1.$$

This finishes the proof of the inversion formula. □

We prove a technical lemma that will be used later.

Lemma 4.11. *For every $n > 1$ and $1 \leq t \leq n$,*

$$\sum_{k=0}^{n-1}(n-k)\sin\frac{\pi kt}{n+1} = \frac{n+(-1)^t}{2\tan\frac{\pi t}{2(n+1)}}.$$

Proof. We have

$$\begin{aligned}\sum_{k=0}^{n-1}(n-k)\sin\frac{\pi kt}{n+1} &= \operatorname{Im}\left(\sum_{k=0}^{n-1}(n-k)\,e^{\frac{\pi kti}{n+1}}\right) = \operatorname{Im}\left(\sum_{l=0}^{n-1}\sum_{k=0}^{l}\, e^{\frac{\pi kti}{n+1}}\right)\\ &= \operatorname{Im}\left(\sum_{l=0}^{n-1}\frac{1-e^{\frac{\pi t(l+1)i}{n+1}}}{1-e^{\frac{\pi ti}{n+1}}}\right) = \operatorname{Im}\left(\frac{n-\sum_{l=1}^{n}e^{\frac{\pi tli}{n+1}}}{1-e^{\frac{\pi ti}{n+1}}}\right)\\ &= \operatorname{Im}\left(\frac{n+1}{1-e^{\frac{\pi ti}{n+1}}}-\frac{1-e^{\pi ti}}{\left(1-e^{\frac{\pi ti}{n+1}}\right)^2}\right).\end{aligned}$$

Similarly to (4.4), we get

$$\operatorname{Im}\left(\frac{n+1}{1-e^{\frac{\pi ti}{n+1}}}\right) = \frac{n+1}{2\tan\frac{\pi t}{2(n+1)}};$$

and

$$\begin{aligned}
\operatorname{Im}\left(\frac{1-e^{\pi ti}}{\left(1-e^{\frac{\pi ti}{n+1}}\right)^2}\right) &= \frac{\operatorname{Im}((1+(-1)^{t+1})(1-e^{-\frac{\pi ti}{n+1}})^2)}{\left(1-e^{\frac{\pi ti}{n+1}}\right)^2\left(1-e^{-\frac{\pi ti}{n+1}}\right)^2} \\
&= \frac{2(1+(-1)^{t+1})(1-\cos\frac{\pi t}{n+1})\sin\frac{\pi t}{n+1}}{(2-2\cos\frac{\pi t}{n+1})^2} \\
&= \frac{(1+(-1)^{t+1})\sin\frac{\pi t}{n+1}}{2(1-\cos\frac{\pi t}{n+1})} = \frac{1+(-1)^{t+1}}{2\tan\frac{\pi t}{2(n+1)}}.
\end{aligned}$$

□

4.6 Cut Norm of Skew-Symmetric Matrices

In this section we will prove that for every linear arrangement π (particularly, for the optimal linear arrangement)

$$\|W_G\|_C \geq \Omega\left(\frac{\operatorname{gain}(G,\pi)|E|}{\log n}\right),$$

which will conclude the proof of Theorem 4.3. Fix an arbitrary linear arrangement π and denote

$$w_{kl} = W(\pi_k, \pi_l).$$

Theorem 4.12. *Let W be an $n \times n$ skew-symmetric matrix. Define*

$$S^+ = S^+(W) = \sum_{1 \le k < l \le n} w_{kl}.$$

Then

$$\|W\|_C \ge \Omega\left(\frac{|S^+|}{\log n}\right).$$

Remark 4.13. *Recall, that* $\mathrm{gain}(G, \pi) = S^+(W_G)/|E|$, *where the columns and rows in W_G are ordered according to the permutation π.*

We need several lemmas.

Lemma 4.14. *Let $\hat{S}_t$ be the discrete Fourier sine transform of a sequence $S_1, \ldots, S_{n-1}$, defined as follows*

$$\hat{S}_t = \sum_{k=1}^{n-1} \sin\Bigl(\frac{\pi k t}{n}\Bigr) S_k.$$

Then

$$\max_t |\hat{S}_t| \ge \frac{\pi}{2 \log n + O(1)} \sum_{k=1}^{n-1} S_k.$$

Proof. The inverse Fourier sine transform is given by

$$S_k = \frac{2}{n} \sum_{t=1}^{n-1} \sin\Bigl(\frac{\pi k t}{n}\Bigr) \hat{S}_t.$$

Hence,

$$\sum_{k=1}^{n-1} S_k = \frac{2}{n} \sum_{k=1}^{n-1} \sum_{t=1}^{n-1} \sin\Bigl(\frac{\pi k t}{n}\Bigr) \hat{S}_t = \frac{2}{n} \sum_{t=1}^{n-1} \left(\sum_{k=1}^{n-1} \sin\Bigl(\frac{\pi k t}{n}\Bigr) \right) \hat{S}_t.$$

By Corollary 4.10,

$$\sum_{k=1}^{n-1} \sin \frac{\pi k t}{n} = \frac{1 - (-1)^t}{2 \tan(\frac{\pi t}{2n})}.$$

Then

$$\begin{aligned}\sum_{k=1}^{n-1} S_k &\le \frac{2}{n}\sum_{t=1}^{n-1}\left|\frac{1-(-1)^t}{2\tan(\frac{\pi t}{2n})}\right| |\hat{S}_t| \\ &\le \frac{2}{n}\left(\sum_{\substack{t=1 \\ t \text{ is odd}}}^{n-1}\frac{2n}{\pi t}\right)\max_t |\hat{S}_t| = \frac{2}{\pi}\left(\log n + O(1)\right)\max_t |\hat{S}_t|,\end{aligned}$$

here we used that $\tan x > x$ for $x \in (0, \pi/2)$. Therefore,

$$\max_t |\hat{S}_t| \ge \frac{\pi}{2\log n + O(1)}\sum_{k=1}^{n-1} S_k.$$

$\square$

Lemma 4.15. *Let W be an $n \times n$ skew-symmetric matrix. Define*

$$S_k = \sum_{j=1}^{n-k} w_{j,j+k}$$

for $1 \le k \le n-1$. And let $\hat{S}_t$ be the discrete Fourier sine transform of S_k. Then

$$\max_{-1\le x_k, y_l \le 1}\sum_{k=1}^{n}\sum_{l=1}^{n} w_{kl}x_k y_l \ge max_t |\hat{S}_t|.$$

Proof. Let $t_0 = \mathrm{argmax}_t\, |\hat{S}_t|$. For every k, l from 1 to n and r from 0 to $n-1$ define

$$\begin{aligned} x_k(r) &= \sin\left(\frac{\pi(k+r)t_0}{n}\right); \\ y_l(r) &= -\cos\left(\pi\frac{(k+r)t_0}{n}\right). \end{aligned}$$

Find the average value of the bilinear form

$$\sum_{k=1}^{n}\sum_{k=1}^{n} w_{kl}x_k(r)y_l(r) \tag{4.5}$$

over r from 0 to $n-1$. Write

$$\frac{1}{n}\sum_{r=0}^{n-1}\left(\sum_{k=1}^{n}\sum_{l=1}^{n} w_{kl}x_k(r)y_l(r)\right) = \frac{1}{n}\sum_{k=1}^{n}\sum_{l=1}^{n}\left(\sum_{r=0}^{n-1} x_k(r)y_l(r)\right) w_{kl}.$$

Observe, that

$$\begin{aligned}\sum_{r=0}^{n-1} x_k(r)y_l(r) &= -\sum_{r=0}^{n-1}\sin\Big(\frac{\pi(k+r)t_0}{n}\Big)\cos\Big(\frac{\pi(l+r)t_0}{n}\Big)\\ &= -\frac{1}{2}\sum_{r=0}^{n-1}\sin\Big(\frac{\pi(k-l)t_0}{n}\Big)+\sin\Big(\frac{\pi(k+l+2r)t_0}{n}\Big)\\ &= \frac{n}{2}\cdot\sin\frac{\pi(l-k)t_0}{n}.\end{aligned}$$

Hence,

$$\begin{aligned}\frac{1}{n}\sum_{r=0}^{n-1}\Big(\sum_{k=1}^{n}\sum_{l=1}^{n} w_{kl}x_k(r)y_l(r)\Big) &=\\ &= \frac{1}{2}\sum_{k=1}^{n}\sum_{l=1}^{n}\sin\Big(\frac{\pi(l-k)t_0}{n}\Big)w_{kl}\\ &= \sum_{1\le k<l\le n}\sin\Big(\frac{\pi(l-k)t_0}{n}\Big)w_{kl}\\ &= \sum_{j=1}^{n-1}\sin\Big(\frac{\pi j t_0}{n}\Big)S_j = \hat{S}_{t_0}.\end{aligned}$$

here we used that the matrix W is skew-symmetric. We got that the average value of bilinear form (4.5) is $\hat{S}_{t_0}$, therefore, there exists r for which the absolute value of

(4.5) is at least $|\hat{S}_{t_0}|$. Since the bilinear form is an odd function as a function of x (when y is fixed), the maximum of (4.5) is at least $|\hat{S}_{t_0}|$. $\square$

Corollary 4.16. *Let W, S_k and $\hat{S}_t$ be as in Lemma 4.15. Then*

$$\max_{x_k, y_l \in \{-1,1\}} \sum_{k=1}^{n} \sum_{l=1}^{n} w_{kl} x_k y_l \geq \max_t |\hat{S}_t|.$$

Proof. The maximum of the bilinear form is attained at a vertex of the cube. $\square$

The following observation is due to Alon and Naor [7]. We prove it here for completeness.

Lemma 4.17 (Alon and Naor [7])**.**

$$\|W\|_C \geq \frac{1}{4} \max_{x_k, y_l \in \{-1,1\}} \sum_{k=1}^{n} \sum_{l=1}^{n} w_{kl} x_k y_l.$$

Proof. Fix x_k and y_l for which the maximum of the bilinear form is attained. Define sets $I^+ = \{k : x_k = 1\}$, $I^- = \{k : x_k = -1\}$, $J^+ = \{l : y_l = 1\}$ and $J^- = \{l : y_l = 1\}$. Now notice that

$$\begin{aligned} \sum_{k=1}^{n} \sum_{l=1}^{n} w_{kl} x_k y_l \quad = \quad & \sum_{\substack{k \in I^+ \\ l \in J^+}} w_{kl} + \sum_{\substack{k \in I^- \\ l \in J^-}} w_{kl} \\ & - \sum_{\substack{k \in I^+ \\ l \in J^-}} w_{kl} - \sum_{\substack{k \in I^- \\ l \in J^+}} w_{kl}. \end{aligned}$$

Each of the terms on the right hand side does not exceed the cut norm $\|W\|_C$ in absolute value. Hence

$$\sum_{k=1}^{n} \sum_{l=1}^{n} w_{kl} x_k y_l \leq 4\|W\|_C.$$

$\square$

Now we are ready to prove Theorem 4.12.

Proof of Theorem 4.12. Let

$$S_k = \sum_{j=1}^{n-k} w_{j,j+k}$$

for $1 \le k \le n-1$. Then

$$S^+ = \sum_{k=1}^{n-1} S_k.$$

By Lemma 4.14 and Corollary 4.16,

$$\max_{x_k, y_l \in \{-1,1\}} \sum_{k=1}^{n} \sum_{l=1}^{n} w_{kl} x_k y_l \ge \frac{\pi}{2\log n + O(1)} \sum_{k=1}^{n-1} S_k.$$

Now by Lemma 4.17,

$$\|W\|_C \ge \frac{\pi}{8\log n + O(1)} \sum_{k=1}^{n-1} S_k = \frac{\pi\ S^+}{8\log n + O(1)}.$$

□

4.7 Lower Bound

We will now prove that the bound in Theorem 4.12 is essentially tight.

Theorem 4.18. *For every $n > 1$, there exists a nonzero $n \times n$ skew-symmetric matrix W such that*

$$\|W\|_C \le O(|S^+|/\log n), \tag{4.6}$$

where $S^+ = S^+(W)$ is defined in Theorem 4.12.

Proof. Consider the matrix W defined by

$$w_{kl} = \sum_{t=1}^{n} \sin \frac{\pi(l-k)t}{n+1}.$$

Clearly, the matrix W is skew-symmetric. Let us compute S^+ (see Lemma 4.11 for details).

$$\begin{aligned} S^+ &= \sum_{1\le k<l\le n} w_{kl} = \sum_{1\le k<l\le n} \sum_{t=1}^{n} \sin \frac{\pi(l-k)t}{n+1} \\ &= \sum_{t=1}^{n} \sum_{k=0}^{n-1} (n-k) \sin \frac{\pi kt}{n+1} = \frac{1}{2} \sum_{t=1}^{n} \frac{n+(-1)^t}{\tan(\frac{\pi t}{2(n+1)})}. \end{aligned}$$

Replace $1/\tan x$ with $1/x + O(1)$,

$$S^+ \ge \frac{1}{2} \sum_{t=1}^{n} \left(\frac{2n^2}{\pi t} + O(n) \right) = \frac{n^2}{\pi} (\log n + O(1)).$$

We are now going to estimate $\|W\|_C$. Pick sets I and J that maximize $\left|\sum_{k\in I;\, l\in J} w_{kl}\right|$. We have

$$\begin{aligned} \|W\|_C &= \Big| \sum_{k\in I;\, l\in J} w_{kl} \Big| = \Big| \sum_{k\in I;\, l\in J} \sum_{t=1}^{n} \sin\Big(\frac{\pi(k-l)t}{n+1}\Big) \Big| \\ &= \Big| \sum_{k\in I;\, l\in J} \sum_{t=1}^{n} \sin\Big(\frac{\pi kt}{n+1}\Big) \cos\Big(\frac{\pi lt}{n+1}\Big) - \cos\Big(\frac{\pi kt}{n+1}\Big) \sin\Big(\frac{\pi lt}{n+1}\Big) \Big| \\ &\le \Big| \sum_{k\in I;\, l\in J} \sum_{t=1}^{n} \sin\Big(\frac{\pi kt}{n+1}\Big) \cos\Big(\frac{\pi lt}{n+1}\Big) \Big| + \Big| \sum_{k\in I;\, l\in J} \sum_{t=1}^{n} \cos\Big(\frac{\pi kt}{n+1}\Big) \sin\Big(\frac{\pi lt}{n+1}\Big) \Big|. \end{aligned}$$

Estimate the first term. Let x_k be the indicator of the set I: $x_k = 1$ if $k \in I$, $x_k = 0$

otherwise. Let y_k be the indicator of the set J. Then the first term equals

$$\begin{aligned}
T_I &\equiv \sum_{t=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} \sin\Bigl(\frac{\pi kt}{n+1}\Bigr)\cos\Bigl(\frac{\pi lt}{n+1}\Bigr) x_k y_l \\
&= \sum_{t=1}^{n}\left(\sum_{k=1}^{n}\sin\Bigl(\frac{\pi kt}{n+1}\Bigr)x_k\right)\left(\sum_{l=1}^{n}\cos\left(\frac{\pi lt}{n+1}\right)y_l\right) \\
&= \sum_{t=1}^{n}\hat{x}_t\hat{y}_t,
\end{aligned}$$

where $\hat{x}_t$ is the Fourier *sine* transform of x_t

$$\hat{x}_t = \sum_{k=1}^{n}\sin\Bigl(\frac{\pi kt}{n+1}\Bigr)x_k, \quad 1 \le t \le n;$$

and $\hat{y}_t$ is the discrete Fourier *cosine* transform of y_k (extended by $y_0 = y_{n+1} = 0$)

$$\hat{y}_t = \sum_{k=1}^{n}\cos\Bigl(\frac{\pi kt}{n+1}\Bigr)y_k, \quad 0 \le t \le n+1.$$

By the Cauchy–Bunyakovski–Schwartz inequality and Bessel's inequality, we have

$$|T_I| \;=\; \Bigl|\sum_{t=1}^{n}\hat{x}_t\hat{y}_t\Bigr| \le \sqrt{\sum_{t=1}^{n}\hat{x}_t^2\sum_{t=1}^{n}\hat{y}_t^2} \le \frac{n+1}{2}\sqrt{\sum_{k=1}^{n}x_k^2\sum_{l=1}^{n}y_l^2} \le \frac{n^2}{2} + O(n).$$

Similarly, the second term is at most $n^2/2 + O(n)$. Hence

$$\|W\|_C = \Bigl|\sum_{k\in I;\, l\in J} w_{kl}\Bigr| \le n^2 + O(n) \le \frac{\pi S^+}{\log n}(1+o(1)).$$

This finishes the proof. $\square$

We presented a matrix W with real entries for which bound (4.6) holds. This matrix corresponds to a directed graph with weighted edges. However, it can be transformed to a matrix with entries -1, 0 and 1, which corresponds to an unweighted directed graph.

Corollary 4.19. *There exists a matrix $\widetilde{W}$ with entries -1, 0 and 1 that satisfies bound (4.6).*

Proof. Let W be the matrix from Theorem 4.18. By scaling, we may assume that the largest entry in W equals 1 in absolute value. Let $N = 4n^4$. We construct the matrix $\widetilde{W}$ by replacing each entry w_{ij} of W with an $N \times N$ block matrix R^{ij} that has the following properties. First, each entry of R^{ij} is either -1, 0 or 1. Second, for every two sets $A, B \subset \{1, \ldots, N\}$,

$$\frac{1}{N^2}\left|\sum_{k\in A,\, l\in B} R^{ij}_{kl} - w_{ij}|A||B|\right| < \frac{2}{\sqrt{N}} = \frac{1}{n^2}. \tag{4.7}$$

We prove that such matrix R^{ij} exists using the probabilistic method (see Alon and Berger [4] for a similar argument). Let every entry of R^{ij} be equal to $\operatorname{sgn}(w_{ij})$ with probability $|w_{ij}|$ and equal to 0 with probability $1 - |w_{ij}|$. Then, by the Bernstein inequality/Chernoff bound, for fixed sets $A, B \subset \{1, \ldots, N\}$, we have

$$\Pr\left(\left|\sum_{k\in A,\, l\in B} R^{ij}_{kl} - w_{ij}|A||B|\right| > \frac{2N^2}{\sqrt{N}}\right) < 2\exp\left(-2\cdot(2N^{3/2})^2/N^2\right) \leq 2e^{-8N}.$$

Since there are 2^{2N} distinct pairs of sets A and B, and $2^{2N}\cdot 2e^{-8N} < 1$, there exists a matrix R^{ij} that satisfies inequality (4.7) for all sets A and B simultaneously. (To ensure that the matrix $\widetilde{W}$ is skew-symmetric, we use this argument to find matrices R^{ij} for $i < j$; we let $R^{ji} = -(R^{ij})^T$.)

We verify that the matrix $\widetilde{W}$ satisfies bound (4.6). Let us estimate the cut norm

$\|\widetilde{W}\|_C$. Let A and B be the sets of indices such that

$$\|\widetilde{W}\|_C = \left| \sum_{k\in A,\, l\in B} \widetilde{w}_{kl} \right|.$$

Denote the restriction of A to the set of row indices of submatrices R^{i*} by A_i; denote the restriction of B to the set of column indices of submatrices R^{*j} by B_j. Then

$$\begin{aligned}
\|\widetilde{W}\|_C &= \left| \sum_{\substack{1\le i\le n\\ 1\le j\le n}} \sum_{\substack{k\in A_i\\ l\in B_j}} R^{ij}_{kl} \right| \\
&\le N^2 \left| \sum_{\substack{1\le i\le n\\ 1\le j\le n}} w_{ij} \frac{|A_i|}{N}\frac{|B_j|}{N} \right| + N^2 \sum_{\substack{1\le i\le n\\ 1\le j\le n}} \frac{1}{n^2} \\
&\le N^2\left(\|W\|_C + 1\right) \le 2N^2\|W\|_C.
\end{aligned}$$

Here, first, we used that, $0 \le |A_i|/N \le 1$ and $0 \le |B_j|/N \le 1$, so $\sum_{ij} w_{ij} \cdot |A_i|/N \cdot |B_j|/N$ does not exceed the cut norm; second, we used that at least one entry in W equals to 1 in absolute value, so $\|W\|_C \ge 1$. On the other hand, we can estimate $S^+(\widetilde{W})$ as follows (assume without loss of generality that $S^+(W) \ge 0$)

$$\begin{aligned}
|S^+(\widetilde{W})| &= \sum_{1\le i<j\le n} \sum_{k,l} R^{ij}_{kl} \ge \sum_{1\le i<j\le n} N^2(w_{ij} - 1/n^2) \\
&\ge N^2(S^+(W) - 1/2) \ge N^2\, S^+(W)(1 - o(1)).
\end{aligned}$$

Here we used that $S^+(W) = \Omega(\log n)\|W\|_C = \Omega(\log n) = \omega(1)$. Combining the bounds for $\|\widetilde{W}\|_C$ and $|S^+(\widetilde{W})|$ with bound (4.6), we get

$$\|\widetilde{W}\|_C \leq O(S^+(\widetilde{W})/\log n).$$

This concludes the proof. $\square$

Chapter 5

Conclusions and Open Questions

In this dissertation, we introduced a new graph parameter $K(G)$ that is equal to the integrality gap of a natural SDP relaxation for the *MAX QP* problem. We showed that

$$C_1 \log w(G) \leq K(G) \leq C_2 \vartheta(\bar{G}),$$

where $w(G)$ is the clique number of G; and $\vartheta(\bar{G})$ is the Lovász theta function of the complement of G (which is always smaller than the chromatic number of G). Even though our lower and upper bounds are equal for several important families of graphs, in general case there is a gap between them. Thus one of the main open questions in this dissertation is whether $K(G)$ always equals $\Theta(\vartheta(\bar{G}))$. From a mathematical point of view it is also interesting to find the exact value of the (classical) Grothendieck constant, and understand for which matrices A and which configurations of vectors $z_1, \dots, z_n$ the (classical) Grothendieck's inequality is tight.

The result of Chapter 3 is optimal (up to a constant factor) assuming the Unique Games Conjecture (UGC). As a consequence any improvements in the approximation guarantees for *MAX k-CSP* will refute the UGC. While it is not clear if the conjecture is true or false, one can try to disprove it by improving the approximation ratio of our algorithm.

In Chapter 4, we presented a new method for finding linear arrangements of directed graphs. Using it we developed a $\log n$-approximation algorithm for the *Advantage over Random for Maximum Acyclic Subgraph* problem. One promising direction for improving the approximation ratio for *Maximum Acyclic Subgraph* is to combine our new method with other methods for rounding SDP relaxations.

Bibliography

[1] A. Agarwal, M. Charikar, K. Makarychev, and Y. Makarychev. $O(\sqrt{\log n})$ approximation algorithms for Min UnCut, Min 2CNF Deletion, and directed cut problems. Proceedings of the 37^{th} ACM STOC, pp. 573–581, 2005.

[2] N. Alon. Explicit Ramsey graphs and orthonormal labelings. The Electronic J. Combinatorics 1 (1994), R12, 8pp.

[3] N. Alon. Covering a hypergraph of subgraphs. Discrete Mathematics 257 (2002), 249-254.

[4] N. Alon and E. Berger. The Grothendieck constant of random and pseudo-random graphs. Discrete Optimization, to appear.

[5] N. Alon, G. Gutin and M. Krivelevich. Algorithms with large domination ratio. J. Algorithms 50 (2004), 118-131.

[6] N. Alon, K. Makarychev, Y. Makarychev, and A. Naor. Quadratic Forms on Graphs. Inventiones Mathematicae, vol. 163, no. 3, pp. 499–522, March 2006. Conference version in the 37^{th} ACM STOC, pp. 486–493, 2005.

[7] N. Alon and A. Naor. Approximating the Cut-Norm via Grothendieck's Inequality. Proceedings of the 36^{th} ACM STOC, pp. 72–80, 2004.

[8] N. Alon and A. Orlitsky. Repeated communication and Ramsey graphs. IEEE Transactions on Information Theory 41 (1995), 1276-1289.

[9] G. Andersson, L. Engebretsen, and J. Håstad. A New Way of Using Semidefinite Programming with Applications to Linear Equations mod p. Journal of Algorithms, vol. 39, pp. 162–204, 2001.

[10] S. Arora and B. Barak. Computational Complexity: A Modern Approach. Manuscript.

[11] P. Austrin. Balanced Max 2-Sat might not be the hardest. In Proceedings of the 39th ACM Symposium on Theory of Computing, pp. 189–197, 2007.

[12] N. Bansal, A. Blum and S. Chawla. Correlation Clustering. Proceedings of the 43^{th} IEEE FOCS, 238-247, 2002.

[13] F. Barahona. On the computational complexity of Ising spin glass models. J. Phys. A: Math. Gen. 15 (1982), 3241-3253.

[14] B. Berger and P. Shor. Approximation algorithms for the maximum acyclic subgraph problem. Proceedings of the first Annual ACM-SIAM SODA, pp. 236–243, 1990.

[15] M. Charikar, V. Guruswami and A. Wirth. Clustering with qualitative information. Proceedings of the 44^{th} IEEE FOCS, 524–533, 2003.

[16] M. Charikar, K. Makarychev, and Y. Makarychev. Near-Optimal Algorithms for Maximum Constraint Satisfaction Problems. Proceedings of the 18th ACM-SIAM SODA, pp. 62–68, 2007.

[17] M. Charikar, K. Makarychev, and Y. Makarychev. On the Advantage over Random for Maximum Acyclic Subgraph. Proceedings of 48^{th} IEEE FOCS, pp. 625-633, 2007.

[18] M. Charikar and A. Wirth. Maximizing quadratic programs: extending Grothendieck's inequality. Proceedings of 45th IEEE FOCS, pp. 54–60, 2004.

[19] G. Ding, P. D. Seymour and P. Winkler. Bounding the vertex cover number of a hypergraph, *Combinatorica* **14** (1994), 23–34.

[20] L. Engebretsen and V. Guruswami. Is constraint satisfaction over two variables always easy? Random Struct. Algorithms, vol. 25, no. 2, pp. 150–178, 2004.

[21] U. Feige and M. Goemans. Approximating the value of two prover proof systems, with applications to MAX 2SAT and MAX DICUT. Proceedings of the 3^{rd} IEEE Israel Symposium on the Theory of Computing and Systems, pp. 182–189, 1995.

[22] A. M. Frieze and R. Kannan. Quick Approximation to matrices and applications, Combinatorica **19** (1999), 175–200.

[23] 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.

[24] A. Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques. Bol. Soc. Mat. Sao paolo, 8:1-79, 1953.

[25] M. Grötschel, L. Lovász and A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. Combinatorica 1 (1981), 169–197.

[26] R. Hassin and S. Rubinstein. Approximations for the maximum acyclic subgraph problem. Source Information Processing Letters, vol. 51, issue 3, pp. 133–140, 1994.

[27] G. Hast. Approximating MAX kCSP - Outperforming a random assignment with almost a linear factor. Proceedings of the 32^{nd} International Colloquium on Automata, Languages and Programming, pp. 956–968, 2005.

[28] G. Hast. Beating a Random Assignment. APPROX 2005, pp. 134–145.

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

[30] J. Håstad. Every 2-CSP allows nontrivial approximation. Proceedings of the 37^{th} ACM STOC, pp. 740–746, 2005.

[31] J. Håstad and V. Srinivasan. On the advantage over a random assignment. Random structures and Algorithms, vol. 25:2, pp. 117–149, 2004.

[32] W. B. Johnson and J. Lindenstrauss. Basic concepts in the geometry of Banach spaces, Handbook of the geometry of Banach spaces, Vol. I, 1–84, North-Holland, Amsterdam, 2001.

[33] F. Juhász. The asymptotic behaviour of Lovász' θ function for random graphs, Combinatorica 2 (1982), 153–155.

[34] V. Kann. On the Approximability of NP-complete Optimization Problems. Ph. D. thesis, Dept. of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, 1992.

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

[36] B. S. Kashin and S. J. Szarek. On the Gram Matrices of Systems of Uniformly Bounded Functions. Proceedings of the Steklov Institute of Mathematics, Vol. 243, 2003, pp. 227–233.

[37] S. Khot. On the power of unique 2-prover 1-round games. Proceedings of 34th ACM STOC, pp. 767–775, 2002.

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

[39] S. Khot and R. O'Donnell. *SDP gaps and UGC-hardness for MaxCutGain.* Proceedings of the 47^{th} IEEE FOCS, pp. 217–226, 2006.

[40] S.-J. Kim, A. Kostochka and K. Nakprasit. On the Chromatic Number of Intersection Graphs of Convex Sets in the Plane, The Electronic J. Combinatorics 11 (2004), R52, 12pp.

[41] J. Krivine. Sur la constante de Grothendieck. C. R. Acad. Sci. Paris Ser. A-B, 284:445-446, 1977.

[42] M. Lewin, D. Livnat, and U. Zwick. Improved Rounding Techniques for the MAX 2-SAT and MAX DI-CUT Problems. Proceedings of the 9^{th} International IPCO Conference on Integer Programming and Combinatorial Optimization, pp. 67–82, 2002.

[43] J. Lindenstrauss and A. Pełczyński. Absolutely summing operators in L_p spaces and their applications. Studia Math. 29, 275-326, 1968.

[44] L. Lovász. Kneser's conjecture, chromatic number and homotopy. Journal of Combinatorial Theory, 25:319–324, 1978.

[45] L. Lovász. On the Shannon capacity of a graph. IEEE Transactions on Information Theory 25(1) (1979), pp. 1–7.

[46] L. Lovász and M. D. Plummer. Matching Theory. North Holland, Amsterdam (1986).

[47] S. Matuura and T. Matsui. New approximation algorithms for MAX 2SAT and MAX DICUT. Journal of Operations Research Society of Japan, vol. 46, pp. 178–188, 2003.

[48] A. Megretski. Relaxation of Quadratic Programs in Operator Theory and System Analysis. In Systems, Approximation, Singular Integral Operators, and Related Topics (Bordeaux, 2000), Basel: Birkhäuser, 2001, pp. 365–392.

[49] A. Nemirovski, C. Roos and T. Terlaky. On Maximization of Quadratic Form over Intersection of Ellipsoids with Common Center. Mathematical Programming, 1999, Vol. 86 Issue 3, pp. 463–473.

[50] Y. Nesterov. Quality of semidefinite relaxation for nonconvex quadratic optimization. CORE Discussion Paper 9719, March 1997.

[51] A. Newman. The Maximum Acyclic Subgraph Problem and Degree-3 Graphs. Proceedings of APPROX 2001, pp. 147–158.

[52] A. Newman. Cuts and Orderings: On Semidefinite Relaxations for the Linear Ordering Problem. Proceedings of APPROX 2004, pp. 195–206.

[53] A. Newman and S. Vempala. Fences Are Futile: On Relaxations for the Linear Ordering Problem. Proceedings of the 8^{th} Conference on Integer Programming and Combinatorial Optimization, pp. 333–347, 2001.

[54] C. H. Papadimitriou and M. Yannakakis. Optimization, Approximation, and Complexity Classes. J. Comput. System Sci. 43, pp. 425–440, 1991.

[55] E. Rietz. A proof of the Grothendieck inequality. Israel J. Math, vol. 19, pp. 271–276, 1974.

[56] A. Samorodnitsky and L. Trevisan. Gowers Uniformity, Influence of Variables, and PCPs. Proceedings of the 38th ACM STOC, pp. 11–20, 2006.

[57] P. D. Seymour. Packing directed circuits fractionally. Combinatorica (15), pp. 281–288, 1995.

[58] M. Talagrand. Spin glasses: a challenge for mathematicians. Cavity and mean field models. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge, 46. Springer-Verlag, Berlin, 2003.

[59] L. Trevisan. Parallel Approximation Algorithms by Positive Linear Programming. Algorithmica, vol. 21, no. 1, pp. 72–88, 1998.

[60] U. Zwick. Finding almost satisfying assignments. Proceedings of the 30th ACM STOC, pp. 551–560, 1998.

[61] U. Zwick. Analyzing the MAX 2-SAT and MAX DI-CUT approximation algorithms of Feige and Goemans. Manuscript. Available at `www.cs.tau.ac.il/~zwick/my-online-papers.html`.