UNIQUE GAMES CONJECTURE : THE BOOLEAN HYPERCUBE AND CONNECTIONS TO GRAPH LIFTS

BY

NAMAN AGARWAL

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Computer Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2014

Urbana, Illinois

Adviser:

Assistant Professor Alexandra Kolla

Abstract

In this thesis we consider two questions motivated by the Unique Games Conjecture . The first question is concerned with the validity of the Unique Games Conjecture when the constraint graph is restricted to the Boolean Hypercube. The Boolean Hypercube is a well studied graph family on which existing spectral methods fail to achieve a sub exponential time bound. We initiate the study of the behaviour of the standard semi-definite program on the Hypercube. We construct an almost optimal integrality gap instance on the Hypercube for the Goemans-Williamson semidefinite program (SDP) for Max-2-LIN($\mathbb{Z}_2$). We conjecture that augmenting the SDP with triangle inequalities makes the SDP exact upto constants on the Hypercube. We further establish connections between the integrality gap of the SDP and Mutlicommodity flow-cut gaps which may lead to an understanding of the behaviour of the SDP on general families of graphs. As a quick corollary we establish that the SDP is exact for planar graphs.

The second question is concerned with spectrum of label extended graphs of Unique Games instances. Such graphs have been extensively studied under the name of Graph Lifts. The main motivation for studying lifts has been understanding Ramanujan expander graphs via two key questions: Is a "typical" lift of an expander graph also an expander; and how can we (efficiently) construct Ramanujan expanders using lifts?

In our work we continue the study of Graph Lifts and show that, for random shift k-lifts, if all the nontrivial eigenvalues of a d-regular graph G are at most λ in absolute value, then with high probability depending only on the number n of nodes of G (and not on k), the absolute value of every nontrivial eigenvalue of the lift is at most $\mathcal{O}(\lambda)$. This improves upon factors of $log(d)$ in the case when $k = 2$. Other results on random lifts have focused on the case when $k \to \infty$ making their results asymptotically true with high probability in the degree of the lift k. To the best of our knowledge, our result is the first upperbound on spectra of lifts for bounded $k > 2$. Our result in particular implies that a typical small lift of a Ramanujan graph is almost Ramanujan, and we believe it will prove crucial in constructing large Ramanujan expanders of all degrees. We also establish a novel characterization of the spectrum of shift lifts by the spectrum of certain k symmetric matrices, that generalize the signed adjacency matrix. We believe that this characterization is of independent interest.

To My Parents and My Friends.

Acknowledgment

First and foremost I would like to thank my advisor Prof. Alexandra Kolla for her invaluable support during my graduate studies at UIUC. She introduced me to problems related to the Unique Games Conjecture and research in general. This thesis would not have been possible without her ideas, guidance and motivation. Apart from research she has always been an extremely helpful source of advice on different aspects of graduate life and I feel extremely lucky to have had her as my advisor.

I thank my collaborators Prof. Luca Trevisan, Prof. Guy Kindler and Vivek Madan for their ideas and support. Discussions with them on various topics including the problem in this thesis have been a great learning experience. I would also like to thank Nikhil Srivastava for enlightening discussions regarding the Graph Lifts problem.

I thank the Computer Science Department at UIUC for providing a wonderful environment for my growth as a graduate student. I would also like to thank the Simons Institute for Theory of Computing where I was a visitor for a semester for giving me the opportunity to attend immensely educational lectures and providing me with a wonderful environment for research.

I am greatly indebted to the Algorithms and Theory Group at UIUC for introducing me to whatever little I know of this area through various courses, seminars and casual discussions. I would like to thank all my friends at UIUC for keeping me 'sane' by being the perfect breakaway from the rigours of graduate life. I would especially like to thank Suvarna Menon for her constant and unwavering support throughout my graduate studies at UIUC. I am grateful to her for being an editor for pretty much everything I have written during my Masters and for being my go-to person for all my problems. Without her care and support this work would have been unimaginable.

Lastly I would like to thank my parents for their undying support towards my studies. They have taught me some wonderful life lessons which have always stood by me through the toughest of times. I feel blessed to have had their support for every endeavour I have ever undertaken.

Table of Contents

List of Figures

Chapter 1

Introduction

Finding *efficient* algorithms for combinatorial optimization problems has been a central focus of Theoretical Computer Science since its inception. The most natural and studied notion of efficiency in this context is the notion of polynomial time algorithms. Indeed we call algorithms efficient if they are guaranteed to run and succeed within time growing as some small degree polynomial of the input size. Many problems including the MINIMUM SPANNING TREE, MAXIMUM MATCHING etc. are known to have such *polynomial* time algorithms. However there is a large collection of such problems including MAX-CUT, GRAPH COLORING, TRAVELLING SALESMAN etc. where it is not known whether such polynomial time algorithms exist. One striking feature of these problems is that they fall in a class which is complete under polynomial time reduction to the 3-SATISFIABILITY problem. Indeed this class referred to as NP forms the basis for the most celebrated open problem in Theoretical Computer Science known as the P vs NP problem.

Most combinatorial optimization problems known to be NP-hard have real world applications where often if the optimal solution is hard to find, even producing an approximation to it suffices. In particular we say that a given algorithm is an ϵ approximation algorithm to a problem P if the output of the algorithm is guaranteed to be within ϵ factor of the optimal. Over the last three decades tremendous research has gone into producing efficient algorithms to many known NP-hard optimization problems with as good an approximation guarantee as possible. One such optimization problem which has attracted a lot of attention is the well known MAX-CUT problem.

Definition 1.0.1 (MAX-CUT)**.** *Given a graph $G = (V, E)$ output a subset of vertices $S \subseteq V$ such that the number of edges leaving the subset $E[S, V - S]$ is maximized.*

The MAX-CUT problem finds applications in statistical physics, circuit layout design, social networks and classification. MAX-CUT was one of the first problems known to be NP-complete [Kar72]. It is easy to see that a trivial randomized approximation algorithm achieves an approximation guarantee of 0.5 in expectation. In a famous paper Goemans-Williamson [GW95] gave a semi-definite programming based algorithm that achieves an approximation ratio $\alpha_{GW} \sim 0.878$. The next natural question to be asked is whether this is tight. Can we achieve a better approximation for MAX-CUT ?

The first and biggest step towards answering the above question was taken by the highly influential work of Arora et al. [ALM^{+}98, AS98] known famously as the PCP Theorem. In particular the work showed that for MAX-CUT and for a lot of other problems the best possible approximation ratio is bounded away from 1. The PCP Theorem initiated a large body of work in the area of Hardness of Approximation. One of the first and the most significant work in this area was by Hastad [Hås01] who proved that many known approximation algorithms achieved the best possible approximation ratio under the assumption that NP is not in P. Perhaps surprisingly the work showed that trivial randomized algorithms for many problems including 3-SAT achieved the best approximation ratio possible. Hastad's work also gave a bound on the optimal approximation ratio of the MAX-CUT problem showing that it was NP-hard to achieve an approximation factor better than $\sim \frac{16}{17}$. Trying to reduce the gap between the known hardness and the achievable approximation ratios remains one of the most fascinating and important questions in Theoretical Computer Science.

Following Hastad's work Khot [Kho02] proposed a generic way to close the gap between many such problems. He proposed the celebrated Unique Games Conjecture which we describe in the next section.

1.1 The Unique Games Conjecture

Khot [Kho02] proposed the following conjecture known famously as the Unique Games Conjecture.

Conjecture 1 (Unique Games Conjecture)**.** *(UGC) For any constants $\epsilon, \delta > 0$, there exists a $k \geq k(\epsilon, \delta)$ such that it is NP-hard to distinguish between instances of Unique Games with alphabet size k where at least $1 - \epsilon$ fraction of constraints are satisfiable and those where at most δ fraction of constraints are satisfiable.*

A full definition of a Unique Games instance appears in Section 2.2, but for now one may think of an instance with alphabet size k as a system of constraints where variables take values in $\mathbb{Z}_k$, and each constraint is a linear equation $\bmod k$ involving two variables(or more generally a permutation mapping $[k]$ to $[k]$). The *constraint graph* of such an instance is a graph that has a vertex for every variable and an edge for every pair of variables that appear together in one of the constraints.

The Unique Games Conjecture has generated a lot of attention over the past 10 years serving as a basis for establishing optimal approximation ratios for a large number of problems. In particular assuming the Unique Games Conjecture tight bounds for the approximability of several fundamental optimization problems, including Vertex Cover [KR03], MaxCUT [KKMO04] and non-uniform sparsest cut [CKK^{+}06, KV05] have been established.

In recent years the UGC has been found to be intimately connected to the power of semi-definite programming (SDP). One can trace the first instance of this connection to the paper by Goemans and Williamson [GW95] on the Max-Cut problem (note that an instance of Max-Cut can be thought of as a linear equation system over $\mathbb{Z}_2$, and thus it is a Unique Games Instance for alphabet size 2). The SDP proposed by Goemans-Williamson can be easily generalized to case of Unique Games over two alphabets. A more convenient way to view the guarantee achieved by the Goemans-Williamson algorithm is that given an instance where the maximum cut is of size $(1-\epsilon)$, it produces a cut that satisfies at least $(1-(2/\pi)\sqrt{\epsilon})$ fraction of the constraints. A matching integrality gap for this relaxation was found by [Kar99] and [FS02], and in [KKMO04] it was proven that if the Unique Games Conjecture is correct, then the Goemans-Williamson algorithm achieves the best possible approximation ratio.

Theorem 1.1.1. *Assume the Unique Games Conjecture. Then for all sufficiently small $\epsilon > 0$, it is NP-hard to distinguish instances of Max-2-LIN($\mathbb{Z}_2$) that are at least $(1-\epsilon)$-satisfiable from instances that are at most $(1-(2/\pi)\sqrt{\epsilon})$-satisfiable.*

The Goemans-Williamson SDP algorithm was later extended to general Unique Games by [CMM06a], and the approximation ratio achieved by [CMM06a] is shown to be tight in [KKMO04]. In a unifying work which best exhibits the relationship between the limitations of Semi-Definite programming and the Unique Games Conjecture Raghavendra [Rag08] showed that for every constraint satisfaction problem there is a polynomial-time, semi-definite programming based algorithm which, if UGC is true, achieves the best possible approximation for the problem. We refer the reader to the survey on Unique Games Conjecture by Khot [Kho10] for a comprehensive survey of these results.

One of the most striking feature about the Unique Games Conjecture is that there is limited evidence in favor of or against it. On one hand it forms the barrier for present algorithmic techniques and on the other there has been steady progress towards its refutation. The most promising algorithmic approaches rely on semidefinite programming relaxations. Various semidefinite programs and rounding schemes have been studied, among which are [Kho02,Tre05, GT06,CMM06a,CMM06b]. Integrality gap results have been sought to understand the limitation of current approaches and as partial evidence of the hardness of unique games. Integrality gap instances have been constructed for the basic relaxation in [KV05,RS09,BGH$^+$11,KPS10], and these instances show that the basic relaxation, and certain extensions of it, cannot be used to refute the UGC. A polynomial-time solvable extension of the basic relaxation, however, solves near-optimally all the known integrality gap instances [BBH$^+$12], and it is a possible candidate for an algorithm refuting the UGC.

In [ABS10], a sub-exponential ($2^{n^{\mathcal{O}(\epsilon)}}$-time) algorithm for general instances is given, based on spectral techniques. Note that the presence of a sub-exponential type algorithm poses one of the most compelling reason to not totally believe in the Unique Games Conjecture . The Exponential Time Hypothesis[ETH] formulated by Impagliazzo et al. [IPZ01] states that there is no sub-exponential algorithm for 3-SAT and use this to show a large class of problems (VERTEX-COVER, K-COLOURING) exhibit the same property. It is conjectured that in fact [BJK05] such a dichotomy holds for constraint problems in general that they either require fully exponential time or are in P. Results in recent years tends to suggest such a dichotomy might hold for approximation algorithms as well. The results of

Moshkovitz et al. [MR08] imply that under the ETH any improvement over the best possible approximation for 3SAT requires nearly exponential time. However the presence of sub-exponential algorithms for the Unique Games Conjecture rules such a dichotomy out for any problem whose hardness has been established using the Unique Games Conjecture .

Unique Games are known to have good polynomial-time or quasi-polynomial-time spectral approximation algorithms for large classes of instances. These include expanders [AKK$^+$08, MM10], local expanders [AIMS10, RS10], and more generally, graphs with few large eigenvalues [Kol10]. Spectral algorithms also solve nearly-optimally, in time at most $2^{n^{o(1)}}$, all the known integrality gap instances mentioned above (here n denotes the number of vertices of the constraint graph), which, as mentioned, are also well approximated in polynomial time by the SDP studied in [BBH$^+$12]. Thus, such instances are known to not be *hard* for Unique Games following the above discussion. We note that instances generated in various semi-random models [KMM11] are solvable in polynomial time.

Algorithmic work on restricted classes of instances can lead to progress toward the resolution of the UGC, irrespective of whether it ultimately proves to be true or false. If the UGC is false, then the algorithmic work on restricted classes of instances represents steady progress toward its refutation. Indeed, the algorithmic breakthrough of [ABS10] came by combining a way of dealing with instances in which the constraint graph has few large eigenvalues, following [Kol10], with a new way to deal with the complementary case of instances in which the constraint graph has several large eigenvalues. If the UGC is true, then the best evidence we can hope for in the short term is the discovery of integrality gaps for polynomial-time solvable relaxations, including the relaxations studied in [BBH$^+$12]. Moreover, the algorithmic work on restricted classes of instances, by identifying which instances are easy, can be used to guide the search for hard families of instances. This motivates our study of Unique Games on the Hypercube on which the first part of the thesis is based.

1.2 Unique Games on Hypercubes

In the first part of the thesis we initiate the study of Unique Games when the constraint graph is the Boolean Hypercube. As discussed above in order to make algorithmic progress on UGC, the next question to ask is what type of graphs are the ones on which known techniques have failed to yield provably good approximation in polynomial or in least $2^{n^{o(1)}}$-time. In this context the Boolean Hypercube turns out to be one such candidate family.

To understand why the Hypercube is a good candidate we need to understand what the guarantees achieved by the spectral algorithms of [ABS10, Kol10] depend upon. The basic idea behind these algorithms is to consider the space spanned by the top eigenvectors of the adjacency matrix and to do a brute force search in this space. In particular given an instance of Unique Games in which $1 - \epsilon$ fraction of edges are satisfiable the algorithm of [ABS10, Kol10] proves that we need to consider the space spanned by eigenvectors with eigenvalue roughly $\geq 1 - \epsilon^{\alpha}$ times the highest eigenvalue (λ_1)(α is an absolute constant here). If D is the dimension of the above space, then we can produce a solution in time roughly $\mathcal{O}(exp(kD))$. Therefore the running time of the algorithm depends exponentially on the dimension of the space of the *large* eigenvectors. As it turns out for the integrality gap instances proposed thus far for SDPs the dimension of this space is roughly $n^{\mathcal{O}(1)}$ which does not establish the full exponential hardness of UGC. Therefore we consider these instances to be *easy* for spectral methods.

The spectrum of the Boolean Hypercube on the other hand proves to be *hard* for the spectral methods in the following sense. It is well known that the number of eigenvalues $\geq (1-\epsilon^{\alpha})\lambda_1$ grows roughly like $n^{\epsilon^{\alpha}}$. Therefore the time taken by spectral algorithms is exponential which makes the Boolean Hypercube *hard* for existing methods.

The purpose of the first part of the thesis is to understand the behaviour of the SDP relaxations when the input constraint graph is the Hypercube. Since the Hypercube is a good representative of the *last frontier* instances on which the Unique Games problem is still not known to be easy, finding an efficient algorithm that solves UG on the Hypercube might give some motivation to suspect that UGC is false. On the other hand, constructing integrality gap instances in which the constraint graph is a Hypercube present certain unique difficulties, and requires constructions of a different nature from the ones that have been developed so far [KV05, RS09, BGH+11, KPS10].

An integrality gap instance for a relaxation of an optimization problem is an instance for which the optimum of the relaxation is $\geq 1-\epsilon$, while the optimum of the problem is $\leq 1-\epsilon'$, for some $\epsilon << \epsilon'$; in all the previous integrality gap instances of unique games (and also integrality gap instances of max cut and other constraint satisfaction problem), the feasible solution witnessing that the optimum of the relaxation is $\geq 1-\epsilon$ is constructed in such a way that every constraint contributes $\geq 1-\epsilon$ to the cost function. In a unique game on the Hypercube, however, if the instance is unsatisfiable then there is an unsatisfiable subset of just four constraints, since for an unsatisfiable instance there need to be four-cycles that are inconsistent. Thus, there cannot be a feasible solution for a relaxation in which every edge contributes more than $3/4$ to the cost function. Being forced to reason about non-symmetric solutions might give new ideas that could be applied in more general settings.

1.2.1 Our results

In the first part of this thesis, we consider the Unique Games problem over two alphabets (also refereed to as the Max-2-LIN($\mathbb{Z}_2$) problem). Recall that, by theorem 1.1.1, an improved approximation algorithm for Max-2-LIN($\mathbb{Z}_2$) for general instances would refute the UGC We study the approximability when restricted to instances whose constraint graph is an Hypercube.

We construct a family of integrality gap instances of Unique Games on the Hypercube constraint graph for the Goemans-Williamson semi-definite program (SDP). We note that unlike the case for all other known integrality gap instances, the state-of-the-art spectral algorithms are known to fail on the cube. We prove the following theorem

Theorem. *(**Main**) For every sufficiently small constant ϵ, and for every $d \geq d(\epsilon)$, there exists a Max-2-LIN($\mathbb{Z}_2$) instance on the Boolean cube Q_d of dimension d such that the UG combinatorial optimal value for that instance is $1 - \Omega(\epsilon)$, and the GW SDP optimal value is $1 - O(\epsilon^{3/2})$.*

Given the analysis of Goemans Williamson [GW95] it is known that given a Max-2-LIN($\mathbb{Z}_2$) instance with combinatorial optimum $1 - \epsilon$ the Goemans Williamson SDP has optimum $1 - \Omega(\epsilon^2)$ in the worst case. In this light the above theorem shows that the Goemans-Williamson SDP has an integrality gap on the Hypercube with a behaviour similar to the gap on general graphs.

Adding triangle inequalities. Adding so called *triangle inequalities* is a standard manipulation of semidefinite programs. We show that adding these constraints to the Goemans-Williamson SDP breaks the integrality gap of our instance. We conjecture that indeed the addition of triangle equalities is sufficient for the Hypercube. The motivation for the conjecture comes from the fact the Hypercube has a high denstiy of short cycles. We further build a connection with the performance of the augmented SDP with the multicommodity flow cut gap on a specific network on a graph. In particular we show that if the flow-cut gap on any family of graph for such a network is $\mathcal{O}(1)$ then in fact Max-2-LIN($\mathbb{Z}_2$) is easy for such a family. The connection described above in particular gives us for free the well known fact that efficient algorithms exist which solve the decision version of the Max-2-LIN($\mathbb{Z}_2$) problem on planar graphs.

1.3 The label extended graph and Graph Lifts

In the second part of the thesis we focus on the notion of random Graph Lifts. Our motivation for Graph Lifts comes from studying the Label Extended Graphs of a particular Unique Games instance. The Label Extended Graph of Unique Games instance on the graph $G = (V, E)$ and the set of constraints $\{\pi_{uv} : [k] \to [k]\}$ over k alphabets is an undirected graph H created by making k copies of every vertex v referred to as the fiber of v $fiber(v)$ and for every edge $(u, v) \in E$ setting up a matching between $fiber(u)$ and $fiber(v)$ corresponding to the permutation $\{\pi_{uv}\}$. We refer to the above constructed graph H as a $k-$lift of a graph and refer to k as a degree of the lift. The Label Extended Graphs have been the main focus of analysis in the spectral techniques for unique games in the works of [Kol10], [ABS10]. In particular the spectral algorithm proposed by [Kol10] shows that if the label extended graph is a relatively good expander(in the sense that it has very few *large* eigenvalues) then one can find the solution of the corresponding Unique Games by simply enumerating the subspace spanned by the first few eigenvectors. It is then a natural question to ask how does the spectrum of a graph lift behave and what kind of relationships can we establish between the spectrum of G and its k-lift H.

Although our study of lifts was motivated by the Unique Games Conjecture , lifts have been a central object of study in Spectral Graph Theory especially as a means to efficiently construct expander graphs. Informally, an expander is a graph where every small subset of the vertices has a relatively large edge boundary. Expander graphs have spawned research in pure and applied mathematics during the last several years, with several applications to multiple fields including complexity theory, the design of robust computer networks, the design of error-correcting codes, de-randomization of randomized algorithms, compressed sensing and the study of metric embeddings. For a comprehensive survey of expander graphs we refer the reader to [Sar06, HLW06].

Most applications are concerned with sparse d-regular graphs G, where the largest eigenvalue of the adjacency matrix A_G is d. The celebrated Cheeger's inequality (stated formally in section 2.4) builds a connection between the combinatorial notion of expansion and the eigenvalues of the adjacency matrix by relating expansion to the gap between the first eigenvalue (which is d also called the trivial eigenvalue) and the second eigenvalue(λ_2). Roughly, the smaller λ_2 is, the better the graph expansion. The Alon-Boppana bound ([Nil91]) proves that $\lambda_2 \geq 2\sqrt{d-1} - o(1)$. Thus graphs with $\lambda_2 \leq 2\sqrt{d-1}$ are optimal expanders and are called *Ramanujan Graphs*. Sometimes the definition of Ramanujan is made tighter by requiring that the maximum absolute eigenvalue of A_G, $\lambda \leq 2\sqrt{d-1}$. This is the definition we use.

A simple probabilistic argument shows the existence of infinite families of expander graphs [Pin73]. Although it is easy to construct Ramanujan graphs of a particular degree d with a small number of vertices(d-regular complete graphs and complete bipartite graphs), constructing an infinite family of such graphs with increasing size explicitly has proven to be a challenging and important task. This was first achieved by the work of Lubotzky, Phillips and Sarnak [LPS88] and Margulis [Mar88]. They built Ramanujan graphs from Cayley graphs. All of their graphs are regular, have degrees $p+1$ where p is a prime, and their proofs rely on deep number theoretic facts. In a recent breakthrough, Marcus, Spielman and Srivastava showed the existence of bipartite Ramanujan graphs of all degrees [MSS13]. A striking result of Friedman [Fri08] and a slightly weaker but more general result of Puder [Pud13], shows that almost every d -regular graph on n vertices is nearly Ramanujan i.e. it has $\lambda = 2\sqrt{d-1} + \mathcal{O}(1)$. It is still unknown whether the event that a random d-regular graph is exactly Ramanujan happens with constant probability. Despite the large body of work on the topic, all attempts to efficiently construct large Ramanujan expanders of any given degree have failed, and exhibiting such constructions remains an intriguing and important open problem.

A combinatorial approach to this problem, initiated by Friedman [Fri03], is to prove that one may obtain new (larger) Ramanujan graphs from smaller ones by taking random graph lifts. A random k-lift H of a graph is a k-lift of a graph G and where every constraint permutation π_{uv} for every edge (u, v) is chosen uniformly and independently. It is easy to see that a k-lift of a d-regular graph is a d-regular and that all of the eigenvalues of the original graph are also eigenvalues of a k-lift . These eigenvalues are referred to as the *old* eigenvalues and the other eigenvalues of H are referred to as the *new* eigenvalues . If λ is the absolute maximum of old eigenvalues except the first one (d), one hopes that the new eigenvalues are bounded roughly by λ. Note that this would imply that the k-lift of a Ramanujan graph is (roughly) Ramanujan. In this spirit there has been a large body of work investing the above problem in different scenarios which we describe next

1.3.1 Previous work

The work on lifts can braodly be divided into two categories. The first is when we consider lifts of very high degrees i.e. $k \to \infty$. Friedman [Fri03] proved that every new eigenvalue of H is $\mathcal{O}(d^{3/4})$ with high probability as $k \to \infty$, and conjectured a bound of $2\sqrt{d-1} + o(1)$, which would be tight (see, e.g. [Gre95]). Linial and Puder [LP10] improved Friedman' s bound to $\mathcal{O}(d^{2/3})$. Lubetzky, Sudakov and Vu [LSV11] showed that the absolute value of every nontrivial eigenvalue of the lift is $O(\lambda \log d)$ which improves on the previous results when G is significantly expanding, Adarrio-Berry and Griffiths [ABG10] further improved the bounds above by showing that every new eigenvalue H is $\mathcal{O}(\sqrt{d})$ and very recently, Puder [Pud13], proved the nearly-optimal bound of $2\sqrt{d-1}+1$. All those results hold with probability tending to 1 as $k \to \infty$, thus the degree k of the lift in question needs to be large.

The second category is the case when $k = 2$. Indeed if one is interested in explicitly constructing Ramanujan graphs using lifts, then one would need to de-randomize the above probabilistic results in some clever way. However, such a de-randomisation might be infeasible if one is looking at lifts of large degree k, where $k \to \infty$ and thus it is essential to look at lifts with low degrees. Bilu and Linial [BL06] were the first to study lifts of graphs with bounded degree $k = 2$ and suggested constructing Ramanujan graphs through a sequence of 2-lifts of a base graph. The idea is to start with a good small d-regular expander graph on some finite number of nodes (e.g. K_{d+1}) and perform a 2-lift operation thus doubling the size of the graph. If there is a way to preserve expansion after lifting, then repeating this operation will give large good expanders of the same bounded degree d. The authors in [BL06] showed that if the starting graph G is significantly expanding so that $|\lambda(G)| \leq \mathcal{O}(\sqrt{d \log d})$, then with high probability in the number of vertices of G, a random 2-lift of G has all its new eigenvalues upper bounded in absolute value by $\mathcal{O}(\sqrt{d \log^3 d})$. In the recent breakthrough work of Marcus, Spielman and Srivastava [MSS13], the authors showed that for every bipartite graph G, there exists a 2-lift of G, such that the new eigenvalues achieve the Ramanujan bound of $2\sqrt{d-1}$. The two results above indicate that understanding the expansion of typical bounded-degree lifts might be the right avenue towards constructing Ramanujan graphs of all degrees. To the best of our knowledge nearly no results are known in the regime when k is a small constant other than 2

1.4 Spectrum of constant degree lifts - Our results

In the second part of the thesis, we continue the study of the behaviour of 2-lifts. We significantly improve the guarantees achieved by Bilu-Linial [BL06] for random 2-lifts with high probability. In particular we show the following theorem

Theorem 1.4.1. *Let G be a d -regular graph with non-trivial eigenvalues at most λ in absolute value, and H be a (uniformly random) 2-lift of G. Let λ_{new} be the largest in absolute value new eigenvalue of H. Then*

$$\lambda_{new} \leq \mathcal{O}(\lambda)$$

with probability at least $1 - e^{-\Omega(n/d^2)}$.

Note that the above theorem gets rid of the multiplicative $\log^{3/2}(d)$ factor present in the Bilu-Linial result. The above theorem, in particular, implies that if we start with G being a small Ramanujan expander, then with high probability a random 2-lift will be almost Ramanujan, having all its non-trivial eigenvalues bounded by $\mathcal{O}(\sqrt{d})$. We note that unlike the case of lifts of degree $k \to \infty$, the dependency on λ is necessary for bounded k. This has previously been observed by the authors in [BL06] who gave the following example: Let G be a disconnected graph on n vertices that consists of $n/(d+1)$ copies of K_{d+1}, and let H be a random 2-lift of G. Then the largest non-trivial eigenvalue of G is $\lambda = d$ and it can be shown that with high probability, $\lambda_{new} = \lambda = d$. Therefore, our results are nearly tight.

We further initiate the study of graph lifts with constant degree $k > 2$. We focus our attention on a slightly restricted notion of lifts where the bijections π_{uv} for each edge (u, v) are chosen uniformly at random from the set of shift permutations on k elements. We call such lifts *shift* lifts and establish the following result about them

Theorem 1.4.2. *Let G be a d -regular graph with non-trivial eigenvalues at most λ in absolute value, and H be a random shift k-lift of G. Let λ_{new} be the be the largest in absolute value new eigenvalue of H. Then*

$$\lambda_{new} \leq \mathcal{O}(\lambda)$$

with probability at least $1 - k \cdot e^{-\Omega(n/d^2)}$.

One of the first steps in the proof of the result in Bilu-Linial [BL06] was the identification of the new eigenvalues as the eigenvalues of a particular *signed* adjacency matrix. We extend the notion of signed adjacency matrix to higher degree shift lifts and exhibit a bijection between the spectrum of shift k-lifts and the spectrum of certain k matrices. We believe this characterization might be of independent interest.

1.5 Organization of this thesis

The rest of the thesis is organised as follows. In Chapter 2 we provide the basic definitions and notations which would be used throughout the thesis. In addition we give a brief introduction to the basic concepts involved and state some well known theorems which would be used in our proofs. In Chapter 3 we present our results on Unique Games on the Hypercube. In particular we provide the full proof of Theorem 3.1.1. We also state and motivate our conjecture about the augmented SDP as well as establish connections to Multicommodity Flow-cut gaps in this chapter. Chapter 4 consists of our results on Graph Lifts. We provide the full proofs of Theorems 1.4.1 and 1.4.2 including the proofs of the required lemmas.

Chapter 2

Preliminaries

In this section we provide the basic notations and definitions that would be used in the rest of the thesis. In addition we introduce the basic concepts and list relevant known theorems and lemmas required by the proofs of theorems in the rest of the thesis. For most standard results we refer the reader to a text describing the proof of the theorem in detail. For certain simple lemmas we provide the proof here itself for completeness.

2.1 Basic Notations and Definitions

Graphs Let $G = (V, E)$ be a graph with vertex set V, $|V| = n$ and edge set E. We will denote with A_G be the adjacency matrix of the graph (if the graph is clear from the context we will shorten this notation to just A). Let $\lambda_1 \geq \lambda_2 \geq \dots \lambda_n$ be n eigenvalues of A. Note that since A is a real, symmetric matrix its eigenvalues are also real. Moreover if G is regular with degree d it is easy to see that $\lambda_1 = d$. We refer to d as the *trivial* eigenvalue and the rest as the *non-trivial* eigenvalues . We define λ to be the maximum non-trivial eigenvalue in absolute value, i.e.

$$\lambda = \max_{i:[2,n]} |\lambda_i|$$

Through most of the paper (especially the section on lifts) G will be a d-regular graph and we will be concerned with eigenvalues of adjacency matrices. Let $Spec(G)$ denote the spectral gap of G defined as $Spec(G) = \lambda_1 - \lambda_2 = d - \lambda_2$. For any two subsets $S, T \subseteq V$ let $E(S, T)$ be the number of edges that go from S to T. For a subset $S \in V$ let $N_G(S)$ denote the neighbourhood of S, i.e. $N_G(S) = \{v \in V | \exists u \in S,\ (u, v) \in E\}$.

A word on the notation. For improving readability capital letters will generally be reserved for sets, small letters for scalars and boldface for vectors. Let $\mathbf{v}(i)$ be the i^{th} coordinate of the vector $\mathbf{v}$. For a vector $\mathbf{x}$ the set $S(\mathbf{x})$ denotes its support, i.e. $\{i | \mathbf{v}(i) \neq 0\}$.

Norms Since we will concern ourselves only with the ℓ_2 norm for vectors and the spectral norm for matrices, let $||\mathbf{x}||$ denote the ℓ_2-norm of a vector $\mathbf{x}$ and let $||M||$ denote the spectral norm of a matrix M. To remind the reader the spectral norm of a matrix is $\max_i |\lambda_i|$. Another very useful characterization of the spectral norm comes from Rayleigh coefficients. Indeed for an $n \times n$ matrix M

$$||M|| = \max_{\mathbf{x} \in \mathbb{C}^n} \frac{|x^* M x|}{x^* x}$$

In case the matrix is real and symmetric this condition can be relaxed to the following

$$||M|| = \max_{\mathbf{x} \in \mathbb{R}^n} \frac{|x^T M x|}{x^T x}$$

We use the above characterization heavily in our proofs. Next we define two operations on vectors

Definition 2.1.1 (Tensor Product)**.** *Given two vectors* $\mathbf{x} \in \mathbb{R}^{n_1}, \mathbf{y} \in \mathbb{R}^{n_2}$ *define the vector* $\mathbf{x} \otimes \mathbf{y} \in \{0,1\}^{n_1 * n_2}$ *as follows*

$$\mathbf{x} \otimes \mathbf{y}(i * n_1 + j) = \mathbf{x}(i)\mathbf{y}(j)$$

It is immediate from the definition that the norm of tensor products is multiplicative,

- $\|\mathbf{x} \otimes \mathbf{y}\|^2 = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$,

Definition 2.1.2 (Diadic Decomposition)**.** *Given a vector* $\mathbf{x} \in [0, \pm 1/2, \pm 1/4 \ldots]^n$ *we define the* diadic decomposition *of* $\mathbf{x}$ *as the set* $\{2^{-i}\mathbf{u_i}\}$ *where each* $\mathbf{u_i} \in [-1, 0, 1]^n$ *is a vector defined as*

$$\mathbf{u_i}(j) = \begin{cases} 1, & \textit{if } x(j) = 2^{-i} \\ -1, & \textit{if } x(j) = -2^{-i} \\ 0, & \textit{otherwise} \end{cases}$$

In Chapter 4 of the thesis we extensively use the standard Chernoff-Hoeffding [Hoe62] bound to estimate the sum of independent random variables. For clarity we state the precise form of Chernoff-Hoeffding bound which we will use below

Theorem 2.1.3 (Chernoff-Hoeffding)**.** *Let* $X_1, X_2 \ldots X_n$ *be independent random variables such that* $Pr(X_i \in [a_i, b_i]) = 1$*. Let* $S = X_1 + X_2 \ldots X_n$*. Then the following holds*

$$Pr(|S - E(S)| \geq t) \leq 2exp\left(-\frac{2t^2}{\sum_i (b_i - a_i)^2}\right)$$

2.1.1 The Boolean Hypercube

Let Q_d be a Boolean Hypercube graph (V_d, E_d) of dimension d. In the context of Hypercubes we will generally denote the number of dimensions by d. This notation is consistent with our notation of d representing the degree. To remind the reader the Boolean Hypercube of dimension d is a graph defined on a set of vertices $v = \{0,1\}^d$. Therefore with every vertex v we can naturally associate a vector $\mathbf{v}$. The edge set E_d is defined as the set of vertices (u, v) such that the corresponding vectors $\mathbf{u}, \mathbf{v}$ differ in exactly one coordinate. For any edge $(u, v) \in E_d$ let $i(\mathbf{u}, \mathbf{v})$ be the coordinate along which the corresponding vectors differ.

We denote by $H(\mathbf{v})$ the hamming weight of the vector $\mathbf{v}$, i.e. the number of $1'$s in $\mathbf{v}$. The Hypercube can therefore be partitioned into layers $L_0 \to L_d$ where $L_i = \{v \in V_d | H(\mathbf{v}) = i\}$. The Hypercube is a highly recursive structure in the sense that for any integer $i \leq d$, it can be viewed as an i-dimensional Hypercube where each vertex is a representative of a $d - i$ dimensional Hypercube. We will appeal to this view quite often in our proofs. Following this view define $Q_{d-k}(\mathbf{x})$ where $\mathbf{x} \in \{0,1\}^k$ be a $d - k$ dimensional sub-cube of Q_d obtained by fixing the first k coordinates to be $\mathbf{x}$.

2.1.2 Fourier Analysis of Boolean Function

A function f defined on the vertices of the Hypercube with range $\{0,1\}$, i.e $f : \{0,1\}^d \to \{0,1\}$ is called a Boolean Function. Note that a Boolean Function is equivalent a subset of vertices in the Hypercube. It turns out in the analysis of Boolean Functions that it is more useful to look at Boolean Functions with the following equivalent but shifted domain-range space $f : \{-1,1\}^d \to \{-1,1\}$. For any subset $S \in [n]$ and for any input $\mathbf{x} \in \{-1,1\}^d$ let

$$\mathbf{x}^S = \prod_{i \in S} \mathbf{x}_i$$

The following is well known and a key fact that leads to an extremely useful of change of basis for analysing Boolean Function

Theorem 2.1.4. *Every function $f : \{-1,1\}^d \to \{-1,1\}$ can be uniquely expressed as the following multilinear polynomial*

$$f(\mathbf{x}) = \sum_{S \subseteq [d]} \widehat{f(S)} \mathbf{x}^S$$

The above expression is referred to as the Fourier Expansion of f and the corresponding coefficients $\widehat{f(S)}$ are referred to as the Fourier coefficients of f. The above change of basis has turned out to be an extremely powerful method to analyse Boolean Functions leading to fundamental results in the field of Hardness of Approximation, Social Choice etc. We refer the reader to an upcoming book by Ryan O Donnell for a comprehensive survey of the Analysis of Boolean Functions [O'D14].

The inner product under the Fourier basis is defined as follows

Definition 2.1.5. *Let $f, g : \{-1,1\}^d \to \{-1,1\}$ be two Boolean functions. The inner product of the two functions is* *defined as*

$$< f, g >= E[f(\mathbf{x})g(\mathbf{x})] = \sum_{S \subseteq [d]} \widehat{f(S)}\widehat{g(S)}$$

where the expectation is over random choices of $\mathbf{x} \in \{-1, 1\}^d$.

An important quantity that has been studied in great detail in the Analysis of Boolean Functions is the notion of the Influence of a Boolean Function. We first define the Influence along the i^{th} dimension. For any vector $\mathbf{x} \in \{-1, 1\}^d$, let $\mathbf{x}^{\oplus i}$ be the vector $\mathbf{x}$ flipped in the i^{th} coordinate.

Definition 2.1.6 (Influence)**.** *Let* $\mathbf{x}$ *be a random vector uniformly chosen from the set* $\{-1, 1\}^n$*, then given any function* $f : \{-1, 1\}^d \rightarrow \{-1, 1\}$ *the influence along the* i^{th} *coordinate* $\mathbb{I}_i(f)$ *is defined as*

$$\mathbb{I}_i(f) = Pr(f(\mathbf{x}) \neq f(\mathbf{x}^{\oplus i}))$$

Definition 2.1.7 (Total Influence)**.** *Given any function* $f : \{-1, 1\}^d \rightarrow \{-1, 1\}$ *the total influence* $\mathbb{I}(f)$ *is defined as*

$$\mathbb{I}(f) = \sum_i \mathbb{I}_i(f)$$

Following from the equivalence between a Boolean Function f and a subset of vertices of the Hypercube $S_f \subseteq V_d$, it is not hard to see that the total influence is related to the edge boundary of S by the following equality

$$\mathbb{I}(f) = \frac{2E[S_f, V_d - S_f]}{2^d}$$

Certain important Boolean Functions have been studied in great detail and one of them with which we would be concerned in this work is the Minority Function $Min_d : \{-1, 1\}^d \rightarrow \{-1, 1\}$. As suggested by the name the function simply returns the value which is in minority in its input.

Following are some well known facts about the Fourier coefficients of Boolean Functions which we use in the thesis and state here without proof. The reader is referred to [O'D14] for proof of these facts.

- $|\widehat{Min_d}(\{i\})| \sim \sqrt{\frac{2}{\pi d}}$
- $\sum_{|S| \geq 2} \widehat{Min_d}(S)^2 \leq (1 - \frac{2}{\pi})$
- $\mathbb{I}(Min_d) \sim \sqrt{\frac{2}{\pi} d}$
- $\sum_i |\widehat{f}(\{i\})| \leq \mathbb{I}(f)$

2.2 Unique Games and Max-LIN definitions

Following is a formal definition of the Unique Games problem.

Definition 2.2.1 (Unique Games)**.** *A Unique Games instance for alphabet size k is specified by an undirected constraint graph $G = (V, E)$, a set of variables $\{x_u\}_{u \in V}$, one for each vertex u, and a set of permutations (constraints) $\pi_{uv} : [k] \to [k]$, one for each (u, v) such that $\{u, v\} \in E$, with $\pi_{uv} = (\pi_{vu})^{-1}$. An assignment of values in $[k]$ to the variables is said to satisfy the constraint on the edge $\{u, v\}$ if $\pi_{uv}(x_u) = x_v$. The optimization problem is to assign a value in $[k]$ to each variable x_u so as to maximize the number of satisfied constraints. We define the value of a solution to be the fraction of constraints that are not satisfied by this solution.*

An optimal solution for a Unique Games instance which satisfies the maximum number of constraints will also be referred to as the *combinatorial solution* and its value will be referred to as the *combinatorial value*. We note that, while it is slightly more common to define the value of a solution to be the fraction of satisfied constraints by the solution, in this paper we find it more convenient to define the value of a solution as the fraction of constraints that is not satisfied by it.

Next we define a special case of the Unique Games problem which is known to be as hard as the Unique Games problem itself (ref [KKMO04]).

Definition 2.2.2 (Max-2-LIN($\mathbb{Z}_q$))**.** *A Max-2-LIN($\mathbb{Z}_q$) instance is a Unique Games instance over alphabet size q where the constraints are of the form $x_u - x_v = c_i \ mod\ q$ and $c_i \in [q]$. Note that to specify a Max-2-LIN($\mathbb{Z}_q$) instance it is sufficient to specify a graph $G = (V, E)$ and a function $f : E \to [k]$.*

The main focus of the thesis would be Max-2-LIN($\mathbb{Z}_2$) instances. Note that they are equivalent to Unique Games over two alphabets. To specify a Max-2-LIN($\mathbb{Z}_2$) instance it is enough to specify a function $f : E \to \{0, 1\}$. We call the edges with the value 0 “equality” edges and the edges with value 1 “inequality” edges in accordance to the constraints implying whether the two labels on the edge should be equal or not. Since we are concerned with instances where the constraint graphs are Hypercubes, we will be viewing a Max-2-LIN($\mathbb{Z}_2$) instance as a function $I : E_d \to \{0, 1\}$.

2.3 Semi-definite programming relaxations for Unique Games

In this section we describe the semidefinite programming relaxations that have been considered for Unique Games and the approximation guarantees they are known to achieve. We first consider the case of Unique Games over two alphabets. Goemans-Williamson [GW95] proposed the following semi-definite program for solving Max-2-LIN($\mathbb{Z}_2$) .

Note that the original paper by Goemans and Williamson defines the SDP for the Max-Cut problem which is essentially a Max-2-LIN($\mathbb{Z}_2$) problem with all edges being inequality edges.

Definition 2.3.1 (GW SDP)**.** *Given a graph $G = (V, E)$ and a Max-2-LIN($\mathbb{Z}_2$) instance $I : E \to \{0, 1\}$ on it, let the set of equality edges be E^+ and the set of inequality edges be E^-. The Goemans-Williamson SDP for the instance is defined as*

$$\textit{minimize } \frac{1}{4|E|}\left(\sum_{(u,v)\in E^+} \|x_u - x_v\|^2 + \sum_{(u,v)\in E^-} \|x_u + x_v\|^2\right)$$

$$\textit{subject to} \quad \|x_u\|^2 = 1 \ \ (\forall\, u \in V)$$

The analysis carried out by Goemans-Williamson for Max-Cut essentially holds for Max-2-LIN($\mathbb{Z}_2$) too. In particular they prove the following theorem

Theorem 2.3.2 (Goemans Williamson [GW95])**.** *Given any Unique Games instance over 2 alphabets such that any labelling unsatisfies at least an ϵ fraction of edges then the GW SDP has its optimum lower bounded by $\Omega(\epsilon^2)$ for sufficiently small ϵ.*

By Theorem 1.1.1 this is indeed tight upto constants under the Unique Games Conjecture. The GW SDP is also known to have an integrality gap which makes the above mentioned theorem tight ([FS02]). For ease of exposition we define the following notation to help us state the combinatorial and SDP optimums of a Unique Games optimum.

Definition 2.3.3 ((α, β)-gap instance)**.** *An infinite family $\mathcal{F}$ of Max-2-LIN($\mathbb{Z}_2$) instances is called* an (α, β)-gap Instance *for Max-2-LIN($\mathbb{Z}_2$) if the combinatorial optimum on any $I \in \mathcal{F}$ is $\Omega(\alpha)$ and the GW SDP has optimum value $\mathcal{O}(\beta)$.*

In particular Theorem 2.3.2 rules out the existence of (ϵ, ϵ^t)-gap Instance for any $t > 2$.

For Unique Games over larger alphabets SDP relaxations have been considered in various different works coupled with rounding techniques that achieve different guarantees depending upon n the number of vertices and k, the number of alphabets. We summarize these results in the following table.

SDP Algorithm	Approximation Guarantee for OPT=$1-\epsilon$
Khot [Kho02]	$1 - \mathcal{O}(k^2\epsilon^{1/5}\sqrt{\log(1/\epsilon)})$
Trevisan [Tre05]	$1 - \mathcal{O}(\sqrt{\epsilon\log(n)})$
Charikar et al [CMM06a]	$\Omega(k^{-\epsilon/(2-\epsilon)})$ $1 - \mathcal{O}(\sqrt{\epsilon\log(k)})$

Following is the standard generalized SDP relaxation for Unique Games over k alphabets.

Definition 2.3.4 (UG SDP)**.** *Given a graph $G = (V, E)$ and a Unique Games instance $\{\pi_{uv} | u, v \in E,\ \pi_{uv} : [k] \to [k]\}$ on it, the generalized SDP for the instance is defined as*

$$\textit{minimize } \frac{1}{2|E|} \sum_{(u,v)\in E} \sum_{i=1}^{k} \|u_i - v_{\pi_{uv}(i)}\|^2$$

subject to

$$u_i \cdot u_j = 0 \ \ (\forall\, u \in V\ \forall\, i \neq j \in [k])$$

$$\sum_{i=1}^{k} \|u_i\|^2 = 1 \ \ (\forall\, u \in V)$$

Charikar et al. [CMM06a] augmented the above mentioned SDP with the following triangle inequalities

$$u_i \cdot v_j \geq 0 \qquad (\forall (u,v) \in E\ i, j \in [k])$$

$$\|u_i\|^2 \geq u_i \cdot v_{\pi_{uv}(i)} \geq 0 \qquad (\forall (u,v) \in E\ i \in [k])$$

They prove that given an instance with SDP optimum ϵ one can guarantee the existence of a combinatorial solution with value $\mathcal{O}(\sqrt{\epsilon \log(k)})$ as well as a combinatorial solution with value $\mathcal{O}(k^{-\epsilon/2-\epsilon})$. The well known integrality gap construction of Khot-Vishnoi [KV05] shows a nearly tight integrality gap of $\mathcal{O}(k^{-\epsilon/9})$. The following theorem of Khot et al. [KKMO04] shows that the results obtained by the SDP relaxations of Charikar et al. are indeed optimal assuming the Unique Games Conjecture

Theorem 2.3.5. *Assume the Unique Games Conjecture. Then for all sufficiently small $\epsilon > 0$ and $k > k(\epsilon)$, it is NP-hard to distinguish instances of Unique Games Conjecture over k alphabets that are at least $(1-\epsilon)$-satisfiable from instances that are at most $(1 - (2/\pi)\sqrt{\epsilon \log(k)})$-satisfiable.*

Theorem 2.3.6. *Assume the Unique Games Conjecture. Then for all sufficiently small $\epsilon > 0$ and $k > k(\epsilon)$, it is NP-hard to distinguish instances of Unique Games Conjecture over k alphabets that are at least $(1-\epsilon)$-satisfiable from instances that are at most $(1 - k^{\epsilon/2-\epsilon})$-satisfiable.*

2.4 Spectral Graph Theory Basics

In this section we give formal definitions of expansion and state some standard results connecting the combinatorial and algebraic notions of expansion.

There are many ways to characterize Expander Graphs, the most common among them being the combinatorial and the algebraic notions of expansion. The combinatorial notion measures the edge boundary of a subset of vertices in the graph. Formally, given a graph $G = (V, E)$ the expansion of the graph $H(G)$ is defined as

Definition 2.4.1 (Combinatorial Expansion)**.**

$$\mathcal{H}(G) = \min_{S \subset V : |S| \leq |V|/2} \frac{E(S, V \backslash S)}{|S|}$$

Since G is a d-regular graph $\mathcal{H}(G) \leq d$. Another way to characterize expansion is via the spectral gap $Spec(G) = d - \lambda_2$, which is referred to as the algebraic expansion of the graph. The following fundamental fact known as the Cheeger's Inequality gives a robust connection between the two notions of expansion above.

Definition 2.4.2 (Cheeger's Inequality)**.**

$$\frac{d - \lambda_2}{2} \leq \mathcal{H}(G) \leq \sqrt{d(d - \lambda_2)}$$

Expanders are also sometimes seen as graphs which are close to random graphs. This idea is quantified by the following well-known fact known as the Expander Mixing Lemma which bounds the deviation between the number of edges between two subsets and the expected number in a random graph.

Theorem 2.4.3 (Expander-Mixing Lemma)**.**

$$(\forall S, T \subseteq V) \ |E(S, T) - \frac{d|S||T|}{n}| \leq \lambda \sqrt{|S||T|}$$

Bilu-Linial [BL06] in their work on lifts showed that the converse of the above statement is almost true as well.

Theorem 2.4.4 (Converse of Expander Mixing Lemma)**.** *Given a graph such that forall* $S, T \subseteq V$

$$|E(S, T) - \frac{d|S||T|}{n}| \leq \alpha \sqrt{|S||T|}$$

Then $\lambda = \mathcal{O}(\alpha(1 + \log(d/\alpha)))$

2.5 Lifts - Basic Definitions and Notations

In this section we formally define k-lifts of graphs and state some of their properties. A k-lift of graph corresponds to a set of permutations $\Pi = \{\pi_{u,v}\}$ which is indexed over the set of edges $E = \{(u, v)\}$, where each $\pi_{u,v} : [k] \to [k]$.

Definition 2.5.1 (k-lift)**.** *Given a graph* $G = (V, E)$ *a k-lift of the graph corresponding to a set of permutations* Π *is* *defined as a graph $H = (V \times [k], E')$ where*

$$E' = \{((x,i),(y,j))|(x,y) \in E, \pi_{x,y}(i) = j\}$$

For every vertex $x \in V$, we define the fiber of x as $fiber(x) = \{x\} \times [k]$. Also let A_H denote the adjacency matrix of H. k would be referred to as the degree of the lift and G is referred to as the base graph.

When the set of permutations Π is chosen randomly (indepenently and uniformly for each edge) the corresponding lift is referred to as a random $k - lift$.

Some initial easy observations can be made about the structure of a $k - lift$. A $k - lift$ is also regular with the same degree as the base graph. Also it is easy to see that $\mathcal{H}(H) \leq \mathcal{H}(G)$ by simply considering the set $S \times [k]$ for each subset $S \subseteq V$ of the original graph. It is easy to see that the eigenvalues of A_G are also eigenvalues of A_H. Therefore we call the n eigenvalues of A_G the *old* eigenvalues and $n(k-1)$ other eigenvalues of A_H the *new* eigenvalues. We will denote by λ_{new} the largest in absolute value new eigenvalue of H, which we also refer to as "first" new eigenvalue for simplicity.

We now define the notion of signing of the adjacency matrix of a graph

Definition 2.5.2 (Signing)**.** *Given an $n \times n$ adjacency matrix A, an $n \times n$ symmetric matrix A_s is a signing of A if for all (i,j) such that $A(i,j) = 1$, $A_s(i,j) \in \{-1, 1\}$ and for all (i,j) such that $A(i,j) = 0$, $A_s(i,j) = 0$.*

An arbitrary signing of A_s is obtained by choosing an arbitrary sign for each edge in A. It is easy to see that there is a simple bijection between 2-lifts and signings, i.e. for every edge there are two permutations to choose from which corresponds to the sign chosen in the signing.

A crucial property of signings observed by Bilu-Linial [BL06] which makes the study of new eigenvalues of a $2 - lift$ convenient is that the new eigenvalues of the lift are exactly the eigenvalues of the signing A_s. To see this first note that the adjacency matrix of a two lift can be written as

$$A_H = 1/2 * \begin{bmatrix} A + A_s & A - A_s \\ A - A_s & A + A_s \end{bmatrix}$$

Now consider any eigenvector v of A_G with eigenvalue α. It is easy to see that $[v, v]$ is an eigenvector of A_H with the same eigenvalue α. This is the set of old eigenvalues. Now consider any eigenvector u of A_s with eigenvalue β. It is easy to see that $[u, -u]$ is an eigenvector of A_H with the same eigenvalue β. Since these are orthogonal eigenvectors we see that the spectrum of A_s is precisely the set of new eigenvalues. Therefore $\lambda_{new} = ||A_s|| = \max_i |\lambda_i(A_s)|$. In this context Bilu and Linial [BL06] made the following conjecture about signings which is still open in its full generality.

Conjecture 2 (Bilu-Linial Conjecture)**.** *For every adjacency matrix A of a d-regular graph there exists a signing A_s such that* $||A_s|| \leq 2\sqrt{d-1}$

2.6 Shift Lifts

In this section section we discuss the notion of Shift Lifts. In addition we give a novel characterization of the *new* eigenvalues of these special class of lifts. The characterization can be seen as an extension of signings to this case.

Definition 2.6.1 (Shift-k-Lift)**.** *A shift k-lift of a graph is a k-lift such that the associated set of permutations Π is such that for all $\pi_{u,v} \in \Pi, \exists s \in [k]$ such that $\pi_{u,v}(i) = (i+s) \mod k$. That is every permutation is a cyclic shift. We denote by $Shift(u,v)$ the "magnitude" of the shift along the edge (u,v). i.e. if $\pi_{u,v}(i) = (i+s) \mod k$, then $Shift(u,v) = s$. Note that (i,j) here is an ordered pair and $Shift(v,u) = -Shift(u,v) \mod k$.*

One of the major reasons that made the study of the new eigenvalues of a 2-lift of a graph G easier, was the ability to characterize its new eigenvalues as eigenvalues of the signed adjacency matrix. This leads to the question of whether such characterization can be extended to k-lifts in general.

A natural avenue towards the characterizing the eigenvalues of such lifts is to look at the roots of unity and for each edge (u,v), assign the value $\omega^{Shift(u,v)}$. Here ω is the k^{th} root of unity. This intuition indeed works. For any given shift k-lift instance, define the following family of Hermitian matrices $A_s(t)$ parameterized by t where t is the k^{th} root of unity.

$$A_s(t)(i,j) = \begin{cases} 0, & \text{if } A(i,j) = 0 \\ t^{Shift(i,j)}, & \text{if } A(i,j) = 1 \end{cases}$$

Theorem 2.6.2. *Let $G(E,V)$ be a graph and H any shift k-lift of G, with the corresponding shifts given by the set $\{Shift(i,j)\}_{(i,j)\in E}$. Let ω be a k^{th} root of unity. Let $\mathbf{v}$ be an eigenvector of the matrix $A_s(\omega)$ above, with eigenvalue α. Then*

$$\mathbf{v}^l(\omega) = [\mathbf{v}, \omega * \mathbf{v}, \omega^2 * \mathbf{v} \ldots \omega^{k-1}\mathbf{v}]$$

is an eigenvector of the adjacency matrix of H with eigenvalue α. Moreover, all eigenvectors created this way using different roots of unity are orthogonal.

Proof: Let A_H be the adjacency matrix of H. Consider the vector $\mathbf{v}^l(\omega)$. Since $\mathbf{v}^l(\omega)$ is a $1 \times kn$ dimensional vector, we will refer to its coordinates as a tuple (x,i) where $x \in [n]$ and $i \in [k]$. Essentially, (x,i) corresponds to the i^{th}

vertex in the fiber of the x^{th} vertex in the original graph. Note that $\mathbf{v}^l(\omega)(x,i) = \omega'^i\mathbf{v}(x)$ Consider the term

$$\begin{aligned} A_H\mathbf{v}^l(\omega)(x,i) &= \sum_{y:(x,y)\in E(G)} \omega^i\omega^{Shift(x,y)}\mathbf{v}(y) \\ &= \omega^i \sum_{y:(x,y)\in E(G)} \omega^{Shift(x,y)}\mathbf{v}(y) \\ &= \alpha\omega^i\mathbf{v}(x) \\ &= \alpha\mathbf{v}^l(\omega)(x,i) \end{aligned}$$

Also note that for any two vectors $\mathbf{v}_1^l(\omega), \mathbf{v}_2^l(\omega')$, $\langle \mathbf{v}_1^l(\omega), \mathbf{v}_2^l(\omega')\rangle = \langle \mathbf{v}_1, \mathbf{v}_2\rangle(1+\beta+\beta^2\ldots)$ where $\beta = \omega^*\omega'$. Note that if $\omega \neq \omega'$, $(1+\beta+\beta^2\ldots) = 0$, otherwise since x, y are orthogonal eigenvectors corresponding to $A_s(\omega)$, therefore $\langle x, y\rangle = 0$ □

2.7 Other Lemmas and proofs

In this section we collect a few simple combinatorial identities and lemmas that are used at various points in the rest of the thesis. The proofs of these lemmas are provided for completeness. We suggest the reader to skip this section for now and read it accordingly when the lemmas are used in the main body of the thesis

Lemma 2.7.1 (Discretization Lemma)**.** *For any $\mathbf{x} \in \mathbb{R}^n$,$||\mathbf{x}||_\infty \leq 1/2$ and M such that the diagonal entries of M are 0, there exists $\mathbf{y} \in \{\pm 1/2, \pm 1/4, \ldots\}^n$ such that $|\mathbf{x}^T M\mathbf{x}| \leq |\mathbf{y}^T M\mathbf{y}|$ and $\|\mathbf{y}\|^2 \leq 4 * \|\mathbf{x}\|^2$. Moreover, each entry of $\mathbf{x}$ between $\pm 2^{-i}$ and $\pm 2^{-i-1}$ is rounded to either $\pm 2^{-i}$ or $\pm 2^{-i-1}$.*

Similarly, for any $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$, $||\mathbf{x}_1||_\infty, ||\mathbf{x}_2||_\infty \leq 1/2$, there exists $\mathbf{y}_1, \mathbf{y}_2 \in \{\pm 1/2, \pm 1/4, \ldots\}^n$ such that $|\mathbf{x}_1^T M\mathbf{x}_2| \leq |\mathbf{y}_1^T M\mathbf{y}_2|$,$\|\mathbf{y}_1\|^2 \leq 4 * \|\mathbf{x}_1\|^2$,$\|\mathbf{y}_2\|^2 \leq 4 * \|\mathbf{x}_2\|^2$ and each entry of $\mathbf{x}_1, \mathbf{x}_2$ between 2^{-i} and 2^{-i-1} is rounded to either 2^{-i} or 2^{-i-1}.

Proof:[Proof of Lemma 2.7.1] To obtain such a vector $\mathbf{y}$ we simply take a vector $\mathbf{x}$ and round its coordinates independently with the following probabilistic rule. Let $\mathbf{x}(i) = \pm(1+\delta_i)2^{-i}$ be the i^{th} coordinate of $\mathbf{x}$. We round $\mathbf{x}(i)$ to $sign(\mathbf{x}(i)) \cdot 2^{-i+1}$ with probability δ_i and $sign(\mathbf{x}(i)) \cdot 2^{-i}$ with probability $1-\delta_i$. Let the rounded vector be $\mathbf{x}'$. Note that $E[\mathbf{x}(i')] = x(i)$. Now since each coordinate is rounded independently and the diagonal entries of M are 0, we get that $E[\mathbf{x}'^T M\mathbf{x}'] = \mathbf{x}^T M\mathbf{x}$. This implies there exists a $\mathbf{y} \in \{\pm 1/2, \pm 1/4, \ldots\}^n$ that can be generated by this rounding such that $|\mathbf{x}^T M\mathbf{x}| \leq |\mathbf{y}^T M\mathbf{y}|$. Also it is easy to see that $\|\mathbf{y}\|^2 \leq 4 * \|\mathbf{x}\|^2$ and by definition in $\mathbf{y}$ every coordinate value between $\pm 2^{-i}$ and $\pm 2^{-i-1}$ is rounded to either $\pm 2^{-i}$ or $\pm 2^{-i-1}$. The proof of the second part of the lemma is the same as the first part. Here we obtain $\mathbf{x}_1'$ and $\mathbf{x}_2'$ by the same procedure and follow the exact same argument to get $\mathbf{y}_1$ and $\mathbf{y}_2$ □

Lemma 2.7.2. *Assuming that $r^t \leq z/2$, $r \geq 2$, $x > 0$, we have the following inequality:*

$$\sum_{i=0}^{i=t} (r^i \log(z/r^i))^x \leq c(r)(r^t \log(z/r^t))^x$$

where $c(r)$ is a constant depending only on r.

Proof:[Proof of Lemma 2.7.2] For all i define $a_i = (r^i \log(z/r^i))^x$. Lets consider the ratio of consecutive terms a_{i+1}/a_i for $i \in [0, t-1]$.

$$\begin{aligned}
\frac{a_{i+1}}{a_i} &= \left(\frac{r^{i+1}\log(z/r^{i+1})}{r^i \log(z/r^i)}\right)^x \\
&= \left(r\left(1 - \frac{\log(r)}{\log(z) - i\log(r)}\right)\right)^x \\
&\geq \left(r\left(1 - \frac{\log(r)}{1 + (t-i)\log(r)}\right)\right)^x && (r^t \leq z/2)
\end{aligned}$$

If $i \leq t-2$ we get that $a_{i+1}/a_i \geq r^x\left(\frac{1+\log(r)}{1+2\log(r)}\right)^x = \alpha(r)$. It is easy to see that $\alpha(r) > 1$ for $r \geq 2$. Also for $i = t-1$ we get that $a_{i+1}/a_i \geq (r/(1+\log(r)))^x \geq 1$.

Now consider the sum S_{-1} defined as

$$\begin{aligned}
&S_{-1} = a_0 + a_1 + \ldots + a_{t-1} \\
&\Rightarrow \alpha(r)S_{-1} = \alpha(r)(a_0 + a_1 + \ldots + a_{t-1}) \\
&\Rightarrow (\alpha(r)-1)S_{-1} = -a_0 + (\alpha(r)a_0 - a_1) + (\alpha(r)a_1 - a_2)\ldots + a_{t-1}\alpha(r) \\
&\Rightarrow (\alpha(r)-1)S_{-1} \leq a_{t-1}\alpha(r) && (a_{i+1} \geq \alpha(r)a_i) \\
&\Rightarrow S_{-1} \leq a_{t-1}\left(\frac{\alpha(r)}{\alpha(r)-1}\right)
\end{aligned}$$

Therefore

$$\sum_{i\in[t]} a_i \leq S{-1} + a_t \leq \left(1 + \left(\frac{\alpha(r)}{\alpha(r)-1}\right)\right) a_t$$

.

Setting $c(r) = \left(1 + \left(\frac{\alpha(r)}{\alpha(r)-1}\right)\right)$ we get the required result. □

Chapter 3

Unique Games on Hypercubes

In this chapter we study the Max-2-LIN($\mathbb{Z}_2$) problem on the Boolean hypercube. We first present the construction of a Max-2-LIN($\mathbb{Z}_2$) instance on the Hypercube which has a large integrality gap on the Goemans Williamson SDP (see Def. 2.3.1). We further show an augmentation of the GW SDP and conjecture that the augmentation indeed works for the hypercube. We motivate our conjecture by showing that our integrality gap no more remains an integrality gap of the augmented SDP. We further build connections to Multicommodity Flows and show that $\mathcal{O}(1)$ Multicommodity Flow-Cut gaps on certain networks on families of graphs is sufficient for the integrality gap of augmented SDP to be $\mathcal{O}(1)$.

3.1 Main Theorem and Proof

We begin by proving the following main theorem

Theorem 3.1.1. *(Main) For every sufficiently small constant ϵ there is an $(\epsilon, \epsilon^{3/2})$ Instance(see Def 2.3.3) on the Hypercube for Max-2-LIN($\mathbb{Z}_2$) .*

Note that the above theorem establishes that there exists a family of Max-2-LIN($\mathbb{Z}_2$) instances on the Hypercube of combinatorial value (note that we are using the minimization objective) $\Omega(\epsilon)$ but the GW SDP has optimum $O(\epsilon^{3/2})$.

Proof Overview We construct our gap instance by starting with an instance for which all edges are equality edges and converting a small number edges to inequalities ensuring that the assignment that assigns the same value to all vertices(say 1) is still roughly the combinatorial optimum assignment, while at the same time the SDP optimum decreases. More concretely we show the following lemma :

Lemma 3.1.2. *(Main Lemma) For every sufficiently large d, there exists a Max-2-LIN($\mathbb{Z}_2$) instance defined over the d dimensional hypercube, whose combinatorial value is $\Omega(d^{-1/2})$ but for which the GW SDP optimal value is $O(d^{-3/4})$.*

The above lemma establishes the existence of a non trivial gap instance for sub-constant $\epsilon \sim d^{-1/2}$. The next task is to blow this instance up to create gap instances for constant ϵ. We do this by showing the following gap preservation lemma:

Lemma 3.1.3 (Gap Preservation)**.** *Suppose that I is a Max-2-LIN($\mathbb{Z}_2$) instance define over the d dimensional hypercube, and let α be the combinatorial value of I and β be the optimal GW SDP value for it. Then for every i there exists an instance I' defined over the $d \cdot i$ dimensional hypercube whose combinatorial value is at least α and whose GW SDP optimal value is at most β.*

Assuming the above mentioned lemmas the proof of Theorem 3.1.1 is straightforward and is presented below

Proof:[Proof of Theorem 3.1.1] Let $\epsilon > 0$ be small enough, and take $d = 1/\epsilon^2$. By lemma 3.1.2 we can find an instance I whose combinatorial value is at least $\Omega(d^{-1/2}) \sim \epsilon$ and whose GW SDP optimal value is at most $\mathcal{O}(d^{-3/4}) \sim \epsilon^{3/2}$. Considering the family of instances that can be obtained from I by applying Lemma 3.1.3 we obtain a $(\epsilon, \epsilon^{3/2})$ Instance of Max-2-LIN($\mathbb{Z}_2$) , proving the theorem. □

Over the next two subsections we prove Lemma 3.1.2 and Lemma 3.1.3

3.1.1 Proof of Lemma 3.1.2

Proof Overview The basic idea behind our proof of Lemma 3.1.2 is to produce a gap instance by making small perturbations to the instance with all edges assigned to 0 (equalities). In the following, we refer to edges of the hypercube connecting vertices that differ in the i-th coordinate as edges *going in the i-th direction*. We define a gap instance $\Delta(k, d)$ on Q_d, where we start from the all-equalities instance on Q_d and introduce inequalities along k directions, for $k \sim \sqrt{d}$. Our goal in choosing which edges to designate as inequality edges is to keep the solutions in which all variables are assigned the same value (say, the value one) to be close to optimal, which implies that the combinatorial optimum is roughly the fraction of inequality edges, while at the same time allowing an SDP solution of value noticeably higher than the fraction of inequalities, thus creating a gap.

We show that if we restrict ourselves to introducing inequality edges in just one direction, then up to about half the edges going in that direction can be changed to inequality while preserving the property that the all-ones assignment to the variables is optimal, and while allowing an SDP solution of smaller cost. We further augment the construction by placing these inequality regions along $\mathcal{O}(\sqrt{d})$ number of directions. The requirement of keeping the all ones solution nearly optimal dictates the choice of the parameter $\mathcal{O}(\sqrt{d})$. In particular in Lemmas 3.1.4 and 3.1.5 we prove that if we place these inequality regions in k directions, the combinatorial optimum grows linearly with k (in particular $\mathcal{O}(\frac{k}{d})$) whereas the SDP optimum grows at most proportionally to $\sqrt{k}$ (in particular $\mathcal{O}(\frac{\sqrt{k}}{d})$). Setting $k \sim \mathcal{O}(\sqrt{d})$ we get a non trivial (super-constant) gap.

Construction of the Instance We need to construct a Max-2-LIN($\mathbb{Z}_2$) instance over the hypercube Q_d of dimension d. Let E_d be the set of edges of the hypercube, and let k be a parameter to be fixed later. In this section we formally describe the construction of our Max-2-LIN($\mathbb{Z}_2$) instance $\Delta[k, d] : E_d \to \{0, 1\}$ supported on a Hypercube of dimension d. To remind the reader E_d is the edge set of a Hypercube of dimension d. The instance is parameterized by a number k which we fix later. For any edge $e = (v_1, v_2)$ let $i(e)$ be the coordinate along which the corresponding vectors $\mathbf{v_1}, \mathbf{v_2}$ differ. Let $H(\mathbf{v}[k])$ be the hamming weight of the vector $\mathbf{v}$ restricted to only coordinates other than the first k coordinates.

We define the function $\Delta[k, d]$ as follows

- If $i(e) > k$, $\Delta[k, d](e) = 0$.
- if $i(e) \leq k$ and if $H(\mathbf{v_1}[k]) > \frac{d-k}{2}$, $\Delta[k, d](e) = 0$.
- Otherwise $\Delta[k, d](e) = 1$.

Following are some observations about our instance. Note that all edges that are assigned to 1 (i.e. are inequality edges) are between vertices (v, v') that differ in one of the first k coordinates. Therefore for any subcube $Q_{d-k}(\mathbf{x})$ defined by fixing the first k coordinates to be $\mathbf{x}$, we have that the edges inside $Q_{d-k}(\mathbf{x})$ are all set to 0 (i.e. are equality edges). Consider two vectors $\mathbf{x_1}, \mathbf{x_2} \in \{0, 1\}^k$ which differ in one coordinate. Every vertex v in the subcube $Q_{d-k}(\mathbf{x_1})$ is connected by an edge to another vertex v' in $Q_{d-k}(\mathbf{x_2})$. The vertex v' can be thought of as a copy of v in the subcube $Q_{d-k}(\mathbf{x_2})$ (restricted to the last $d - k$ coordinates the two vertices are the same). The edge connecting (v, v') is an inequality or equality edge depending on which side of the majority cut v belongs to in its corresponding subcube. Specifically if more than half of the last $d - k$ coordinates of v are 1, then (v, v') is an equality edge, otherwise it is an inequality edge.

We now bound the GW-SDP optimum and the combinatorial optimum of $\Delta[k, d]$ in the following two lemmas.

Lemma 3.1.4. *For $k \leq \mathcal{O}(\sqrt{d})$, $\Delta[k, d]$ has combinatorial optimum $\Omega(\frac{k}{d})$.*

Lemma 3.1.5. *$\Delta[k, d]$ has GW-SDP optimum $O(\frac{\sqrt{k}}{d})$.*

Proof:[Proof of Lemma 3.1.2] The proof of Lemma 3.1.2 is immediate from Lemmas 3.1.2 and 3.1.4 and by setting $k = \Theta(\sqrt{d})$. □

We now give the proofs of the above mentioned Lemmas

3.1.2 Proof of Lemma 3.1.5

To prove the lemma it is enough to exhibit a valid solution to the GW-SDP which achieves a value of $O(\frac{\sqrt{k}}{d})$. To this end we exhibit a two dimensional solution $S : V_d \rightarrow \mathbb{R}^2$. Our solution will map every vertex v to a unit vector in $\mathbb{R}^2$ and therefore it is enough to just specify the angles α_v between v and the x-axis.

The solution S is symmetric with respect to the $d - k$ dimensional subcubes $Q_{d-k}(\mathbf{x})$ and depends only upon the parity of the k dimensional vector $\mathbf{x}$. Within a subcube, the vector assigned to a vertex depends only on the hamming weight of the vertex restricted to the subcube. Let $L_i(\mathbf{x})$ be the layer in the $d - k$ dimensional subcube $Q_{d-k}(\mathbf{x})$ of hamming weight i (vertices with i ones in the last $d-k$ coordinates). Formally a vertex $v \in L_i(\mathbf{x})$ if $v \in Q_{d-k}(\mathbf{x})$ and $H(\mathbf{v}[k]) = i$ (as a reminder, $H(\mathbf{v}[k])$ is the hamming weight of the vector $\mathbf{v}$ restricted to only coordinates after the k^{th} coordinate).

We now define our solution S to the GW-SDP parametrized by t. We will find a suitable value for t when we analyze the value of the solution.

- For every k-length vector $\mathbf{x}^+$ of parity 1, and for all $v \in L_i(\mathbf{x}^+)$

$$\alpha_v = \begin{cases} 0 & \text{if } i \leq \frac{d-k}{2} - t \\ \frac{\pi}{4}\left(1 - \frac{\frac{(d-k)}{2} - i}{t}\right) & \text{if } i \in (\frac{d-k}{2} - t, \frac{d-k}{2} + t) \\ \frac{\pi}{2} & \text{if } i \geq \frac{d-k}{2} + t \end{cases}$$

- For every k-length vector $\mathbf{x}^-$ of parity -1, and for all $v \in L_i(\mathbf{x}^-)$ assign α_v to be $\pi -$ the corresponding value for its neighboring vertex $\mathbf{x}^+$ of parity 1. i.e.

$$\alpha_v = \begin{cases} \pi & \text{if } i \leq \frac{d-k}{2} - t \\ \pi - \frac{\pi}{4}\left(1 - \frac{\frac{(d-k)}{2} - i}{t}\right) & \text{if } i \in (\frac{d-k}{2} - t, \frac{d-k}{2} + t) \\ \frac{\pi}{2} & \text{if } i \geq \frac{d-k}{2} + t \end{cases}$$

Following is a schematic of the solution described above. L represents layers of subcubes with parity 1 and the L' represent their counterparts in subcubes of parity -1

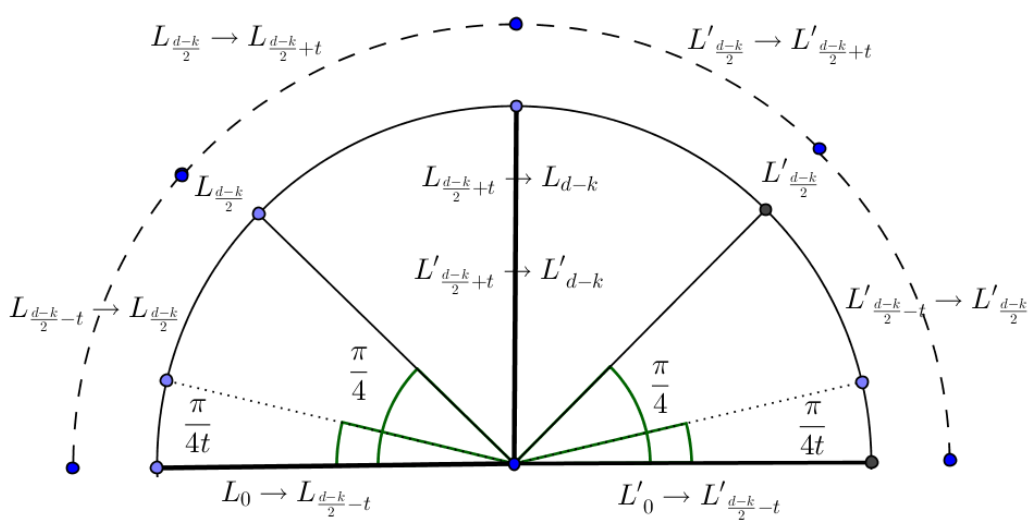

Figure 3.1: Schematic for SDP solution

We first compute the contribution of a fixed subcube $Q_{d-k}(\mathbf{x})$ to the GW-SDP objective. Consider a vertex $v \in L_i(\mathbf{x})$ where $i \in [0, \frac{(d-k)}{2} - t]$. This vertex is connected with equalities to its neighbours inside the subcube and with inequalities to its neighbours outside the subcube. Since all its neighbours inside the subcube are mapped to the same vector, the contribution to the SDP value of those edges is zero. Moreover all neighbors of v in different subcubes are mapped to the antipodal point of the vector v is mapped to (since neighboring subcubes have different parity). Therefore the contribution of every edge connected to this vertex is 0. Similarly, for a vertex $v \in L_i(\mathbf{x})$ where $i \in [(d-k)/2 + t, (d-k)]$ the contribution of all its edges is 0.

Consider a vertex $v \in L_i(\mathbf{x})$ where $i \in ((d-k)/2 - t, (d-k)/2)$. The total contribution of the neighbors of this vertex comes from the inequalities going out of the subcube, which is

$$k(1 + cos(\pi - 2\alpha_v)) \leq 2k$$

and from the equalities inside the subcube, which is

$$(d-k)(1 - cos(\frac{\pi}{4t}))$$

The total contribution of edges adjacent to v therefore is

$$2k + (d-k)(1 - cos(\frac{\pi}{4t})) \leq \mathcal{O}(k + \frac{(d-k)}{t^2})$$

The total fraction of vertices contained in layers $L_i(\mathbf{x})$ for $i = ((d-k)/2 - t, (d-k)/2 + t)$ is $\mathcal{O}(t/\sqrt{d-k})$ (for t=1 it is $\theta(\frac{1}{\sqrt{d-k}})$ and that is the layer with the largest fraction of vertices). Therefore the total contribution of a fixed subcube $Q_{d-k}(x)$ is bounded by

$$|V_{d-k}(x)|\mathcal{O}(\frac{t}{\sqrt{d-k}}\left(k + \frac{(d-k)}{t^2}\right)$$

Substituting $t = \sqrt{\frac{d-k}{k}}$ and summing the contribution over all subcubes $Q_{d-k}(x)$ we get that the fractional value of this SDP feasible solution is $\mathcal{O}(\frac{\sqrt{k}}{d})$.

3.1.3 Proof of Lemma 3.1.4

In this section we prove Lemma 3.1.4. Consider the $d-k$ dimensional subcubes $Q_{d-k}(\mathbf{x})$ where $\mathbf{x}$ is a k dimensional vector. We begin by two observations which can be made without loss of generality about any optimal assignment $\Gamma :: V_d \rightarrow \{0, 1\}$. The proofs of the observations are provided from completeness

Observation 1. *Without loss of generality we can assume that for any optimum assignment, the assignment restricted* *on any subcube $Q_{d-k}(\mathbf{x})$ can be assumed to be only dependent on the parity of $\mathbf{x}$. In other words, if $\mathbf{x}, \mathbf{y}$ are k dimensional vectors with the same parity then the assignments on the subcubes $Q_{d-k}(\mathbf{x}), Q_{d-k}(\mathbf{y})$ will be the same.*

Proof:[Proof of Observation 1]

We prove this by contradiction. Let Γ be an optimum assignment. For a subset of edges $E \subseteq E_d$ let $Val_\Gamma(E)$ be the number of unsatisfied edges in E. Let $Val_\Gamma(Q_{d-k}(\mathbf{x}))$ be the number of unsatisfied edges in the subcube $Q_{d-k}(\mathbf{x})$.

Let S be the set of pairs of k dimensional vectors $\mathbf{x_1}, \mathbf{x_2}$ which differ in one coordinate. Given any two such vectors $\mathbf{x_1}, \mathbf{x_2}$, let $E(\mathbf{x_1}, \mathbf{x_2})$ be the set of edges (u, v) that go between the corresponding subcubes i.e. $u \in Q_{d-k}(\mathbf{x_1}), v \in Q_{d-k}(\mathbf{x_2})$. Therefore the total combinatorial value of the assignment Γ(i.e. total number of unsatisfied edges) is

$$\sum_{\mathbf{x}} Val_\Gamma(Q_{d-k}(\mathbf{x})) + \sum_{(\mathbf{x_1},\mathbf{x_2})\in S} Val_\Gamma(E(\mathbf{x_1}, \mathbf{x_2})) = \sum_{(\mathbf{x_1},\mathbf{x_2})\in S} \left(\frac{1}{k}\Big(Val_\Gamma(Q_{d-k}(\mathbf{x_1})) + Val_\Gamma(Q_{d-k}(\mathbf{x_2}))\Big) + Val_\Gamma(E(\mathbf{x_1}, \mathbf{x_2}))\right) \quad (3.1)$$

Given the above expression let $\mathbf{x_1'}, \mathbf{x_2'}$ be vectors such that the quantity inside the summation in the RHS above is minimum. Now consider the assignment in which for every vector $\mathbf{x}$ which has the same parity as $\mathbf{x_1'}$, the subcube $Q_{d-k}(\mathbf{x})$ has the same assignment as the subcube $Q_{d-k}(\mathbf{x_1'})$ in Γ. We do the same with $\mathbf{x_2'}$. It is easy to see that the above described assignment satisfies at least as many edges as Γ. □

By the above argument for any optimal assignment Γ it is enough to specify two assignment functions $\Gamma_+ : Q_{d-k} \to \{0,1\}$ and $\Gamma_- : Q_{d-k} \to \{0,1\}$, one for subcubes for which the first k coordinates have parity 1 and one for subcubes for which the first k coordinates have parity -1. Let $Val(\Gamma_+)$ and $Val(\Gamma_-)$ be the number of edges not satisfied within the subcubes of positive and negative parity respectively. Let $Val(\Gamma_+, \Gamma_-)$ denote the number of edges not satisfied between a fixed subcube of positive parity and a fixed subcube of negative parity. The total number of edges not satisfied by the assignment Γ therefore is

$$2^{k-1}(Val(\Gamma_+) + Val(\Gamma_-)) + 2^{k-1}k(Val(\Gamma_+, \Gamma_-))$$

Observation 2. *Without loss of generality the assignment Γ_- can be assumed to be the all 1's assignment $\mathbf{1}$ i.e. $\mathbf{1}(v) = 1$ for all $v \in Q_{d-k}$.*

Proof:[Proof of Observation 2] Consider any optimal assignment (Γ'_+,Γ'_-). Consider the assignment such that $\Gamma_- = \mathbf{1}$ and $\Gamma_+ = \Gamma'_+ \oplus \Gamma'_-$. Note that $Val(\Gamma'_+, \Gamma'_-) = Val(\mathbf{1}, \Gamma'_+ \oplus \Gamma'_-)$. Also note that $Val(\mathbf{1}) = 0$ and $Val(\Gamma'_+ \oplus \Gamma'_-) \leq Val(\Gamma'_+) + Val(\Gamma'_-)$. Therefore the assignment $(\mathbf{1}, \Gamma'_+ \oplus \Gamma'_-)$ is at least as good as (Γ'_+, Γ'_-). □

In accordance with Observations 1 and 2 for an optimal assignment it is enough to specify the assignment Γ_+ for the positive parity subcubes.

Consider an optimal solution $(\Gamma'_+, \mathbf{1})$. Let $d' = d - k$. Let $V_0 \subseteq V_{d'}$ be the set of vertices v such that $\Gamma_+(v) = 0$ and $H_{1/2} \in V_{d'}$ be the set of vertices v such that $H(\mathbf{v}) \leq d'/2$. Now it is easy to see that the number of edges unsatisfied by the assignment $(\Gamma'_+, \mathbf{1})$ is

$$\begin{aligned}
&2^{k-1}\Big(Val(\Gamma_+) + Val(\mathbf{1}) + k * Val(\Gamma_+, \mathbf{1})\Big) \\
&= 2^{k-1}\Big(E[V_0, V_{d'} - V_0] + 0 + k(|H_{1/2} - V_0| + |V_0 - H_{1/2}|)\Big) \\
&= 2^{k-1}\Big(E[V_0, V_{d'} - V_0] + 0 + k(|H_{1/2}| - |H_{1/2} \cap V_0| + |V_0 - H_{1/2}|)\Big) \\
&= 2^{k-1}\Big(\frac{k * 2^{d'}}{2} - A(k, d')\Big)
\end{aligned}$$

where $A(k, d') = k(|V_0 \cap H_{1/2}| - |V_0 - H_{1/2}|) - E[V, V_{d'} - V_0]$. We show in lemma 3.1.6 that the above defined quantity $A(k, d') \leq \frac{k\alpha}{2} * 2^{d'}$ for $k \leq \mathcal{O}(\sqrt{d})$ where α is a universal constant.

Therefore the total fraction of edges unsatisfied by the any optimum assignment is $\mathcal{O}(k/d)$ for $k = \mathcal{O}(\sqrt{d})$. This completes the proof of Lemma 3.1.4 assuming Lemma 3.1.6 which we prove next

Lemma 3.1.6. *Let Q_d be the hypercube of dimension d. Let V_d be the vertex set of the cube and let $V \subseteq V_d$. Let $H_{1/2}$ be the set of vertices with hamming weight $\leq d/2$. Let $k \leq \frac{\pi}{2}\mathbb{I}(Maj_d)$, where $\mathbb{I}(Maj_d) = \Theta(\sqrt{d})$ is the influence of the majority function on d coordinates. Then*

$$A(k, d) \stackrel{def}{=} k(|V \cap H_{1/2}| - |V - H_{1/2}|) - E[V, V_d - V] \leq \alpha \frac{k2^d}{2}$$

where α is a constant < 1.

Proof: Let Min_d, Maj_d be the minority/majority Boolean function over d variables. Let $\mathbb{I}(Maj_d) = \mathbb{I}(Min_d)$ be the influences of the functions Maj_d, Min_d. Note that Min_d is the indicator function of the $H_{1/2}$. Let $f : \{-1, 1\}^n \to \{-1, 1\}$ be the associated Boolean function which is the indicator function of the set V. Then

$$\begin{aligned}
A(k, d) &= k(|V \cap H_{1/2}| - |V - H_{1/2}|) - E[V, V_d - V] \\
&= \frac{2^{d'}}{2}\Big(k\langle f, Min_{d'}\rangle) - \mathbb{I}(f)\Big)
\end{aligned}$$

We now show the following inequality on Boolean functions over the cube, which proves the lemma when $k \leq \frac{\pi}{2}\mathbb{I}(Min_d) = \mathcal{O}(\sqrt{d})$.

Claim. *There exists a constant $\alpha < 1$, such that for any Boolean function f on the d-dimensional cube,*

$$\frac{\pi}{2} \cdot \mathbb{I}(Min_d) \cdot (\langle f, Min_d\rangle - \alpha) \leq \mathbb{I}(f)$$

Proof: The proof of the claim is a direct application of the following facts whose proof can be found in any standard text on Boolean Functions such as [O'D14]

$$|\widehat{Min_d}(\{i\})| \sim \sqrt{\frac{2}{\pi d}} \tag{3.2}$$

$$\sum_{|S|\geq 2} \widehat{Min_d}(S)^2 \leq (1 - \frac{2}{\pi}) \tag{3.3}$$

$$\mathbb{I}(Min_d) \sim \sqrt{\frac{2}{\pi} d} \tag{3.4}$$

$$\sum_i |\widehat{f}(\{i\})| \leq \mathbb{I}(f) \tag{3.5}$$

Using the above facts we have that

$$\begin{aligned}
\frac{\pi}{2}\mathbb{I}(Min_d)\langle f, Min_d\rangle &= \frac{\pi}{2}\mathbb{I}(Min_d)\Big(\sum \widehat{f}(S)\widehat{Min_d}(S)\Big) \\
&\leq \frac{\pi}{2}\mathbb{I}(Min_d)\left(\sum_{|S|=1} |\widehat{f}(S)||\widehat{Min_d}(S)| + \sum_{|S|\geq 2} \widehat{f}(S)\widehat{Min_d}(S)\right) \\
&= \frac{\pi}{2}\mathbb{I}(Min_d)\left(\sum_{|S|=1} |\widehat{f}(S)|\sqrt{\frac{2}{\pi d}} + \sum_{|S|\geq 2} \widehat{f}(S)\widehat{Min_d}(S)\right) && \text{(by Equation 3.2)} \\
&\leq \frac{\pi}{2}\mathbb{I}(Min_d)\left(\sum_{|S|=1} |\widehat{f}(S)|\sqrt{\frac{2}{\pi d}} + \sqrt{\sum_{|S|\geq 2} \widehat{f}^2(S)}\sqrt{\sum_{|S|\geq 2} \widehat{Min_d}^2(S)}\right) && \text{(by Cauchy Schwartz)} \\
&\leq \sum_{|S|=1} |\widehat{f}(S)| + \frac{\pi}{2}\mathbb{I}(Min_d)\left(\sqrt{1-\frac{2}{\pi}}\right) && \text{(by Equations 3.3,3.4)} \\
&\leq \mathbb{I}(f) + \frac{\pi}{2}\mathbb{I}(Min_d)\left(\sqrt{1-\frac{2}{\pi}}\right) && \text{(by Equation 3.5)}
\end{aligned}$$

Putting $\alpha = \sqrt{1-\frac{2}{\pi}}$ proves the claim, and thus also completes the proof of Lemma 3.1.6. □

□

3.1.4 Proof of Lemma 3.1.3

In this section, we prove the gap preservation lemma 3.1.3. To prove the lemma we define the following general operation on Max-2-LIN($\mathbb{Z}_2$) instances on the cube.

Definition 3.1.7 (Tensor Max-2-LIN($\mathbb{Z}_2$))**.** *Given two Max-2-LIN($\mathbb{Z}_2$) instances $\Gamma_1 : E_{d_1} \to \{0,1\}$ and $\Gamma_2 : E_{d_2} \to \{0,1\}$ supported on Hypercubes of dimensions d_1 and d_2, define a Max-2-LIN($\mathbb{Z}_2$) instance $\Gamma_1 \otimes \Gamma_2 : E_{d_1+d_2} \to \{0,1\}$ on the hypercube of dimension $d_1 + d_2$ as follows. For an edge (v_1, v_2) let $i(v_1, v_2)$ be the coordinate on which* *the corresponding vectors* $\mathbf{v}_1, \mathbf{v}_2$ *differ. Let* $\mathbf{v}_i^{d_1}$,$\mathbf{v}_i^{d_2}$ *be the vector* $\mathbf{v}_i$ *restricted on the first* d_1 *coordinates and the last* d_2 *coordinates respectively. Then*

$$\Gamma_1 \otimes \Gamma_2((v_1, v_2)) = \begin{cases} \Gamma_1((\mathbf{v}_1^{d_1}, \mathbf{v}_2^{d_1})) & \textit{if } i(v_1, v_2) \in [0, d_1 - 1] \\ \\ \Gamma_2((\mathbf{v}_1^{d_2}, \mathbf{v}_2^{d_2})) & \textit{if } i(v_1, v_2) \in [d_1, d_1 + d_2 - 1] \end{cases}$$

Note that the above tensor product defines an edge according to the first instance or the second instance depending upon the coordinate along which the edge crosses. We prove the following lemmas about the above defined tensor product.

Lemma 3.1.8. *Let* Γ_1 *have combinatorial optimum* $\geq \beta_1$ *and* Γ_2 *have combinatorial optimum* $\geq \beta_2$*. Then the combinatorial optimum of* $\Gamma_1 \otimes \Gamma_2$ *is* $\geq \frac{d_1\beta_1 + d_2\beta_2}{d_1 + d_2}$.

Lemma 3.1.9. *Let* Γ_1 *have SDP optimum* $\leq \alpha_1$ *and* Γ_2 *have SDP optimum* $\leq \alpha_2$*. Then the SDP optimum of* $\Gamma_1 \otimes \Gamma_2$ *is* $\leq \frac{d_1\alpha_1 + d_2\alpha_2}{d_1 + d_2}$.

Note that given any (α, β) Max-2-LIN($\mathbb{Z}_2$) Instance Γ on the cube of dimension d define $\Gamma_i = \otimes_1^i \Gamma$. By lemmas 3.1.9 and 3.1.8 we get that Γ_i is an (α, β) Instance on the hypercube of dimension $i.d$. This proves lemma 3.1.3.

Proof:[Proof of lemma 3.1.8] We prove the lemma by proving that the instance $\Gamma_1 \otimes \Gamma_2$ can be partitioned into edge disjoint copies of the instances Γ_1 and Γ_2. Consider any $d_1 + d_2$ dimensional vector. Fix the last d_2 coordinates and vary the first d_1 coordinates within the space $\{0, 1\}^{d_1}$. Note that the vectors generated by the above process naturally define a subset of edges of the cube $Q_{d_1+d_2}$. Also note that the subset of edges generated is an exact copy of Γ_1. Therefore repeating the above process for all choices of the last d_2 coordinates gives us 2^{d_2} edge disjoint copies of Γ_1 within $\Gamma_1 \otimes \Gamma_2$. Fixing the first d_1 coordinates and repeating the same line of argument as above we get 2^{d_1} edge disjoint copies of Γ_2 within $\Gamma_1 \otimes \Gamma_2$. Note that the copies described above form an edge partition of $\Gamma_1 \otimes \Gamma_2$. Therefore the combinatorial optimum of the instance is

$$\begin{aligned} &\geq \frac{1}{E_{d_1+d_2}}\left(2^{d_2}\frac{d_1 2^{d_1}}{2}\beta_1 + 2^{d_1}\frac{d_2 2^{d_2}}{2}\beta_2\right) \\ &\geq \frac{d_1\beta_1 + d_2\beta_2}{d_1 + d_2} \end{aligned}$$

□

Proof:[Proof of lemma 3.1.9] It is enough to give one SDP solution which has the required value. Let the optimal SDP solution for Γ_1 be $S_1 : V_{d_1} \to \mathbb{R}^{2^{d_1}}$ and for Γ_2 be $S_2 : V_{d_2} \to \mathbb{R}^{2^{d_2}}$. Define the following solution $S_1 \otimes S_2 : V_{d_1+d_2} \to \mathbb{R}^{2^{d_1+d_2}}$.

$$S_1 \otimes S_2(\mathbf{v}) = S_1(\mathbf{v}^{d_1}) \otimes S_2(\mathbf{v}^{d_2})$$

It is immediate by the properties of tensor products of vectors that $S_1 \otimes S_2$ is a valid SDP solution. We now compute the SDP value of $S_1 \otimes S_2$. Let $E_1 \in E_{d_1+d_2}$ be the set of edges which go through the first d_1 coordinates and $E_2 \in E_{d_1+d_2}$ be the set of edges which go through the last d_2 coordinates. Note that $|E_1| = \frac{d_1}{2} 2^{d_1+d_2}$ and $|E_2| = \frac{d_2}{2} 2^{d_1+d_2}$. Therefore the SDP value achieved by the solution is

$$\begin{aligned}
&\frac{1}{E_{d_1+d_2}} \left(\sum_{(v_1,v_2)\in E_1} \|S_1 \otimes S_2(v_1) \pm S_1 \otimes S_2(v_2)\|^2 + \sum_{(v_1,v_2)\in E_2} \|S_1 \otimes S_2(v_1) \pm S_1 \otimes S_2(v_2)\|^2 \right) \\
&= \frac{1}{E_{d_1+d_2}} \left(\sum_{(v_1,v_2)\in E_1} \|S_2(\mathbf{v}_1^{d_2})\|^2 \|\|S_1(\mathbf{v}_1^{d_1}) \pm S_2(\mathbf{v}_2^{d_1})\|^2 + \sum_{(v_1,v_2)\in E_2} \|S_1(\mathbf{v}_1^{d_1})\|^2 \|\|S_2(\mathbf{v}_1^{d_2}) \pm S_2(\mathbf{v}_2^{d_2})\|^2 \right) \\
&= \frac{1}{E_{d_1+d_2}} \left(\alpha_1 d_1 2^{d_1+d_2} + \alpha_2 d_2 2^{d_1+d_2} \right) \\
&= \frac{d_1\alpha_1 + d_2\alpha_2}{d_1 + d_2}
\end{aligned}$$

□

3.2 Towards solving Max-2-LIN($\mathbb{Z}_2$) on Hypercubes

In this section we propose a candidate algorithm for solving Max-2-LIN($\mathbb{Z}_2$) on the Hypercube. By solving we mean solving the decision version of the problem. Specifically for our purpose it suffices to show that given a Max-2-LIN($\mathbb{Z}_2$) instance whose combinatorial optimum is ϵ, the SDP optimum is $\Omega(\epsilon)$. Our candidate algorithm is simply augmenting the Goemans Williamson SDP with appropriate triangle inequalities. We conjecture that the augmented SDP is strong enough to solve Max-2-LIN($\mathbb{Z}_2$) on the Hypercube. In particular we show that the augmented SDP has $\mathcal{O}(1)$ integrality gap on our proposed instance $\Delta(k, d)$. The motivation behind our conjecture comes from the fact which we show next that on a cycle of any length the augmented SDP has no integrality gap. This implies in particular that an inconsistent cycle in a graph acts as a certificate for an unsatisfied edge in the SDP solution as well. Therefore the property of necessarily having many inconsistent cycles makes Max-2-LIN($\mathbb{Z}_2$) instances on a graph solvable by the augmented SDP. We show that our instance indeed has a lot of inconsistent cycles and by conjecturing that in fact any instance on the Boolean Hypercube satisfies this property. Using this motivation of finding inconsistent cycles we build a connection between multicommodity flow cut gaps on graphs and the integrality gap of the augmented SDP. This connection automatically gives us an exact algorithm for Max-2-LIN($\mathbb{Z}_2$) on planar graphs.

We begin by defining our augmented SDP.

Definition 3.2.1 (GW+). *Given a graph $G = (V, E)$ and a Max-2-LIN($\mathbb{Z}_2$) instance $I : E \to \{0, 1\}$ on it, let the set of equality edges be E^+ and the set of inequality edges be E^-. The augmented Goemans-Williamson SDP for the instance is defined as*

$$\textit{minimize } \frac{1}{4|E|}\left(\sum_{(u,v)\in E^+} \|x_u - x_v\|^2 + \sum_{(u,v)\in E^-} \|x_u + x_v\|^2\right)$$

$$\textit{subject to} \quad \|x_u\|^2 = 1 \ \ (\forall\, u \in V)$$

$$\|a_i - a_j\|^2 \le \|a_i - a_k\|^2 + \|a_k - a_j\|^2 \ \ (\forall\, i, j, k \in V, a_i = \pm x_i, a_j \pm x_j, a_k \pm x_k)$$

One way in which the GW+algorithm improves on the Goemans-Williamson algorithm is that is takes inconsistent cycles into account, as is formalised below.

Definition 3.2.2 (inconsistent cycles). *Let I be a Max-2-LIN($\mathbb{Z}_2$) instance defined on a graph G. A cycle in G is said to be inconsistent if no assignment can satisfy all edges of the cycle (note that there is always an assignment that satisfies all edges of a cycle but one).*

We show in the following theorem that GW+has no integrality gap on cycles and therefore GW+has a high value on graphs which have a good number of such cycles.

Theorem 3.2.3. *Consider a Max-2-LIN($\mathbb{Z}_2$) instance I defined on a graph $G = (V, E)$, and suppose that there are $\epsilon \cdot |E|$ edge-disjoint inconsistent cycles in the instance. Then given I, the value returned by GW+is at least ϵ.*

Proof: First consider the case where the instance just contains one cycle C. If it is consistent, it is easy to see that the SDP achieves a value of 0. We now focus on inconsistent cycles. Let $u_0, u_1 \ldots u_{n-1}$ be the vertices of the cycle in order and let $u_n = u_0$. Let E_i be edge connecting $u_i \to u_{i+1}$ and let $C(E_i)$ be defined to be 1 if there is an equality constraint on E_i and -1 otherwise. Define

$$sign(i) = \Pi_{j=0}^{i} C(E_i)$$

Note that $sign(0) = 1$ and $sign(n) = -1$ the latter because the cycle is inconsistent. The objective function of GW+now is thus

$$\begin{aligned}\frac{1}{4}\left(\sum_{i=0}^{n}\|sign(i)X_{u_i}-sign(i+1)X_{u_{i+1}}\|^2\right) &\geq \frac{1}{4}\left(\|sign(0)X_{u_0}-sign(n)X_{u_n}\|^2\right)\\ &= \frac{1}{4}\|2(X_{u_0})\|^2\\ &= 1\end{aligned}$$

The first inequality follows from the triangle inequalities added to GW+. The above implies that GW+has no gap on a cycle(inconsistent or consistent).

Now for a general instance, note that the above implies that any inconsistent cycle in the given instance must contribute at least 1 to the value of the objective function in GW+. In particular if we can find $\epsilon \cdot |E|$ inconsistent edge disjoint cycles in the given instance we can be assured that the GW+optimum is at least ϵ, as required. □

Theorem 3.2.3 naturally leads us to the interesting question whether there are instances of Max-2-LIN($\mathbb{Z}_2$) on the hypercube which are ϵ unsatisfiable, and yet there are not enough disjoint inconsistent cycles that certify the value to be at least $\Omega(\epsilon)$. We conjecture that in fact there are no such instances, and therefore that the GW+SDP has at best a $\mathcal{O}(1)$ integrality gap for Max-2-LIN($\mathbb{Z}_2$) on the hypercube.

Conjecture 3. *Given a Max-2-LIN($\mathbb{Z}_2$) instance on the Hypercube (V_d, E_d) such that in any labeling at least ϵ fraction of its edges are unsatisfied, then there are at least $\Omega(\epsilon|E_d|)$ edge disjoint inconsistent cycles in the instance.*

One motivation behind Conjecture 3 is the presence of a large number of short cycles containing every edge – each edge is contained in d 4-cycles, d^2 6-cycles and so on. We now show that Conjecture 3 is true for our instance $\Delta(k, d)$, defined previously. Recall that when $k \leq \mathcal{O}(\mathcal{I}(Maj_d))$, the combinatorial optimum is $\Theta(\frac{k}{d})$

Theorem 3.2.4. *Let $I = \Delta(k, d)$ be the Max-2-LIN($\mathbb{Z}_2$) instance defined in Section 3.1, where $k \leq \mathcal{O}(I(Maj_d)) = \mathcal{O}(\sqrt{d})$. Then there are at least $\Omega(\frac{k}{d} \cdot |E|)$ edge disjoint inconsistent cycles in I, where E is the set of edges in I.*

Proof: We first investigate the number of inconsistent edge disjoint cycles in our instance between two subcubes of dimension $d-k$. The inconsistent edge disjoint cycles in the whole instance will just be the union of these cycles over all subcubes.

Edge-disjoint paths To find the required cycles, we first consider a $d-k$ dimensional subcube $Q_{d-k}(\mathbf{x})$ inside our instance, and let $H_{1/2}$ be the set of vertices with Hamming weight $\leq \frac{d-k}{2}$ inside it (we only consider the Hamming weight relative to the subcube). We would like to find many edge disjoint simple paths in the cube Q_{d-k} such that for every vertex $v \in H_{1/2}$ there are at least $\ell = \Theta(\sqrt{d})$ paths that start from it and end in a vertex outside of $H_{1/2}$. For every vertex $v' \in V_{d-k} - H_{1/2}$ we also require that at most ℓ paths end at any one vertex.

Note that if P is a path of this type, and if Q is taken to be the same path but on a neighbouring subcube $Q_{d-k}(\sigma_i(\mathbf{x}))$ (σ_i flips the i'th bit of $\mathbf{x}$), then the two paths can be joined to create an inconsistent cycle.

Note that the above problem is equivalent to the following flow system. Let every edge within Q_{d-k} have capacity 1, and add a source s that connects to every vertex $v \in H_{1/2}$ with an edge of capacity ℓ and a target t that connects to every vertex outside of $H_{1/2}$ with an edge of capacity ℓ. If this system has a flow that saturates the edges going out of s and into t, then we can find the needed paths in our instance: that follows since if such a flow exists there must also be an equivalent integral flow. Once an integral flow is achieved, it is easy to see that it can be broken into edge-independent paths inside the subcube.

To see whether the flow system is satisfiable or not we simply need to check the whether every $s-t$ cut is flow sufficient. Consider any cut $V \subset V_{d-k}$. Note that the demand of the cut is $\ell|(|V - H_{1/2}| - |H_{1/2} - V|)|$ and the capacity of the cut is $E(V, V_{d-k} - V)$. Note that lemma 3.1.6 implies that for $\ell \leq \mathcal{O}(\sqrt{d})$ the cut is flow sufficient.

Stitching paths together. For every path $P = P(\mathbf{x})$ that we found in $Q_{d-k}(\mathbf{x})$, we can take a corresponding path $P(\mathbf{y})$ in any other subcube. We thus have a system of disjoint paths in the subcubes of our instance. Let us show how to stitch them together to get edge disjoint cycles. For this purpose, consider the graph G on the subcube $Q_{d-k}(\mathbf{x})$ which connects two points when they are connected by one of our chosen paths. G is a bipartite graph, and because of the way the paths were selected, it is regular and each vertex has degree ℓ. It is well known that such a graph can always be partitioned into ℓ matchings (e.g. using Hall's theorem): this means that we can choose a color $i = i(P)$ for each path, $i \in \{1, \ldots, \ell\}$, such that no vertex connects to two paths with the same color.

Now each path $P(\mathbf{y})$ in a subcube $Q_{d-k}(\mathbf{y})$ can be matched to the similar path $P(\mathbf{y}')$ in $Q_{d-k}(\mathbf{y}')$, where $\mathbf{y}' = \sigma_i(\mathbf{y})$ and $i = i(P)$ is the index chosen by P (since $\ell \leq k$, also $i \leq k$). Joining the endpoints of those paths creates an inconsistent cycle, and it is easy to verify that this indeed gives a system of edge-disjoint inconsistent cycles in $\Delta(k, d)$.

Counting cycles. As we constructed $\ell \cdot 2^{d-k}$ disjoint paths in each subcube, and since each cycle consists of two such paths, the total number of cycles is $\ell \cdot 2^{d-k} \cdot 2^k/2 = \Omega(k \cdot 2^d)$. Since the number of edges in $\Delta(k, d)$ is $d \cdot 2^d$, the number of cycles is $\Omega(\frac{k}{d} \cdot |E|)$ as required. □

3.3 Max-2-LIN($\mathbb{Z}_2$) and Multi-commodity Flow-Cut Gaps

In this section we build a connection between the integrality gap of the augmented Goemans-Williamson SDP defined in the previous section (refer Def 3.2.1) and the multicommodity cut flow gaps on certain special graphs. Our motivation comes from the connections between the existence of a large number of edge disjoint cycles and the integrality gap of GW+which we described in the previous section.

To remind the reader a Multicommodity Flow network defined on a graph $G = (V, E)$ is specified by a capacity function $C : E \to \mathbb{R}_+$ and a set of demand pairs $D = \{(s_i, t_i) \in V \times V\}$ with associated demand d_i. Let P^i be the set of paths connecting s_i to t_i. A flow in the network is a function f which for every path $p \in P^i$ assigns a positive value also referred to as the flow for the i^{th} commodity along the path p. A flow is feasible if for every edge the total flow along every path that the edge appears in for every commodity is bounded above the by the capacity along that edge. The objective is to set up a flow system in the graph routing as large a fraction of every demand as possible. This fraction is usually referred to as λ. For formal definitions of the multicommodity flow problem and the associated algorithms we refer the reader to the following survey [Shm97]

For a particular Multicommodity Flow instance G, C, D we say that the cut condition is satisfied if for any subset of vertices $S \subseteq V$, the total capacity of the edges leaving the set is greater than or equal to the total demand that goes across the set. Clearly for λ to be 1 it is necessary that the cut condition is satisfied but it is well known that it is not sufficient. The ratio $1/\lambda$ is referred to as the Cut-Flow gap of the network.

Now given any Max-2-LIN($\mathbb{Z}_2$) instance I supported on a graph G and any assignment $\Gamma : V \to \{0, 1\}$ we consider the following Multicommodity Flow network. Let $\hat{E} \subseteq E$ be the set of edges that are not satisfied by the assignment Γ. Define the capacity and demand function as follows

$$C_\Gamma(e) = \begin{cases} 0 & \text{if } e \in \hat{E} \\ 1 & \text{if } e \in E - \hat{E} \end{cases}$$

$$D_\Gamma = \{(s_i, t_i) | (s_i, t_i) \in \hat{E}\}$$

Note that the above described instance can be seen as a generalization of the edge disjoint cycles problem described in Conjecture 3. In particular if any assignment Γ does not satisfy ϵ fraction of the edges and a certificate of that fact can be obtained by producing $\alpha\epsilon|E|$ cycles such that each cycle contain exactly one edge from $\hat{E}$ then it is easy to see that an α fraction of the demand can be routed in the above network.

The Multicommodity flow network also provides a neat way of characterizing the optimal combinatorial solution $\hat{\Gamma}$. Specifically we prove the following claim

Claim. *Given a Max-2-LIN($\mathbb{Z}_2$) instance I supported on $G = (V, E)$, let $\hat{\Gamma}$ be an optimal assignment for I then for the multicommodity flow network corresponding to $\hat{\Gamma}$ the cut condition is satisfied.*

Proof:[Proof of Claim 3.3] Consider any subset $S \in V$. Let $\delta(S)$ be the set of edges (u, v) such that $u \in S$ and $v \notin S$. Let $\hat{E}$ be the set of edges that are not satisfied by the assignment $\hat{\Gamma}$. Let $\hat{\delta(S)} = \hat{E} \cap \delta(S)$. By the definition of the Multicommodity flow network it follows that the cut condition is equivalent to $|\delta(S) - \hat{\delta(S)}| \geq |\hat{\delta(S)}|$. We will show that the above condition is true for all S. If not consider any $S \in V$ that violates this condition. Now consider the assignment $\hat{\Gamma}[S]$ defined as follows

$$\hat{\Gamma}[S](v) = \begin{cases} \hat{\Gamma}[S](v) & \text{if } v \notin S \\ \sim \hat{\Gamma}[S](v) & \text{if } v \in S \end{cases}$$

It is now easy to see that the above defined assignment $\hat{\Gamma}[S]$ has strictly more number of edges satisfied than $\hat{\Gamma}$ which contradicts the optimality of $\hat{\Gamma}$ □

Having established the relationship between cut-condition and optimality we can now relate the Integrality Gap of GW+and the Flow-Cut Gaps of the Multicommodity flow networks induced by the assignments. The theorem can be seen as a fractional relaxation of Theorem 3.2.3

Theorem 3.3.1. *Let I be a Max-2-LIN($\mathbb{Z}_2$) instance supported on the a graph $G = (V, E)$. Let $\hat{\Gamma}$ be the optimal assignment and let ϵ be the fraction of edges unsatisfied. Let $G, \hat{C}, \hat{D}$ be the Multicommodity Flow network induced by the assignment $\hat{\Gamma}$ and λ be the optimum fractional multicommodity flow network. Then GW+has optimum at least $\lambda\epsilon$.*

Proof: Let $\hat{E}$ be the edges unsatisfied by $\hat{\Gamma}$. These edges form the demand pairs in the network $G, \hat{C}, \hat{D}$. Consider an optimum flow f achieving optimum λ. For every edge $(u, v) \in \hat{E}$ let P_{uv} be the set of paths connecting u, v. Now by definition for every path $p \in P_{u,v}$, $f(p)$ is defined to be a non-negative value. Also $\sum_p f(p) = \lambda$. The above leads to the following simple decomposition of the GW+objective function on the instance I

$$\begin{aligned}
\sum_{(u,v)\in E} \|x_u \pm x_v\|^2 &= \sum_{(u,v)\in \hat{E}} \|x_u \pm x_v\|^2 + \sum_{(u,v)\in E-\hat{E}} \|x_u \pm x_v\|^2 \\
&\geq \sum_{(u,v)\in \hat{E}} \lambda\|x_u \pm x_v\|^2 + \sum_{(u,v)\in E-\hat{E}} \|x_u \pm x_v\|^2 \\
&= \sum_{(u,v)\in \hat{E}} \sum_{p\in P_{u,v}} f(p)\|x_u \pm x_v\|^2 + \sum_{(u,v)\in E-\hat{E}} \|x_u \pm x_v\|^2 \\
&\geq \sum_{(u,v)\in \hat{E}} \sum_{p\in P_{u,v}} f(p)\|x_u \pm x_v\|^2 + \sum_{(u,v)\in \hat{E}} \sum_{p\in P_{u,v}} \sum_{(u',v')\in p} f(p)\|x_{u'} \pm x_{v'}\|^2 \\
&= \sum_{(u,v)\in \hat{E}} \sum_{p\in P_{u,v}} f(p)\left(\|x_u \pm x_v\|^2 + \sum_{(u',v')\in p} \|x_{u'} \pm x_{v'}\|^2\right) \\
&\geq \sum_{(u,v)\in \hat{E}} \sum_{p\in P_{u,v}} 4f(p) = 4\lambda\epsilon|E|
\end{aligned}$$

The fourth line above follows from the edge capacity constraints applied to each edge in a feasible flow. The second last inequality follows from the fact that if for each edge $(u, v) \in \hat{E}$ we consider the cycle the cycle formed by a path $p \in P_{uv}$ and (u, v) this cycle is inconsistent(as it contains exactly one unsatisfied edge) and therefore the contribution of the cycle to the SDP optimum is 4 (this was proved in the proof of Theorem 3.2.3). □

The above theorems prove that the flow cut gap on these networks provides an upper bound on the Integrality Gap of GW+SDP relaxation. Therefore we make the following statement.

Corollary 3.3.2. *Given any graph family F consider Multicommodity Flow networks created on a graph $G \in F$ by designating some edges to be demand pairs with demand 1 and the rest to be supply edges with capacity 1. If the cut-flow gap on any $G \in F$ for such networks is bounded by λ then the GW+has an integrality gap of λ*

Note that our above corollary requires the demand edges to be selected out of the edges of the base graph itself. This is a key restriction which we hope might help to achieve $\mathcal{O}(1)$ flow cut gaps. Note that it is known that the flow cut gap in such a network for planar graphs is 1 ([Sch03]). This in particular implies that in case of planar graphs the SDP optimum = Combinatorial Optimum.

3.4 Conclusion and Open Problems

In this section we studied the Max-2-LIN($\mathbb{Z}_2$) problem on the Boolean Hypercube. In particular we showed that the standard Goemans Williamson SDP has a non-trivial integrality gap when the constraint graph is a Boolean Hypercube. Unlike previous integrality gap instances, the construction of the integrality gap naturally led us to reason about SDP solutions which are asymmetrical in the sense that not every edge contributes to the solution equally. This is a feature that is necessary for a constraint graph like the Hypercube. Further we augmented the GW SDP with triangle inequalities and showed that our instance does not survive this augmentation. Based on the dense cycle structure of the Hypercube we conjectured that the augmentation is enough for solving Max-2-LIN($\mathbb{Z}_2$) for the Hypercube.

Next we established a connection between MultiCoomodity Cut-Flow gaps and the integrality gap of the augmented SDP. In particular we showed that if a certain class of Multicommodity Flow networks on a graph have a constant cut-flow gap then the the integrality gap of the augmented SDP can be bounded by a constant.

Heading forward there are some clear open problems. First and foremost is our conjecture that GW+indeed is tight on the Hypercube. Further, the connection with Multicommodity Cut-Flow gaps gives a neat way to bound the integrality gap of GW+. The next step would be to find families of graphs for which the condition stated in Corollary 3.3.2 is true. As far as the author is aware the most general family for which this is known are graphs that exclude K_5 as a minor.

Chapter 4

Spectrum of Graph Lifts

In this section we prove Theorem 1.4.1 and 1.4.2. Note since our results for high probabilities require that $n >> d$, we can without loss of generality assume that $\lambda > \sqrt{d}$. This condition can be seen to be true on any graph for which diameter is greater than 4. Therefore if $n >> d$ in particular $n > d^5$ this condition holds. Therefore we will assume in the rest of the section that $\lambda > \sqrt{d}$.

We will in general refer to an event happening with high probability if the event happens with probability $\geq 1 - e^{-\Omega_d(n)}$ where $\Omega_d(n)$ is a function growing with n with possibly some polynomial factors of d (this function in our case would in general be $\frac{n}{d^c}$ for some c). The above definition of high probability makes sense because in construction of Ramanujan Graphs we concern ourselves with fixed small degree d and a growing n.

4.1 Proof Overview

In this section we provide an overview of our proof of Theorems 1.4.1 and 1.4.2

We remind the reader that $2-$ lifts are equivalent to signings of the adjacency and that the new eigenvalues of the lift H are bounded in absolute value by the spectral radius of A_s [refer Section 2.5] and therefore it is enough to provide an upper bound on the spectral radius of A_s that holds with high probability (in n). Note that the spectral radius of A_s is defined as follows:

$$||A_s|| = \max_{\mathbf{x} \in \mathbb{R}^n} \frac{|\mathbf{x}^T A_s \mathbf{x}|}{\|\mathbf{x}\|^2}$$

We start by "rounding" each vector $\mathbf{x}$ to a vector $\mathbf{y}$, such that $\mathbf{y} \in \{\pm 1/2, \pm 1/4 \ldots\}$. It can be shown that $\frac{|\mathbf{y}^T A_s \mathbf{y}|}{\|\mathbf{y}\|^2}$ approximates $\frac{|\mathbf{x}^T A_s \mathbf{x}|}{\|\mathbf{x}\|^2}$ with a loss of at most a factor of 4. We next consider the diadic decomposition of $\mathbf{y}$ to vectors $\mathbf{u}_i \in \{0, \pm 1\}^n$, such that $\mathbf{y} = \sum_i 2^{-i}\mathbf{u}_i$ (refer Definition 2.1.2). Now it is easy to see that

$$|\mathbf{y}^T A_s \mathbf{y}| = |\sum_{i,j} (2^{-i}\mathbf{u}_i)^T A_s (2^{-j}\mathbf{u}_j)|$$

.

Lets consider an individual term $(2^{-i}\mathbf{u}_i)^T A_s (2^{-j}\mathbf{u}_j)$ in this sum. Over random choices of the signing, the product

$(2^{-i}\mathbf{u}_i)^T A_s (2^{-j}\mathbf{u}_j)$ is a sum of independent, zero-mean random variables and a simple application of the Chernoff bound gives that

$$\Pr[|\mathbf{u}_i^T A_s \mathbf{u}_j| \geq \sqrt{d \log d |S(\mathbf{u}_i)||S(\mathbf{u}_j)|}] \leq d^{-(|S(\mathbf{u}_i)|+|S(\mathbf{u}_j)|)} \tag{4.1}$$

Here, for a vector $\mathbf{u}$ we denote its support by $S(\mathbf{u})$.

Application of this simple bound was sufficient to obtain previous results in [BL06], which allowed for a factor of $\log d$ loss. However, in order to obtain our tight results, we are faced with two significant challenges. First, we need our argument to hold with high probability in n, and the probability term $d^{-(|S(\mathbf{u}_i)|+|S(\mathbf{u}_j)|)}$ is clearly not sufficient in the case where the supports of both vectors $\mathbf{u}_i$ and $\mathbf{u}_j$ are small. Second, we cannot afford to lose the factor of $\log d$ in the above bound. To remedy these problems, we separate the sum $|\mathbf{y}^T A_s \mathbf{y}| = |\sum_{i,j}(2^{-i}\mathbf{u}_i)^T A_s (2^{-j}\mathbf{u}_j)|$ into different parts and apply different bounds at each of those parts.

First, we look at vectors $\mathbf{u}_i$ and $\mathbf{u}_j$ with small support, i.e. $|S(\mathbf{u}_i)|, |S(\mathbf{u}_j)| \leq \frac{n}{d^2}$. For such vectors we use a trivial bound and show that their total contribution to the (absolute value of the) sum is less than $\lambda\|\mathbf{y}\|^2$.

Second, we look at the remaining part which consists of terms in which at least one of the $\mathbf{u}_i, \mathbf{u}_j$ has large ($> n/d^2$) support. In order to avoid the $\log d$ factor loss, we need to further separate this remaining sum into parts. One part contains the set of all (i, j) such that at least one of the three guarantees holds.

- $|j - i| > \frac{1}{2}\log d$
- $|S(\mathbf{u}_i)| > E(S(\mathbf{u}_j), V \setminus S(\mathbf{u}_j))$
- $|S(\mathbf{u}_j)| > E(S(\mathbf{u}_i), V \setminus S(\mathbf{u}_i))$

Here, for any two sets of nodes A, B we denote by $E(A, B)$ the number of edges with one endpoint in A and one endpoint in B. The last two cases represent the event where the support of one of the vectors $\mathbf{u}_i$ or $\mathbf{u}_j$ is larger than the total number of edges that leave the support of the other. We show, by using again a trivial bound, that the total contribution to the (absolute value of the) sum from terms that fall into one of the three cases above is no more than $\sqrt{d}\|\mathbf{y}\|^2$. We note that both of the trivial bounds that we have used so far are non-probabilistic.

For the part of the sum that remains we need to employ a tighter bound on the deviation of the zero mean quantity $|\mathbf{u}_i^T A_s \mathbf{u}_j|$. As noted before, $\mathbf{u}_i^T A_s \mathbf{u}_j$ is a sum of independent variables whose total number is at most $E(S(\mathbf{u}_i), S(\mathbf{u}_j))$. We upper bound $E(S(\mathbf{u}_i), S(\mathbf{u}_j))$ by $d|S(\mathbf{u}_i)||S(\mathbf{u}_j)|/n + \lambda\sqrt{|S(\mathbf{u}_i)||S(\mathbf{u}_j)|}$, using the Expander Mixing Lemma 2.4.3 and separate the following cases depending upon which of the two terms in EML is dominating i.e

Case 1: $\lambda\sqrt{|S(\mathbf{u}_i)||S(\mathbf{u}_j)|} \leq d|S(\mathbf{u}_i)||S(\mathbf{u}_j)|/n \Rightarrow E(S(\mathbf{u}_i), S(\mathbf{u}_j)) \leq 2d|S(\mathbf{u}_i)||S(\mathbf{u}_j)|/n$

Case 2: $\lambda\sqrt{|S(\mathbf{u}_i)||S(\mathbf{u}_j)|} \geq d|S(\mathbf{u}_i)||S(\mathbf{u}_j)|/n \Rightarrow E(S(\mathbf{u}_i), S(\mathbf{u}_j)) \leq 2\lambda\sqrt{|S(\mathbf{u}_i)||S(\mathbf{u}_j)|}$.

For **Case 1** we prove that with probability at least $1 - e^{-\Omega(\frac{n}{d^2})}$ we have for each relevant term of the sum:

$$|\mathbf{u}_i^T A_s \mathbf{u}_j| \le 8\sqrt{\lambda\sqrt{|S(\mathbf{u}_i)||S(\mathbf{u}_j)|}|S(\mathbf{u}_j)|\log(\frac{2d|S(\mathbf{u}_i)|}{|S(\mathbf{u}_j)|})} \tag{4.2}$$

The quantity $\sqrt{\lambda\sqrt{|S(\mathbf{u}_i)||S(\mathbf{u}_j)|}|S(\mathbf{u}_j)|\log(\frac{2d|S(\mathbf{u}_i)|}{|S(\mathbf{u}_j)|})}$ is chosen such that the term $\lambda\sqrt{|S(\mathbf{u}_i)||S(\mathbf{u}_j)|}$ cancels out the term in the denominator which appears in the probability guarantee we get from the Chernoff bound and the term $|S(\mathbf{u}_j)|\log(\frac{2d|S(\mathbf{u}_i)|}{|S(\mathbf{u}_j)|})$ allows us to apply the union bound.

Case 2 is slightly more complicated than **Case 1**, as we need to consider multiple terms $|\sum_i \mathbf{u}_i A_s \mathbf{u}_j|$ for a fixed $\mathbf{u}_j$. If instead we considered each term separately, then for each $\mathbf{u}_j$ the term $|S(\mathbf{u}_i)|$ would get counted $\log d$ times, which would result in a $\log d$ factor loss we cannot afford. Instead we show that with probability at least $1 - e^{-\Omega(\frac{n}{d^2})}$ we have for each relevant $\mathbf{u}_j$:

$$|\sum_i \mathbf{u}_i^T A_s \mathbf{u}_j| \le 8\sqrt{1/n * d|S(\mathbf{u}_j)|^2(\sum_i |S(\mathbf{u}_i)|2^{2i})\log(\frac{2n}{|S(\mathbf{u}_j)|})} \tag{4.3}$$

Combining these two bounds we prove the following lemma which bounds the total contribution of the sum of terms that remain after removing vectors with small supports.

Lemma. *Let $\mathbf{u}_1, \mathbf{u}_2, \ldots \in \{0, \pm 1\}^n$, $\mathbf{v}_1, \mathbf{v}_2 \ldots \in \{0, \pm 1\}^n$ be two families of vector sets such that for all (i, j), $S(\mathbf{u}_i) \cap S(\mathbf{u}_j) = S(\mathbf{v}_i) \cap S(\mathbf{v}_j) = \emptyset$ and either for all i, $|S(\mathbf{v}_i)| > \frac{n}{d^2}$ or for all i, $|S(\mathbf{u}_i)| > \frac{n}{d^2}$. Let A_s be a random signing matrix. The following holds with high probability over random choices of signing.*

$$|\sum_{i \le j}(2^{-i} * \mathbf{u}_i^T)A_s(2^{-j} * \mathbf{v}_j)| \le \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d}))\sum_i |S(\mathbf{u}_i)|2^{-2i} + (\frac{\lambda}{5} + \mathcal{O}(\sqrt{d}))\sum_j |S(\mathbf{v}_j)|2^{-2j} \tag{4.4}$$

In the next section we combine the bound obtained by the above lemma and the bound on vectors with small support to prove Theorem 1.4.1

For the proof of Theorem 1.4.2, we follow a similar path. However, we are no longer able to exploit the relation between the spectrum of lifts and the spectral radius of signed matrices. Instead, as presented in section 2.6, we find a novel complete characterization of the spectrum of shift k-lifts by the spectrum of certain k matrices which can be seen as a generalization of the signed matrix.

4.2 Proof of Theorem 1.4.1

We restate Theorem 1.4.1 here a with slightly stronger statement which implies Theorem 1.4.1

Theorem 4.2.1. *Let G be a d -regular graph with non-trivial eigenvalues at most λ in absolute value, and H be a (uniformly random) 2-lift of G. Let λ_{new} be the largest in absolute value new eigenvalue of H. Then*

$$\lambda_{new} \leq 4\lambda + \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d}))$$

with probability $1 - e^{-\Omega(n/d^2)}$.

Note that the above Theorem implies Theorem 1.4.1 for any value of $\lambda > \sqrt{d}$. We will present the proof of Theorem 4.2.1 assuming the following lemma whose proof we defer till Section 4.4.

Lemma 4.2.2. *Let $\mathbf{u}_1, \mathbf{u}_2, \ldots \in \{0, \pm 1\}^n$, $\mathbf{v}_1, \mathbf{v}_2 \ldots \in \{0, \pm 1\}^n$ be any two families of vector sets such that for all $(i, j), S(\mathbf{u}_i) \cap S(\mathbf{u}_j) = S(\mathbf{v}_i) \cap S(\mathbf{v}_j) = \emptyset$ (To remind the reader $S(\mathbf{v})$ is the support of $\mathbf{v}$). Additionally either of the following conditions hold for all the system vectors*

- *either for all i, $|S(\mathbf{v}_i)| > \frac{n}{d^2}$*
- *or for all i, $|S(\mathbf{u}_i)| > \frac{n}{d^2}$.*

Let A_s be a random signing of A. The following holds with probability at aleast $1 - e^{-\Omega(n/d^2)}$ over random choices of signings.

$$|\sum_{i \leq j} (2^{-i} * \mathbf{u}_i^T) A_s (2^{-j} * \mathbf{v}_j)| \leq \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d})) \sum_i |S(\mathbf{u}_i)| 2^{-2i} + (\frac{\lambda}{5} + \mathcal{O}(\sqrt{d})) \sum_j |S(\mathbf{v}_j)| 2^{-2j} \quad (4.5)$$

Also if we know that for all (i, j) $|S(u_i)| \geq |S(v_j)|$, then the following holds with probability at least $1 - e^{-\Omega(n/d^2)}$

$$|\sum_{i \leq j} (2^{-i} * \mathbf{u}_i^T) A_s (2^{-j} * \mathbf{v}_j)| \leq \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d}))(\sum_i |S(\mathbf{u}_i)| 2^{-2i} + \sum_j |S(\mathbf{v}_j)| 2^{-2i}) \quad (4.6)$$

Proof:[Proof of Theorem 4.2.1] For any given vector $\mathbf{x} \in \mathbb{R}^n$ let $R(\mathbf{x}) = \frac{|\mathbf{x}^T A_s \mathbf{x}|}{\mathbf{x}^T \mathbf{x}}$. We know that $\lambda_{new} = ||A_s|| = \max_{\mathbf{x} \in \mathbb{R}^n} R(\mathbf{x})$. To prove an upper bound on λ_{new} we will prove that the quantity $R(\mathbf{x})$ is bounded for all $\mathbf{x}$. In particular we will show that for all $\mathbf{x}$, $|\mathbf{x}^T A_s \mathbf{x}| \leq 4 * (\lambda + \mathcal{O}(\sqrt{d}))\mathbf{x}^T \mathbf{x}$ with high probability. Note that due to scaling we can look at only those vectors $\mathbf{x}$ for which $\|\mathbf{x}\|_\infty \leq 1/2$.

Now given any vector $\mathbf{x}$ we will first obtain its discretized form $\mathbf{y} \in \{\pm 1/2, \pm 1/4, \ldots\}^n$ as promised by Lemma 2.7.1. Note that $|\mathbf{x}^T A_s \mathbf{x}| \leq |\mathbf{y}^T A_s \mathbf{y}|$ and $|\mathbf{y}^2| \leq 4|\mathbf{x}^2|$. We will prove an upper bound on $|\mathbf{y}^T A_s \mathbf{y}|$.

Consider the diadic decomposition of $\mathbf{y} = \{2^{-i}\mathbf{u}_i\}$ (refer Definition 2.1.2)where $i \geq 1$ and $\mathbf{u}_i \in \{-1, 0, 1\}^n$. Partition the vectors $\{\mathbf{u}_i\}$ into two sets A and B such that $A = \{\mathbf{u}_i \big| |S(\mathbf{u}_i)| \leq \frac{n}{d^2}\}$ and $B = \{\mathbf{u}_i \big| |S(\mathbf{u}_i)| > \frac{n}{d^2}\}$. Let $\mathbf{y}_A = \sum_{i: \mathbf{u}_i \in A} 2^{-i}\mathbf{u}_i$ and $\mathbf{y}_B = \sum_{i: \mathbf{u}_i \in B} 2^{-i}\mathbf{u}_i$. Note that $\mathbf{y} = \mathbf{y}_A + \mathbf{y}_B$ and $|\mathbf{y}|^2 = |\mathbf{y}_A|^2 + |\mathbf{y}_B|^2 = \sum_i 2^{-2i}|S(\mathbf{u}_i)|$

Now we have that $|\mathbf{y}^T A_s \mathbf{y}| \leq |\mathbf{y}_A^T A_s \mathbf{y}_B| + 2|\mathbf{y}_A^T A_s \mathbf{y}_B| + |\mathbf{y}_B^T A_s \mathbf{y}_B|$. We will now consider each part of the above summation separately.

Part 1 - $|\mathbf{y}_A^T A_s \mathbf{y}_A|$ Lets consider $|\mathbf{y}_A^T A_s \mathbf{y}_A|$ first. Note that

$$|\mathbf{y}_A^T A_s \mathbf{y}_A| \leq \mathbf{y}_A'^T A \mathbf{y}_A'$$

where $\mathbf{y}_A'$ is defined as the vector obtained by making each coordinate of $\mathbf{y}_A$ positive. Note the use of A in the RHS of the above inequality. Let J be the $n \times n$ matrix with entries equal to 1. Therefore

$$\begin{aligned}
\mathbf{y}_A'^T A \mathbf{y}_A' &= \mathbf{y}_A'^T (A - \frac{d}{n} J) \mathbf{y}_A' + \mathbf{y}_A'^T \left(\frac{d}{n} J\right) \mathbf{y}_A' \\
&\leq \lambda \|\mathbf{y}_A'\|^2 + \mathbf{y}_A'^T \left(\frac{d}{n} J\right) \mathbf{y}_A'
\end{aligned}$$

Lets look at the term $\mathbf{y}_A'^T \left(\frac{d}{n} J\right) \mathbf{y}_A'$. Let the diadic decomposition of $\mathbf{y}_A' = \{2^{-i}\mathbf{u}_i\}$. Note that since $|S(\mathbf{u}_i)| \leq \frac{n}{d^2}$

$$\begin{aligned}
\mathbf{y}_A'^T \left(\frac{d}{n} J\right) \mathbf{y}_A' &= 2 \sum_i \sum_{j \geq i} \frac{d}{n} 2^{-i} |S(\mathbf{u}_i)| 2^{-j} |S(\mathbf{u}_j)| \\
&\leq 2 \sum_i \frac{1}{d} 2^{-2i} |S(\mathbf{u}_i)| \sum_{j \geq i} 2^{i-j} \\
&\leq \frac{4}{d} \|\mathbf{y}_A'\|^2
\end{aligned}$$

Part 2 - $|\mathbf{y}_B^T A_s \mathbf{y}_B|$ Consider the diadic decomposition of $\mathbf{y}_B = \{2^{-i}\mathbf{u}_i\}$. We have that $|\mathbf{y}_B^T A_s \mathbf{y}_B| = |2 \sum_{i \leq j} (2^{-i}\mathbf{u}_i) A_s (2^{-j}\mathbf{u}_j)|$. Now since $|S(\mathbf{u}_i)| > \frac{n}{d^2}$, we can now apply Lemma 4.2.2 and we get that

$$\begin{aligned}
\frac{1}{2} |\mathbf{y}_B^T A_s \mathbf{y}_B| &\leq \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d})) \sum_i |S(\mathbf{u}_i)| 2^{-2i} + (\frac{\lambda}{5} + \mathcal{O}(\sqrt{d})) \sum_j |S(\mathbf{u}_j)| 2^{-2j} \\
&\leq \left(\frac{\lambda}{5} + \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d}))\right) \|\mathbf{y}_B\|^2
\end{aligned}$$

Part 3 - $|\mathbf{y}_A^T A_s \mathbf{y}_B|$ Consider the diadic decomposition of $\mathbf{y}_A = \{2^{-i}\mathbf{u}_i\}$ and $\mathbf{y}_B = \{2^{-j}\mathbf{v}_j\}$. Therefore $|\mathbf{y}_A^T A_s \mathbf{y}_B| = |\sum_{i \leq j} (2^{-i}\mathbf{u}_i) A_s (2^{-j}\mathbf{v}_j) + \sum_{i < j} (2^{-i}\mathbf{v}_i) A_s (2^{-j}\mathbf{u}_j)|$. Now since $|S(\mathbf{v}_i)| > \frac{n}{d^2}$(by definition) and for all (i, j) $|S(\mathbf{v}_i)| \geq$

$|S(\mathbf{u}_j)|$, we can now apply lemma 4.2.2 and we get that

$$
\begin{aligned}
|\mathbf{y}_A^T A_s \mathbf{y}_B| &\leq |\sum_{i\leq j}(2^{-i}\mathbf{u}_i)A_s(2^{-j}\mathbf{v}_j)| + |\sum_{i<j}(2^{-i}\mathbf{v}_i)A_s(2^{-j}\mathbf{u}_j)| \\
&\leq \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d}))\sum_i |S(\mathbf{u}_i)|2^{-2i} + (\frac{\lambda}{5} + \mathcal{O}((\sqrt{d}))\sum_j |S(\mathbf{v}_j)|2^{-2j} \\
&\quad + \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d}))\sum_j (|S(\mathbf{v}_j)|2^{-2j} + |S(\mathbf{u}_j)|2^{-2j}) \\
&\leq \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d}))\mathbf{y}_A\|^2 + (\frac{\lambda}{5} + (\mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d})))\|\mathbf{y}_B\|^2
\end{aligned}
$$

Putting it all together Putting the above inequalities together we get that

$$
\begin{aligned}
|\mathbf{y}^T A_s \mathbf{y}| &\leq (\lambda + \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d})))\|\mathbf{y}_A\|^2 + (\frac{4\lambda}{5} + \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d})))\mathbf{y}_B\|^2 \\
&\leq \lambda + \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d}))\|\mathbf{y}\|^2.
\end{aligned}
$$

Therefore we have

$$
\begin{aligned}
|\mathbf{x}^T A_s \mathbf{x}| &\leq |\mathbf{y}^T A_s \mathbf{y}| \\
&\leq (\lambda + \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d})))\|\mathbf{y}\|^2 \\
&\leq 4 * (\lambda + \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d})))\|\mathbf{x}\|^2
\end{aligned}
$$

□

4.3 Generalization to Shift Lifts

In this section we present the proof of Theorem 1.4.2. Note that Theorem 1.4.2 is a generalization of Theorem 1.4.1 to the case of k-shift lifts. As in the case of 2-lifts we will prove the following stronger statement which implies Theorem 1.4.2 (assuming $\lambda > \sqrt{d}$).

Theorem 4.3.1. *Let G be a d -regular graph with non-trivial eigenvalues at most λ in absolute value, and H be a (uniformly random) shift k-lift of G. Let λ_{new} be the largest in absolute value new eigenvalue of H. Then*

$$\lambda_{new} \leq 16(\lambda + \mathcal{O}(\max(\sqrt{\lambda \log d}, \sqrt{d})))$$

with probability at least $1 - k * e^{-\Omega(n/d^c)}$.

To prove the above theorem, we state a slightly general form of Theorem 4.2.1

Theorem 4.3.2. *Let G be a d -regular graph with non-trivial eigenvalues at most λ in absolute value with adjacency matrix A. Let A' be a random real matrix each of whose entries A'_{ij} is a random variable with the following properties*

- $\forall i, j,\ E[A'_{ij}] = 0$ *and* $|A'_{ij}| \leq 1$ *with probability 1*
- $\forall i, j$ *if* $A_{ij} = 0$ *then* $A'_{ij} = 0$ *with probability 1*

Then with probability at least $1 - e^{-\Omega(n/d^2)}$

$$||A_s|| \leq 4 * (\lambda + \mathcal{O}(\max \sqrt{\lambda \log(d)}, \sqrt{d}))$$

The proof of the above Theorem is exactly the same as the proof of Theorem 4.2.1. The only difference is that every entry in A' may now have a smaller magnitude but that does not affect any of the arguments in the proof of Theorem 4.2.1. Using theorem 4.3.2, we will now prove theorem 4.3.1.

Proof: Note that for a shift lift $\lambda_{new} = \max\limits_{\omega, \omega \neq 1} ||A_s(\omega)||$ where ω is the k^{th} root of unity (To remind the reader $A_s(\omega)$ is the extended signed adjacency matrix for k-shift lifts. Refer Section 2.5). Therefore,

$$P(\lambda_{new} \geq 16(\lambda + \mathcal{O}(\max \sqrt{\lambda \log(d)}, \sqrt{d}))) \leq \sum_{\omega, \omega \neq 1} P(||A_s(\omega)|| \geq 16(\lambda + \mathcal{O}(\max \sqrt{\lambda \log(d)}, \sqrt{d})))$$

Therefore if we can show that for a fixed ω

$$P(||A_s(\omega)|| \geq 16(\lambda + \mathcal{O}(\max \sqrt{\lambda \log(d)}, \sqrt{d}))) \leq e^{-\Omega(n/d^2)}$$

By union bound we have that

$$P(\lambda_{new} \geq 16(\lambda + \mathcal{O}(\max \sqrt{\lambda \log(d)}, \sqrt{d}))) \leq (k-1)e^{-\Omega(n/d^2)}$$

which implies the theorem. Therefore it is enough to show that for a fixed ω,

$$P(||A_s(\omega)|| \geq 16(\lambda + \mathcal{O}(\max \sqrt{\lambda \log(d)}, \sqrt{d}))) \leq e^{-\Omega(n/d^2)}$$

The spectral radius of $A_s(\omega) = \max\limits_{\mathbf{x} \in \mathbb{C}^n} \frac{|\mathbf{x}^* A_s(\omega) \mathbf{x}|}{|\mathbf{x}^* \mathbf{x}|}$. We first split the vector $\mathbf{x}$ and matrix $A_s(\omega)$ into its real and imaginary parts. Let $\mathbf{x} = \mathbf{x}_1 + i\mathbf{x}_2$ and $A_s(\omega) = A_s^1(\omega) + iA_s^2(\omega)$ where $\mathbf{x}_1, \mathbf{x}_2$ are real vectors and $A_s^1(\omega)$ and $A_s^2(\omega)$ are real matrices. By theorem 4.3.2 we have that with high probability,

$$|\mathbf{x}_i^T A_s^k(\omega) \mathbf{x}_j| \leq 4(\lambda + \mathcal{O}(\max \sqrt{\lambda \log(d)}, \sqrt{d}))(||\mathbf{x}_i|| ||\mathbf{x}_j||)$$

Therefore

$$
\begin{aligned}
|\mathbf{x}^* A_s(\omega)\mathbf{x}| &\leq \sum_{i,j,k\in\{1,2\}} |\mathbf{x}_i^T A_s^k(\omega)\mathbf{x}_j| \\
&\leq 8(\lambda + \mathcal{O}(\max\sqrt{\lambda\log(d)}, \sqrt{d}))(||\mathbf{x}_1||^2 + ||\mathbf{x}_2||^2) + 16(\lambda + \mathcal{O}(\max\sqrt{\lambda\log(d)}, \sqrt{d}))||\mathbf{x}_1||||\mathbf{x}_2|| \\
&\leq 16(\lambda + \mathcal{O}(\max\sqrt{\lambda\log(d)}, \sqrt{d}))(||\mathbf{x}_1||^2 + ||\mathbf{x}_2||^2) \\
&\leq 16(\lambda + \mathcal{O}(\max\sqrt{\lambda\log(d)}, \sqrt{d}))|\mathbf{x}^*\mathbf{x}|
\end{aligned}
$$

Therefore we have that $\forall\omega$,

$$||A_s(\omega)|| = \max_{x\in\mathbb{C}^n} \frac{|x^* A_s(\omega)x|}{|x^* x|} \leq 16(\lambda + \mathcal{O}(\max\sqrt{\lambda\log(d)}, \sqrt{d}))$$

w.p. greater than $1 - e^{-\Omega(\frac{n}{d^2})}$ which proves the theorem. □

4.4 Proof of Lemma 4.2.2

In this section we prove the main technical lemma of our lifts result which is Lemma 4.2.2. Our proof is based on the following two probabilistic lemmas. Note that these lemmas are the places where we use the argument of high probability. So once the conditions guaranteed by these lemmas are satisfied then the rest of the proof follows and we ensure that these conditions are met by a random lift with high probability.

Lemma 4.4.1. *For a random 2-lift, let A_s be the signed adjacency matrix the following property holds with probability $1 - e^{-\Omega(\frac{n}{d^2})}$*

Let $\mathbf{u}, \mathbf{v} \in \{0, \pm 1\}^n$ s.t. $|S(\mathbf{u})| \leq |S(\mathbf{v})| \leq d|S(\mathbf{u})|$, $S(\mathbf{v}) > \frac{n}{d^2}$ and $\frac{d}{\lambda}\sqrt{|S(\mathbf{u})||S(\mathbf{v})|} < n$. Then,

$$|\mathbf{u}^T A_s \mathbf{v}| \leq 8\sqrt{\lambda\sqrt{|S(\mathbf{u})||S(\mathbf{v})|}|S(\mathbf{v})|\log(\frac{2d|S(\mathbf{u})|}{|S(\mathbf{v})|})} \tag{4.7}$$

Lemma 4.4.2. *For a random 2-lift, let A_s be the signed adjacency matrix the following property holds with probability $1 - e^{-\Omega(\frac{n}{d^2})}$*

Let $v, u_0, u_1, \cdots \in \{0, \pm 1\}^n$ s.t. $|S(\mathbf{v})| \geq 2^{2i}|S(\mathbf{u}_i)|$, and $\frac{d}{\lambda}\sqrt{|S(\mathbf{u}_i)||S(\mathbf{v})|} \geq n$. Let $\mathbf{u} = \sum_i \mathbf{u}_i.2^i$. Then,

$$|\mathbf{v}^T A_s \mathbf{u}| \leq 8\sqrt{\frac{d}{n}|S(\mathbf{v})|^2(\sum_i |S(\mathbf{u}_i)|2^{2i})\log(\frac{2n}{|S(\mathbf{v})|})} \tag{4.8}$$

Firstly, using lemma 4.4.1 and 4.4.2, we will prove lemma 4.2.2 and then prove these lemmas independently.

4.4.1 Proof of lemma 4.2.2

Proof: For the ease of presentation in this section we denote $|S(\mathbf{u}_i)|$ with y_i and $|S(\mathbf{v}_j)|$ with z_j Since conditions 4.7 and 4.8 of lemma 4.4.1 and 4.4.2 hold true w.h.p. we can assume that both of the conditions hold for the matrix A_s. To prove the lemma we need to bound the quantity

$$X = |\sum_{i \leq j} (2^{-i}\mathbf{u}_i) A_s (2^{-j}\mathbf{v}_j)|$$

We will prove an upperbound on this quantity by partitioning the sum into multiple parts and proving upper bounds for all those parts. In the rest of the section we use C_I to denote sets of tuples (i, j) of integers (that satisfy some conditions), and we use X_I to denote sums of the form $|\sum_{(i,j) \in C_I} \mathbf{u}_i^T A_s \mathbf{v}_j|$.

We first partition the sum into two parts X_1 and X_2 where we show that the part X_2 can be easily bound by using a trivial bound of $|\mathbf{u}_i^T A_s \mathbf{v}_j| \leq d \min(y_i, z_j)$ on each individual term of X_2.

$$\begin{aligned}
C_1 &= \{(i,j) \big| (i \leq j < i + \frac{1}{2}\log(d)) \wedge (\max(y_i, z_j) < d \min(y_i, z_j))\} \\
C_2 &= \bar{C_1} = \{(i,j) \big| (j \geq i + \frac{1}{2}\log(d)) \vee (y_i \geq d z_j) \vee (z_j \geq d y_i)\} \\
X_1 &= |\sum_{(i,j) \in C_1} 2^{-i-j} \mathbf{u}_i^T A_s \mathbf{v}_j| \\
X_2 &= |\sum_{(i,j) \in C_2} 2^{-i-j} \mathbf{u}_i^T A_s \mathbf{v}_j| \\
X &\leq X_1 + X_2
\end{aligned}$$

We further analyze the sum X_1 by breaking it into two parts X_3 and X_4 such that X_3 contains the part of the sum where $y_i \geq z_j$ and X_4 contains the part of the sum where $y_i < z_j$.

$$\begin{aligned}
C_3 &= C_1 \cap \{(i,j) \big| (y_i \geq z_j)\} \\
C_4 &= C_1 \cap \bar{C_3} = C_1 \cap \{(i,j) \big| (y_i < z_j)\} \\
X_3 &= |\sum_{(i,j) \in C_3} 2^{-i-j} \mathbf{u}_i^T A_s \mathbf{v}_j| \\
X_4 &= |\sum_{(i,j) \in C_4} 2^{-i-j} \mathbf{u}_i^T A_s \mathbf{v}_j| \\
X_1 &\leq X_3 + X_4
\end{aligned}$$

We show the required bound on X_3 then further analyze the sum X_4. As guaranteed by Expander Mixing lemma, number of edges between $S(\mathbf{u}_i)$ and $S(\mathbf{v}_j)$ is bounded by $\frac{d y_i z_j}{n} + \lambda\sqrt{y_i z_j}$. As indicated in the proof overview we split our analysis based on which of the two terms in the right hand side of the above expression is the dominating one. Therefore we split X_4 into two cases X_5 and X_6 where X_5 contains the terms where $\frac{dy_iz_j}{n} < \lambda\sqrt{y_iz_j}$ and X_6 contains the terms where $\frac{dy_iz_j}{n} \geq \lambda\sqrt{y_iz_j}$. This separation helps us apply one of the two bounds in Lemma 4.4.1 and Lemma 4.4.2.

$$
\begin{aligned}
C_5 &= C_4 \cap \{(i,j)\big|(\frac{d}{\lambda}\sqrt{y_iz_j} < n)\} \\
C_6 &= C_4 \cap \bar{C}_5 = C_4 \cap \{(i,j)\big|(\frac{d}{\lambda}\sqrt{y_iz_j} \geq n)\} \\
X_5 &= |\sum_{(i,j)\in C_5} 2^{-i-j}\mathbf{u}_i^T A_s \mathbf{v}_j| \\
X_6 &= |\sum_{(i,j)\in C_6} 2^{-i-j}\mathbf{u}_i^T A_s \mathbf{v}_j| \\
X_4 &\leq X_5 + X_6
\end{aligned}
$$

To ease the calculations, we further split the sum X_5 into two parts depending on whether $y_i 2^{-2i}$ is significantly (in terms of λ) greater than $z_j 2^{-2j}$. In this regard we make the following separation of X_5 into X_7 and X_8. In the analysis of X_7 and X_8 we use the bound given by lemma 4.4.2 for individual entries.

$$
\begin{aligned}
C_7 &= C_5 \cap \{(i,j)\big|(y_i 2^{-2i} < \frac{\lambda}{\sqrt{d}} z_j 2^{-2j})\} \\
C_8 &= C_5 \cap \bar{C}_7 = C_5 \cap \{(i,j)\big|(y_i 2^{-2i} \geq \frac{\lambda}{\sqrt{d}} z_j 2^{-2j})\} \\
X_7 &= |\sum_{(i,j)\in C_7} 2^{-i-j}\mathbf{u}_i^T A_s \mathbf{v}_j| \\
X_8 &= |\sum_{(i,j)\in C_8} 2^{-i-j}\mathbf{u}_i^T A_s \mathbf{v}_j| \\
X_5 &\leq X_7 + X_8
\end{aligned}
$$

Similar to the case of X_5 we separate X_6 into two parts X_9 and X_{10} on the sizes of $2^{-2i}y_i$ and $2^{-2j}z_j$. In the analysis of X_9 and X_{10} we use the bound given by lemma 4.4.1 for individual entries.

$$
\begin{aligned}
C_9 &= C_6 \cap \{(i,j)\big|(y_i 2^{-2i} \leq z_j 2^{-2j})\} \\
C_{10} &= C_6 \cap \bar{C}_9 = C_6 \cap \{(i,j)\big|(y_i 2^{-2i} > z_j 2^{-2j})\} \\
X_9 &= |\sum_{C_9} 2^{-i-j}\mathbf{u}_i^T A_s \mathbf{v}_j| \\
X_{10} &= |\sum_{C_{10}} 2^{-i-j}\mathbf{u}_i^T A_s \mathbf{v}_j| \\
X_6 &\leq X_9 + X_{10}
\end{aligned}
$$

These cases are summarized in the following figure

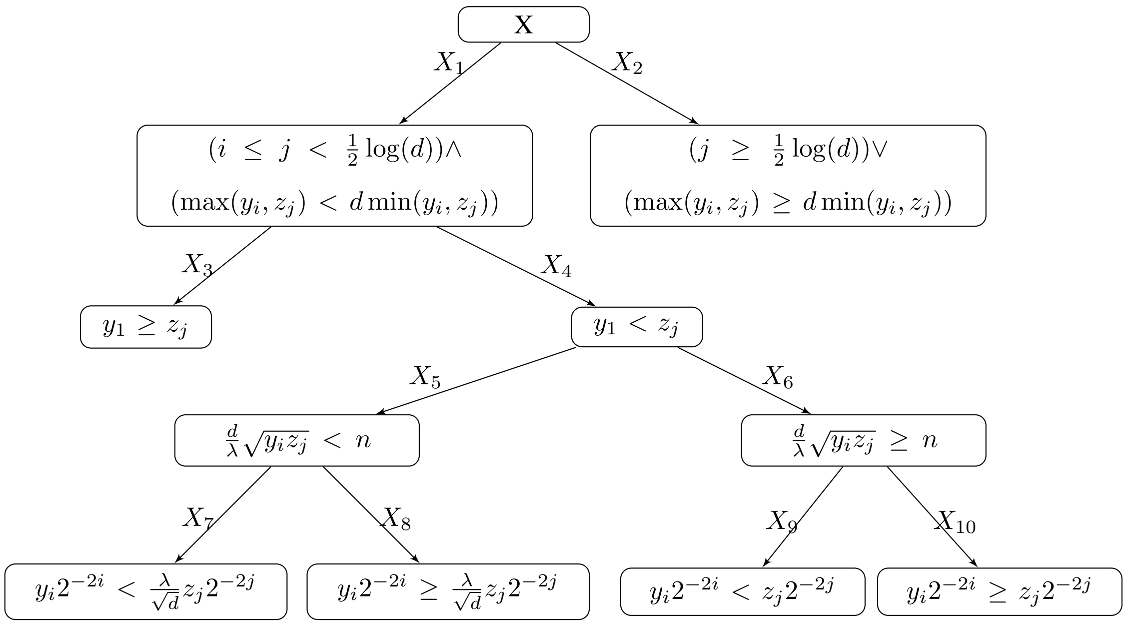

Figure 4.1: Case division for the analysis

We now prove bounds on the leaves $X_2, X_3, X_7, X_8, X_9, X_{10}$ of the above tree

Upper bound on X_2

$$X_2 \leq \Big| \sum_{j \geq i + \frac{1}{2}\log(d)} (2^{-i}\mathbf{u}_i)^T A_s (2^{-j}\mathbf{v}_j) \Big| \qquad (X_2')$$

$$+ \Big| \sum_{\substack{i \leq j < i + \frac{1}{2}\log(d), \\ \max(y_i, z_j) \geq d\min(y_i, z_j)}} (2^{-i}\mathbf{u}_i)^T A_s (2^{-j}\mathbf{v}_j) \Big| \qquad (X_2'')$$

Note that since the number of edges out of any set S is bounded by $d|S|$, we have that $|\mathbf{u}_i^T A \mathbf{v}_j| \leq d\min(y_i.z_j)$ for any $\mathbf{u}_i, \mathbf{v}_j \in \{-1, 0, +1\}^n$. We avoid writing the complete conditions from the sum when otherwise understood.

$$\begin{aligned} X_2' &\leq \sum_i \sum_{j = i + \frac{1}{2}\log(d)}^{\infty} 2^{-i} 2^{-j} |\mathbf{u}_i^T A_s \mathbf{v}_j| \\ &\leq \sum_i \sum_{j = i + \frac{1}{2}\log(d)}^{\infty} 2^{-i} * 2^{-j} d y_i \\ &\leq \mathcal{O}(\sqrt{d}) \sum_i 2^{-2i} y_i \end{aligned}$$

$$\begin{aligned} X_2'' &\leq \sum 2^{-i-j} |\mathbf{u}_i^T A_s \mathbf{v}_j| \\ &\leq \sum 2^{-i-j} max(y_i, z_j) \\ &\leq \sum_{i \leq j < i + \frac{1}{2}\log(d)} 2^{-i-j}(y_i + z_j) \\ &\leq \mathcal{O}(1) \sum_i y_i 2^{-2i} + \mathcal{O}(\sqrt{d}) \sum_j z_j 2^{-2j} \end{aligned}$$

Combining X_2' and X_2'', we get

$$\boldsymbol{X_2 \leq \mathcal{O}(\sqrt{d})(\sum_j z_j 2^{-2j} + \sum_i y_i 2^{-2i})} \tag{4.9}$$

Upper bound on X_3

X_3 is the sum conditioned over the following set of i, j

$$C_3 = \{(i,j) | (i \leq j < i + \frac{1}{2}\log(d)) \wedge (\max(y_i, z_j) \leq d \min(y_i, z_j)) \wedge (y_i \geq z_j)\}$$

If $\frac{d}{\lambda}\sqrt{y_i z_j} \geq n$, then by lemma 4.4.2(substituting $\mathbf{v} = \mathbf{u}_i, \mathbf{u}_0 = \mathbf{v}_j, \mathbf{u}_1 = \mathbf{u}_2, \cdots = \phi$), and the fact that $\frac{y_i}{n}\log(\frac{2n}{y_i}) \leq 2$, we get that

$$|\mathbf{u}_i^T A_s \mathbf{v}_j| \leq \mathcal{O}(\sqrt{d}) y_i$$

And if $\frac{d}{\lambda}\sqrt{y_i z_j} < n$, then by lemma 4.4.1(substituting $\mathbf{u} = \mathbf{v}_j, \mathbf{v} = \mathbf{u}_i$), and the fact that since $y_i \leq n$, we get that

$$|\mathbf{u}_i^T A_s \mathbf{v}_j| \leq \mathcal{O}(\sqrt{\lambda \log(2d)}) y_i$$

Therefore we have that,

$$\begin{aligned} X_3 &\leq \sum_{(i,j) \in C_3} 2^{-i-j} |u_i^T A_s v_j| \\ &\leq \sum_i \sum_{j=i}^{\infty} 2^{-i-j} \mathcal{O}(\max(\sqrt{d}, \sqrt{\lambda \log(d)})) y_i \end{aligned}$$

Therefore we get that

$$\boldsymbol{X_3 \leq \mathcal{O}(\max(\sqrt{d}, \sqrt{\lambda \log(d)})) \sum_i y_i 2^{-2i}} \tag{4.10}$$

Upper bound on X_7

X_7 is the sum conditioned over the following set of i, j

$$C_7 = \left\{(i,j) | (i \leq j) \wedge (j \leq i + \frac{1}{2}\log(d)) \wedge (y_i \leq z_j \leq d y_i) \wedge \boldsymbol{(\frac{d}{\lambda}\sqrt{y_i z_j} < n)} \wedge \left(\boldsymbol{y_i 2^{-2i} < \frac{\lambda}{\sqrt{d}} z_j 2^{-2j}}\right)\right\}$$

We will use lemma 4.4.1(substituting $\mathbf{u} = \mathbf{u}_i$ and $\mathbf{v} = \mathbf{v}_j$) to bound $|\mathbf{u}_i^T A_s \mathbf{v}_j|$. Therefore

$$\begin{aligned}
X_7 &\leq \mathcal{O}(1) \sum_{(i,j)\in C_7} 2^{-i-j} \sqrt{\lambda \sqrt{y_i z_j} z_j \log\left(\frac{2dy_i}{z_j}\right)} \\
&\leq \mathcal{O}(1) \sum_{C_7} \frac{(\lambda)^{3/4}}{d^{1/8}} z_j 2^{-i-j} \sqrt{2^{-j-i} 2 \log\left(\frac{\sqrt{2\sqrt{d}\lambda}}{2^{j-i}}\right)} && \left(y_i 2^{-2i} < \frac{\lambda}{\sqrt{d}} z_j 2^{-2j}\right) \\
&\leq \mathcal{O}(1) \frac{(\lambda)^{3/4}}{d^{1/8}} \sum_j z_j 2^{-2j} \sum_{i=j-\frac{1}{2}\log(d)+1}^{i=j} \sqrt{2^{j-i} \log\left(\frac{\sqrt{2\sqrt{d}\lambda}}{2^{j-i}}\right)} \\
&\leq \mathcal{O}(1) \frac{\lambda^{3/4}}{d^{1/8}} \sqrt{\sqrt{d} \log\left(\sqrt{\frac{2\lambda}{\sqrt{d}}}\right)} \sum_j z_j 2^{-2j} && (\text{by lemma 2.7.2 and } \lambda \geq \sqrt{d}) \\
&= \mathcal{O}(1) \lambda \sqrt{\sqrt{\frac{\sqrt{d}}{\lambda}} \log\left(\sqrt{\frac{2\lambda}{\sqrt{d}}}\right)} \sum_j z_j 2^{-2j}
\end{aligned}$$

It is easy to see that for every $c_1 > 0$, there exists c_2 s.t. $\sqrt{\sqrt{\frac{\sqrt{d}}{\lambda}} \log\left(\sqrt{\frac{2\lambda}{\sqrt{d}}}\right)} \leq (c_1 + c_2\sqrt{d}/\lambda)$ where c_1, c_2 are constants. Hence, we can chose c_1 s.t.

$$\boldsymbol{X_7 \leq (\frac{\lambda}{5} + \mathcal{O}(\sqrt{d})) \sum_j z_j 2^{-2j}} \tag{4.11}$$

Upper bound on X_8

X_8 is the sum conditioned over the following set of i, j

$$C_8 = \left\{(i,j)|(i \leq j) \wedge (j < i + \frac{1}{2}\log(d)) \wedge (y_i \leq z_j < dy_i) \wedge \boldsymbol{(\frac{d}{\lambda}\sqrt{y_i z_j} < n)} \wedge \left(\boldsymbol{y_i 2^{-2i} \geq \frac{\lambda}{\sqrt{d}} z_j 2^{-2j}}\right)\right\}$$

We will again using lemma 4.4.1(substituting $\mathbf{u} = \mathbf{u}_i$ and $\mathbf{v} = \mathbf{v}_j$) to bound $|\mathbf{u}_i^T A_s \mathbf{v}_j|$.

$$\begin{aligned}
X_8 &\leq \mathcal{O}(1) \sum_{(i,j)\in C_8} 2^{-i-j} \sqrt{\lambda \sqrt{y_i z_j} z_j \log\left(\frac{2dy_i}{z_j}\right)} \\
&= \mathcal{O}(1) \sum_{(i,j)\in C_8} 2^{-i-j} \sqrt{\lambda} y_i \sqrt{\left(\frac{z_j}{y_i}\right)^{\frac{3}{2}} \log\left(\frac{2dy_i}{z_j}\right)} \\
&= \mathcal{O}(1) \sum_{(i,j)\in C_8} 2^{-i-j} \frac{d^{3/8}}{\lambda^{1/4}} y_i \sqrt{2^{3j-3i} \log\left(\frac{2\lambda\sqrt{d}}{2^{2j-2i}}\right)} && \left(y_i 2^{-2j} \geq \frac{\lambda}{\sqrt{d}} z_j 2^{-2j}\right)
\end{aligned}$$

Above holds since $x^{\frac{3}{2}} \log\left(\frac{c}{x}\right)$ is increasing if $x \leq \frac{c}{2}$

$$\leq \mathcal{O}(1) \sum_i \frac{d^{3/8}}{\lambda^{1/4}} y_i 2^{-2i} \sum_{i=j}^{j=i+\frac{1}{2}\log(d)-1} \sqrt{2^{j-i} 2 \log\left(\frac{\sqrt{2\lambda\sqrt{d}}}{2^{j-i}}\right)}$$

$$\leq \mathcal{O}(1) \sum_i \frac{d^{3/8}}{\lambda^{1/4}} y_i 2^{-2i} \sqrt{\sqrt{d} \log\left(\sqrt{2\frac{\lambda}{\sqrt{d}}}\right)} \qquad \text{(by lemma 2.7.2)}$$

$$\leq \mathcal{O}(1) \sum_i d^{\frac{1}{2}} y_i 2^{-2i} \sqrt{\sqrt{\frac{\sqrt{d}}{\lambda}} \log\left(\sqrt{2\frac{\lambda}{\sqrt{d}}}\right)}$$

Using $\lambda \geq \sqrt{d}$, we have $\sqrt{\frac{\sqrt{d}}{\lambda} \log\left(\sqrt{\frac{2\lambda}{\sqrt{d}}}\right)} \leq \mathcal{O}(1)$

$$\boldsymbol{X_8 \leq \mathcal{O}(\sqrt{d}) \sum_i y_i 2^{-2i}} \tag{4.12}$$

Upper bound on X_9

X_9 is the sum conditioned over the following set of i, j

$$C_9 = \left\{(i,j) | (i \leq j) \wedge (j < i + \frac{1}{2}\log(d)) \wedge (y_i \leq z_j \leq d y_i) \wedge \boldsymbol{(\frac{d}{\lambda}\sqrt{y_i z_j} \geq n)} \wedge \boldsymbol{\left(y_i 2^{-2i} < z_j 2^{-2j}\right)}\right\}$$

In this case, we will use lemma 4.4.2 to bound $|\sum_{i=j-1/2\log(d)}^{i=j} 2^{-i+j} \mathbf{u}_i^T A_s \mathbf{u}_j|$. We group $\mathbf{v}_j$ according to support sizes and then sum them together. For $c = 0, 1, 2, \ldots, \log(n)$, define

$$\begin{aligned} J_c &= \{j | \frac{n}{2^c} \leq z_j < \frac{2n}{2^c}\} \\ j_c &= \min(J_c) \end{aligned}$$

Therefore we can get the following bound

$$X_9 \leq \mathcal{O}(1) \sum_j 2^{-2j} \sqrt{\frac{d z_j^2}{n} \sum_{i=j-\frac{1}{2}\log(d)}^{i=j} y_i 2^{-2i+2j} \log\left(\frac{2n}{z_j}\right)}$$

$$\leq \mathcal{O}(\sqrt{d}) \sum_j \sqrt{2^{-2j} \frac{z_j^2}{n} \log\left(\frac{2n}{z_j}\right) \sum_{i=j-\frac{1}{2}\log(d)}^{i=j} y_i 2^{-2i}}$$

$$\leq \mathcal{O}(\sqrt{d}) \sum_c \sum_{j \in J_c} \sqrt{4n 2^{-2j-2c} \log(2 \cdot 2^c) \sum_{i=j-1/2\log(d)+1}^{i=j} y_i 2^{-2i}} \qquad \left(\frac{n}{2^c} \leq z_j < \frac{2n}{2^c}\right)$$

$$\leq \mathcal{O}(\sqrt{d}) \sum_c \sum_{j \in J_c} \frac{1}{2} \left(\left(4n.2^{-j-j_c-c}\right) + \left(2^{-j+j_c-c} \log(2 \cdot 2^c) \sum_{i=j-1/2\log(d)+1}^{i=j} y_i 2^{-2i}\right)\right) \qquad (A.M. \geq G.M.)$$

$$\leq \mathcal{O}(\sqrt{d})\sum_c \sum_{j\in J_c}\left(\frac{1}{2}\left(4n.2^{-j-j_c-c}\right)+\sqrt{d}\sum_c \sum_{j\in J_c}\sum_i 2^{-j+j_c-c}\log(2\cdot 2^c)y_i 2^{-2i}\right)$$

$$\leq \mathcal{O}(\sqrt{d})\left(4\sum_c \frac{n}{2^c}\sum_{j\in J_c}2^{-j-j_c}+\sum_i y_i 2^{-2i}\sum_c \frac{\log(2\cdot 2^c)}{2^c}\sum_{j\in J_c}2^{-j+j_c}\right)$$

Summing up the geometric sums, we get

$$\boldsymbol{X_9 \leq \mathcal{O}(\sqrt{d})\left(\sum_j z_j 2^{-2j}+\sum_i y_i 2^{-2i}\right)} \tag{4.13}$$

Upper bound on X_{10}

X_{10} is the sum conditioned over the following set of i, j

$$C_{10}=\left\{(i,j)|(i\leq j)\wedge(j\leq i+\frac{1}{2}\log(d))\wedge(y_i\leq z_j\leq dy_i)\wedge\boldsymbol{(\frac{d}{\lambda}\sqrt{y_i z_j}\geq n)}\wedge\boldsymbol{\left(y_i 2^{-2i}\geq z_j 2^{-2j}\right)}\right\}$$

We divide X_{10} into two parts depending on the value of i and j.

$$X_{10}\leq |\sum_{(i,j)\in C_{11}}2^{-i-j}\mathbf{u}_i^T A_s \mathbf{v}_j| \tag{X_{11}}$$

$$+|\sum_{(i,j)\in C_{12}}2^{-i-j}\mathbf{u}_i^T A_s \mathbf{v}_j| \tag{X_{12}}$$

$$C_{11}=\{(i,j)|(i,j)\in C_{10},(j<i+\frac{1}{2}\log(n/y_i))\}$$

$$C_{12}=\{(i,j)|(i,j)\in C_{10},(j\geq i+\frac{1}{2}\log(n/y_i))\}$$

First we analyze X_{11}. We use lemma 4.4.2(substituting $\mathbf{u}_0=\mathbf{u}_i, \mathbf{v}=\mathbf{v}_j, \mathbf{u}_1=\mathbf{u}_2=\cdots=\emptyset$) for bounding $|\mathbf{u}_i^T A_s \mathbf{v}_j|$.

$$X_{11}\leq \sum_{(i,j)\in C_{11}}2^{-i-j}|\mathbf{u}_i^T A_s \mathbf{v}_j|$$

$$\leq \mathcal{O}(1)\sum_{(i,j)\in C_{11}}2^{-i-j}\sqrt{\frac{dy_i z_j^2}{n}\log\left(\frac{2n}{z_j}\right)}$$

$$\leq \mathcal{O}(\sqrt{d})\sum_{(i,j)\in C_{11}}2^{-2i}y_i\sqrt{2^{-2j+2i}\frac{1}{ny_i}z_j^2\log\left(\frac{2n}{z_j}\right)}$$

$$\leq \mathcal{O}(\sqrt{d})\sum_i y_i 2^{-2i}\sum_{j=i}^{j=i+\frac{1}{2}\log(\frac{n}{y_i})}\sqrt{\frac{y_i 2^{2j-2i}}{n}\log\left(\frac{2n}{y_i 2^{2j-2i}}\right)} \qquad (z_j 2^{-2j}\leq y_i 2^{-2i})$$

Above holds because $x^2 \log\left(\frac{c}{x}\right)$ is increasing function if $x \leq \frac{c}{2}$

$$\leq \boldsymbol{\mathcal{O}(\sqrt{d}) \sum_i y_i 2^{-2i}} \qquad \text{(lemma 2.7.2)}$$

Next, we analyze X_{12}. We again use lemma 4.4.2(substituting $\mathbf{u}_0 = \mathbf{u}_i, \mathbf{v} = \mathbf{v}_j, \mathbf{u}_1 = \mathbf{u}_2 = \cdots = \emptyset$) for bounding $|\mathbf{u}_i^T A_s \mathbf{v}_j|$.

$$\begin{aligned}
X_{12} &\leq \sum_{(i,j)\in C_{12}} 2^{-i-j} |\mathbf{u}_i^T A_s \mathbf{v}_j| \\
&\leq \mathcal{O}(1) \sum_i \sum_{j=i+\frac{1}{2}\log(n/y_i)}^{\infty} 2^{-i-j} \sqrt{d y_i z_j} \sqrt{\frac{z_j}{n} \log\left(\frac{2n}{z_j}\right)} \\
&\leq \mathcal{O}(\sqrt{d}) \sum_i y_i 2^{-2i} \sum_{j=i+\frac{1}{2}\log(n/y_i)}^{\infty} 2^{-j+i} \sqrt{\frac{n}{y_i}} && \left(\frac{z_j}{n} \log\left(\frac{2n}{z_j}\right) \leq 1, z_j \leq n\right) \\
&\leq \boldsymbol{\mathcal{O}(\sqrt{d}) \sum_i y_i 2^{-2i}}
\end{aligned}$$

Combining X_{11} and X_{12} we get,

$$\boldsymbol{X_{10} \leq \mathcal{O}(\sqrt{d}) \sum_i y_i 2^{-2i}} \tag{4.14}$$

Putting it all together

Next we put together the multiple calculations in Equations 4.9,4.10,4.11,4.12,4.13,4.14

$$\begin{aligned}
X_6 \leq &\quad X_9 + X_{10} &&\leq \mathcal{O}(\sqrt{d}) \sum_i y_i 2^{-2i} + \mathcal{O}(\sqrt{d}) \sum_j z_j 2^{-2j} \\
X_5 \leq &\quad X_7 + X_8 &&\leq (\frac{\lambda}{5} + \mathcal{O}(\sqrt{d})) \sum_j z_j 2^{-2j} + \mathcal{O}(\sqrt{d}) \sum_i y_i 2^{-2i} \\
X_4 \leq &\quad X_5 + X_6 &&\leq (\frac{\lambda}{5} + \mathcal{O}(\sqrt{d})) \sum_j z_j 2^{-2j} + \mathcal{O}(\sqrt{d}) \sum_i y_i 2^{-2i} \\
X_1 \leq &\quad X_3 + X_4 &&\leq \mathcal{O}(\max(\sqrt{\lambda \log(d)}, \sqrt{d})) \sum_i y_i 2^{-2i} + (\lambda/5 + \mathcal{O}(\sqrt{d})) \sum_j z_j 2^{-2j} \\
X \leq &\quad X_1 + X_2 &&\leq \mathcal{O}(\max(\sqrt{\lambda \log(d)}, \sqrt{d})) \sum_i y_i 2^{-2i} + (\lambda/5 + \mathcal{O}(\sqrt{d})) \sum_j z_j 2^{-2j}
\end{aligned}$$

which proves the first of Lemma 4.2.2. But, if we know that for all $\boldsymbol{i, j, y_i \geq z_j}$, it is easy to see from the analysis that the following inequalities hold which prove the second part of the lemma 4.2.2

$$
\begin{aligned}
X_4 &= 0 \\
X_1 &\leq \mathcal{O}(\max(\sqrt{\lambda \log(d)}, \sqrt{d})) \sum_i y_i 2^{-2i} \\
X &\leq \mathcal{O}(\max(\sqrt{\lambda \log(d)}, \sqrt{d}))(\sum_i y_i 2^{-2i} + \sum_j z_j 2^{-2j})
\end{aligned}
$$

□

4.4.2 Proof of Lemma 4.4.1

For the sake of presentation we make a slight change of notation here. Let $exp(x)$ represent e^x

Proof: Without loss of generality we can assume that $S(\mathbf{v}) \subseteq N_G(S(\mathbf{u}))$(the neighbour set of the support of $\mathbf{u}$). If not we can simply look at the restriction of $\mathbf{v}$ on the set i.e $S(\mathbf{v}) \cap N_G(S(\mathbf{u}))$.

Let $Bad(\mathbf{u}, \mathbf{v})$ be the event which represents the event that

$$
|\mathbf{u}^T A_s \mathbf{v}| > 8\sqrt{\lambda \sqrt{|S(\mathbf{u})||S(\mathbf{v})|}|S(\mathbf{v})| \log\left(\frac{2d|S(\mathbf{u})|}{|S(\mathbf{v})|}\right)}
$$

Lemma 4.4.1 requires us to bound the following probability

$$
P(\cup_{\mathbf{u},\mathbf{v}} Bad(\mathbf{u}, \mathbf{v}))
$$

Note that the sum $\mathbf{u}^T A_s \mathbf{v}$ can be written as

$$
\mathbf{u}^T A_s \mathbf{v} = 2\sum_{i<j} [\mathbf{u}]_i [A_s]_{ij} [\mathbf{v}]_j
$$

Over a random signing A_s the RHS above is a sum of independent variables with maximum value ± 2 or ± 1 and mean 0. The maximum number of non-zero entries in this sum could be $E(S(u), S(v))$, i.e. the number of edges which go from $S(u), S(v)$ when they are seen as subsets of vertices of the original graph.

Therefore for a fixed u, v by applying Chernoff bounds we get that

$$
\begin{aligned}
P(Bad(\mathbf{u}, \mathbf{v})) = Pr\left(|\mathbf{u}^T A_s \mathbf{v}| > 8\sqrt{\lambda \sqrt{|S(\mathbf{u})||S(\mathbf{v})|}|S(\mathbf{v})| \log\left(\frac{2d|S(\mathbf{u})|}{|S(\mathbf{v})|}\right)}\right) \\
\leq 2 * exp\left(-2 * \frac{64\lambda\sqrt{|S(\mathbf{u})||S(\mathbf{v})|}|S(\mathbf{v})| \log\left(\frac{2d|S(\mathbf{u})|}{|S(\mathbf{v})|}\right)}{4E(S,T)}\right) \quad (4.15)
\end{aligned}
$$

Now given the condition of the lemma and the expander mixing lemma we have that

$$\begin{aligned} E(S(\mathbf{u}), S(\mathbf{v})) &= \frac{d|S(\mathbf{u})||S(\mathbf{v})|}{n} + \lambda\sqrt{|S(\mathbf{u})||S(\mathbf{v})|} \\ &\leq 2\lambda\sqrt{|S(\mathbf{u})||S(\mathbf{v})|} \end{aligned}$$

Putting this in the previous expression we get that the probability is bounded by

$$2exp\left(-16|S(\mathbf{v})|\log\left(\frac{2d|S(\mathbf{u})|}{|S(\mathbf{v})|}\right)\right)$$

Note that we want to put an upper bound on $P(\cup Bad(\mathbf{u}, \mathbf{v}))$ for all choices of $\mathbf{u}, \mathbf{v}$. For this purpose we would first fix the size of support of $\mathbf{u}, \mathbf{v}$ and union bound over all possible choices of $\mathbf{u}, \mathbf{v}$ of that fixed support and then union bound over all choices of the support. For fixed support sizes $|S(\mathbf{u})|, |S(\mathbf{v})|$, note that the total number of choices for the support sets for $\mathbf{u}$ are $\binom{n}{|S(\mathbf{u})|}$. Now since $S(\mathbf{v})$ is a subset of $N_G(S(\mathbf{u}))$ number of choices of $S(\mathbf{v})$ for a fixed $S(\mathbf{u})$ are bounded by $\binom{d|S(\mathbf{u})|}{|S(\mathbf{v})|}$. Also since each entry in $\mathbf{u}, \mathbf{v}$ is 0 or ± 1 the total number of choices for $\mathbf{u}$ and $\mathbf{v}$ are bounded by

$$\begin{aligned} \binom{n}{|S(\mathbf{u})|} * 2^{|S(\mathbf{u})|} * \binom{d|S(\mathbf{u})|}{|S(\mathbf{v})|} * 2^{|S(\mathbf{v})|} \leq exp\left(|S(\mathbf{u})|\log\left(\frac{n}{|S(\mathbf{u})|}\right) + (ln(2)+1)|S(\mathbf{u})|\right) \\ * exp\left(|S(\mathbf{v})|\log\left(\frac{d|S(\mathbf{u})|}{|S(\mathbf{v})|}\right) + (ln(2)+1)|S(\mathbf{v})|\right) \quad (4.16) \end{aligned}$$

We will first show upper bounds on each of these terms. Note that since $|S(\mathbf{v})| \geq \frac{n}{d^2}$ and hence $|S(\mathbf{u})| \geq \frac{n}{d^3}$ (since $|S(\mathbf{v})| \leq d|S(\mathbf{v})|$) we get that (assuming $d \geq 2$)

$$\begin{aligned} exp\left(|S(\mathbf{u})|\log\left(\frac{n}{|S(\mathbf{u})|}\right) + (ln(2)+1)|S(\mathbf{u})|\right) &\leq exp\left(3|S(\mathbf{u})|\log(d)\right) \\ &= exp\left(3\frac{\frac{|S(\mathbf{u})|}{|S(\mathbf{v})|}\log(d)}{\log\left(2d\frac{|S(\mathbf{u})|}{|S(\mathbf{v})|}\right)} * |(S(\mathbf{v})|\log\left(2d\frac{|S(\mathbf{u})|}{|S(\mathbf{v})|)}\right)\right) \\ &\leq exp\left(3 * |S(\mathbf{v})|\log\left(2d\frac{|S(\mathbf{u})|}{|S(\mathbf{v})|}\right)\right) \end{aligned}$$

The last line follows by noting that $x\log(d)/\log(2dx)$ is bounded by 1 for $x \in [1/d, 1]$ and that $\frac{|S(\mathbf{u})|}{|S(\mathbf{v})|} \in [1/d, 1]$ Also note that

$$exp\left(|S(\mathbf{v})|\log(\frac{2d|S(\mathbf{u})|}{|S(\mathbf{v})|}) + (ln(2)+1)|S(\mathbf{v})|\right) \leq exp\left(3 * |S(\mathbf{v})|\log(\frac{2d|S(\mathbf{u})|}{|S(\mathbf{v})|}\right)$$

Therefore by union bound we get that the probability of a bad event for fixed support sizes $|S(\mathbf{u})|, |S(\mathbf{v})|$ is bounded by

$$2 * exp\left(-6|S(\mathbf{v})|\log(\frac{2d|S(\mathbf{u})|}{|S(\mathbf{v})|})\right) \leq exp\left(-6\frac{n}{d^2}\log(2)\right)$$

Now the number of choices of the supports are n^2 at best and we get that

$$\begin{aligned} Pr\left(|\mathbf{u}^T A_s \mathbf{v}| \leq 8\sqrt{\lambda\sqrt{|S(\mathbf{u})||S(\mathbf{v})|}|S(\mathbf{v})|\log(\frac{2d|S(\mathbf{u})|}{|S(\mathbf{v})|})}\right) &\geq 1 - 2n^2 exp\left(-6\frac{n}{d^2}\log(2)\right) \\ &\geq 1 - exp\left(-\Omega(\frac{n}{d^2})\right) \end{aligned}$$

Hence proved. □

4.4.3 Proof of Lemma 4.4.2

Proof: As in the proof of the previous lemma. We will once again use the chernoff bound to bound the probability of bad events. We fill fix the size of the supports of $\mathbf{v}, \mathbf{u}_1, \mathbf{u}_2 \ldots$ and prove that the probability is small and then union bound over the choices of the support.

Lets first fix $\mathbf{v}, \mathbf{u}_1, \mathbf{u}_2, \ldots$. The sum $|\mathbf{v}^T A_s \mathbf{u}|$ is once again a sum of independent random variables with mean 0. This is so because note that the intersection between any two sets in $\{u_i\}$ is ϕ. It is easy to see that the sum of squares of the maximum values of these variables is

$$\leq \sum_i 4 * E(S(\mathbf{u}_i), S(\mathbf{v})) * 2^{2i}$$

Now we know that given the conditions of the lemma and the Expander Mixing Lemma

$$E(S(\mathbf{u}_i), S(\mathbf{v})) \leq 2\frac{d|S(\mathbf{u}_i)||S(\mathbf{v})|}{n}$$

Therefore using the above we get (via Chernoff Bound) that for a fixed $\mathbf{v}, \mathbf{u}_1, \mathbf{u}_2, \ldots$

$$Pr\left(|\mathbf{v}^T A_s \mathbf{u}| > 8\sqrt{\frac{d}{n}S(\mathbf{v})^2(\sum_i |S(\mathbf{u}_i)|2^{2i})\log(\frac{2n}{|S(\mathbf{v})|})}\right) \leq 2exp\left(-16S(\mathbf{v})\log(\frac{2n}{|S(\mathbf{v})|})\right)$$

Now fixing the values of the support sizes $|S(\mathbf{v})|, |S(\mathbf{u}_1)|, |S(\mathbf{u}_2)|, \ldots$ the number of possible choices for $\mathbf{v}$ are

$$\binom{n}{|S(\mathbf{v})|} * 2^{|S(\mathbf{v})|} \leq exp\left(3|S(\mathbf{v})|\log(\frac{2n}{|S(\mathbf{v})|})\right)$$

Similarly the number of possible choices for each $\mathbf{u}_i$ are

$$\leq \exp\left(3|S(\mathbf{u}_i)|\log(\frac{2n}{|S(\mathbf{u}_i)|})\right)$$

Therefore the total number of choices for all u_i are

$$exp\left(\sum_i 3|S(\mathbf{u}_i)|\log(\frac{2n}{|S(\mathbf{u}_i)|})\right)$$

Note that since each $|S(\mathbf{u}_i)|, |S(\mathbf{v})| \leq n$ we can replace each $|S(\mathbf{u}_i)|$ by its upper bound $\frac{|S(\mathbf{v})|}{2^{2i}}$ Therefore

$$\begin{aligned}\exp\left(\sum 3|S(\mathbf{u}_i)|\log\left(\frac{2n}{|S(\mathbf{u}_i)|}\right)\right) &\leq exp\left(3\sum\frac{|S(\mathbf{v})|}{2^{2i}}\log(\frac{2n}{\frac{|S(\mathbf{v})|}{2^{2i}}})\right)\\ &\leq exp\left(10|S(\mathbf{v})|\log\left(\frac{2n}{|S(\mathbf{v})|}\right)\right)\end{aligned}$$

The last inequality follows by applying Lemma 2.7.2 Therefore the total number of choices of $\mathbf{v}, \mathbf{u}_1 \ldots$ fixing $|S(\mathbf{v})|, |S(\mathbf{u}_1)| \ldots$ are bounded by

$$\exp\left(13|S(\mathbf{v})|\log\left(\frac{2n}{|S(\mathbf{v})|}\right)\right)$$

Therefore by union bound fixing the support sizes the probability of the bad event is bounded by

$$exp\left(-3|S(\mathbf{v})|\log\left(\frac{2n}{|S(\mathbf{v})|}\right)\right) \leq exp\left(-3\frac{n}{d^2}\log(d)\right)$$

Now the number of choices for sizes of these supports are at best $n * n^{\log(n)}$. To see this since the size of each $|S(\mathbf{u}_i)|$ decreases exponentially there can be at best $\log(n)$ such sets. Therefore putting together the union bound we get that the total probability of the bad event is bounded by

$$exp\left((\log(n))(\log(n)+1)\right) * exp\left(-3\frac{n}{d^2}\log(d)\right) \leq exp\left(-\Omega(\frac{n}{d^2})\right)$$

This proves the lemma.

□

4.5 Conclusion and Open Problems

In this section we showed that the non-trivial part of the spectrum of a random 2-lift with high probability is bounded by the spectrum of the base graph upto constant factors. This naturally suggests that in order to to create bigger

Ramanujan Graphs one can expect to take small Ramanujan graphs and repeatedly take a random 2-lift. This strategy however fails because of the presence of a constant factor in our results. In particular after $\log(d)$ such lifts our upper bounds essentially become redundant. Although the existence of a Ramanujan lift is known (for bipartite graphs), the precise behaviour with random lifts seems hard to pin down.

Another important open question in this regime is the existence of efficient algorithms that given a Ramanujan graph G in polynomial time can produce a Ramanujan lift of the graph. It is conceivable that a derandomization of our result leads to Ramanujan graphs (upto constant loss in the spectrum), however the real challenge is to be able to preserve the Ramanujan bound exactly by taking lifts. Efficient constructions of such families remains an important open question in Spectral Graph Theory.

References

[ABG10] L. Addario-Berry and S. Griffiths. The spectrum of random lifts. Arxiv preprint arXiv:1012.4097, 2010.

[ABS10] S. Arora, B. Barak, and D. Steurer. Subexponential algorithms for unique games and related problems. In *Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, FOCS '10, pages 563–572, 2010.

[AIMS10] Sanjeev Arora, Russell Impagliazzo, William Matthews, and David Steurer. Improved algorithms for unique games via divide and conquer. In *In Electronic Colloquium on Computational Complexity*, pages TR10–041, 2010.

[AKK$^+$08] Sanjeev Arora, Subhash Khot, Alexandra Kolla, David Steurer, Madhur Tulsiani, and Nisheeth Vishnoi. Unique Games on expanding constraint graphs are easy. In *Proceedings of the 40th ACM Symposium On The Theory Of Computing*, volume 40, pages 21–28, 2008.

[ALM$^+$98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, May 1998.

[AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of np. *J. ACM*, 45(1):70–122, January 1998.

[BBH$^+$12] Boaz Barak, Fernando G. S. L. Brandão, Aram Wettroth Harrow, Jonathan A. Kelner, David Steurer, and Yuan Zhou. Hypercontractivity, sum-of-squares proofs, and their applications. In *Proceedings of the 44th Symposium on Theory of Computing Conference*, pages 307–326, 2012.

[BGH$^+$11] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer. Making the long code shorter, with applications to the unique games conjecture. *Electronic Colloquium on Computational Complexity (ECCC)*, 18:142, 2011.

[BJK05] Andrei Bulatov, Peter Jeavons, and Andrei Krokhin. Classifying the complexity of constraints using finite algebras. *SIAM J. Comput.*, 34(3):720–742, March 2005.

[BL06] Y. Bilu and N. Linial. Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica*, 26(5):495–519, 2006.

[CKK$^+$06] Shuchi Chawla, Robert Krauthgamer, Ravi Kumar, Yuval Rabani, and D. Sivakumar. On the hardness of approximating multicut and sparsest-cut. In *Computational Complexity*, volume 15(2), pages 94–114, 2006.

[CMM06a] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for unique games. In *Proceedings of the ACM Symposium on the Theory of Computing*, volume 38, pages 205–214, 2006.

[CMM06b] Eden Chlamtac, Konstantin Makarychev, and Yury Makarychev. How to play unique games using embeddings. In *Proceedings of the IEEE Symposium on Foundations of Computer Science*, volume 48, pages 687–696, 2006.

[Fri03] J. Friedman. Relative expanders or weakly relatively ramanujan graphs. *Duke Math. J*, 118:2003, 2003.

[Fri08] J. Friedman. A proof of alon's second eiganvalue conjecture and related problems. *Mem. Amer. Math,Soc*, 195(910), 2008.

[FS02] U. Feige and G. Schechtman. On the optimality of the random hyperplane rounding technique for max-cut. *Random Structures and Algorithms*, 20:403–440, 2002.

[Gre95] Y. Greenberg. On the spectrum of graphs and their universal coverings. Ph.D Thesis, 1995.

[GT06] Anupam Gupta and Kunal Talwar. Approximating unique games. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, volume 24, pages 99–106, 2006.

[GW95] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6), 1995.

[Hås01] Johan Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.

[HLW06] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc*, 43(4):439–561, 2006.

[Hoe62] Wassily Hoeffding. Probability inequalities for sums of bounded random variables, 1962.

[IPZ01] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *J. Comput. Syst. Sci.*, 63(4):512–530, 2001.

[Kar72] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103, 1972.

[Kar99] H. Karloff. How good is the goemans-williamson max-cut algorithm? *SIAM Journal on Computing*, 29(1):336–350, 1999.

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the ACM Symposium on the Theory of Computing*, volume 34, pages 767–775, 2002.

[Kho10] Subhash Khot. On the unique games conjecture (invited survey). In *IEEE Conference on Computational Complexity*, pages 99–121, 2010.

[KKMO04] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for max-cut and other 2-variable CSPs? In *Proceedings of the IEEE Symposium on Foundations of Computer Science*, volume 45, pages 146–154, 2004.

[KMM11] Alexandra Kolla, Konstantin Makarychev, and Yury Makarychev. How to play Unique Games against a semi-random adversary. In *Proceedings of the 43nd ACM Symposium on Foundations of Computer Science*, 2011.

[Kol10] Alexandra Kolla. Spectral algorithms for Unique Games. In *Proceedings of the 25th ACM Conference on Computational Complexity*, volume 25, pages 122–130, 2010.

[KPS10] Subhash Khot, Preyas Popat, and Rishi Saket. Approximate lasserre integrality gap for unique games. In *APPROX-RANDOM*, Lecture Notes in Computer Science, pages 298–311. Springer, 2010.

[KR03] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within $2-\varepsilon$. In *Proceedings of the IEEE Conference on Computational Complexity*, volume 18, pages 379–386, 2003.

[KV05] Subhash Khot and Nisheeth K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1. In *Proceedings of the IEEE Symposium on Foundations of Computer Science*, volume 46, pages 53–62, 2005.

[LP10] N. Linial and D. Puder. Word maps and spectra of random graph lifts. *Random Struct. Algorithms*, 37(1)):100–135, 2010.

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[LSV11] E. Lubetzky, B. Sudakov, and V. Vu. Spectra of lifted ramanujan graphs. *Advances in Mathematics*, 227:16121645, 2011.

[Mar88] G.A Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Probl. Inf. Transm*, 24(1):39–46, 1988.

[MM10] Konstantin Makarychev and Yury Makarycev. How to play unique games on expanders. In *In Proceedings of the eighth Workshop on Approximation and Online Algorithms*, 2010.

[MR08] Dana Moshkovitz and Ran Raz. Two-query pcp with subconstant error. *J. ACM*, 57(5):29:1–29:29, June 2008.

[MSS13] A. Marcus, D. Spielman, and N. Srivastava. Interlacing families i: Ramanujan graphs of all degrees. In Proceedings, FOCS 2013, 2013.

[Nil91] A. Nilli. On the second eigenvalue of a graph. *Discrete Math*, 91(2):207–210, 1991.

[O'D14] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.

[Pin73] M. Pinsker. On the complexity of a concentrator. *7th International Teletraffic Conference*, pages 318/1–318/4, 1973.

[Pud13] D. Puder. Expansion of random graphs: New proofs, new results. Arxiv Preprint:1212.5216, 2013.

[Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every csp? In *Proceedings of the 40th ACM Symposium On The Theory Of Computing*, volume 40, pages 245–254, 2008.

[RS09] Prasad Raghavendra and David Steurer. Integrality gaps for strong sdp relaxations of unique games. In *FOCS*, pages 575–585. IEEE Computer Society, 2009.

[RS10] Prasad Raghavendra and David Steurer. Graph expansion and the Unique Games conjecture. In *Proceedings of the 42nd ACM Symposium On The Theory Of Computing*, volume 42, pages 755–764, 2010.

[Sar06] P. Sarnak. What is an expander? *Notices Amer. Math. Soc*, 51(7):762–763, 2006.

[Sch03] A. Schrijver. *Combinatorial Optimization: Polyhedra and Efficiency. Disjoint paths, hypergraphs*. Number v. 3 in Algorithms and Combinatorics. Springer, 2003.

[Shm97] David Shmoys. Cut problems and their application to divide-and-conquer. In Dorit S. Hochbaum, editor, *Approximation Algorithms for NP-hard Problems*. PWS Publishing Co., Boston, MA, USA, 1997.

[Tre05] Luca Trevisan. Approximation algorithms for unique games. In *Proceedings of the IEEE Symposium on Foundations of Computer Science*, volume 46, pages 197–205, 2005.