Algorithms and Explicit Constructions via Spectral Techniques

Jun-Ting Hsieh

CMU-CS-25-131

August 2025

Computer Science Department
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:
Pravesh K. Kothari, Chair (CMU/Princeton University)
Ryan O'Donnell
Jason Li
Venkatesan Guruswami (UC Berkeley)
David Steurer (ETH Zurich)

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

This research was sponsored by the National Science Foundation under award numbers CCF2047933 and CCF2211971. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of any sponsoring institution, the U.S. government or any other entity.

Keywords: Spectral Methods, Spectral Expanders, Vertex Expanders, Constraint Satisfaction Problems, Independent Set

To my parents Tsang-Jen and Chui-Hui, my sister Jun-Shiuan, and my wife Bingbin

Abstract

Spectral methods have become ubiquitous in computer science. By analyzing the eigenvalues and eigenvectors of matrices naturally associated with a graph, such as its adjacency matrix, one can extract useful information about the graph's structure. Such methods have yielded the best-known results for a wide range of foundational problems.

In this thesis, we apply this "spectral lens" to prove new results in graph theory, design algorithms, and construct explicit vertex expanders. The results are divides into three parts:

- **Part I**. We develop spectral techniques to obtain new results in graph theory. These results are not only of independent interest but are also key ingredients in later parts of the thesis.
- **Part II**. We present algorithms for both refuting semirandom constraint satisfaction problems and recovering solutions in planted ones, both utilizing spectral information of the underlying hypergraph. Moreover, we give algorithms to find large independent sets in spectral expanders.
- **Part III**. We construct explicit constant-degree vertex expanders using the *tripartite line product*. First, we obtain explicit unique-neighbor expanders by instantiating the product using Ramanujan graphs — the optimal spectral expanders. Then, by replacing Ramanujan graphs with the incidence graphs of Ramanujan cubical complexes, we obtain the first explicit *lossless* vertex expanders.

Acknowledgments

First and foremost, I would like to thank my advisor Pravesh K. Kothari. I first met him when he was still a postdoc at Princeton, just about to move to CMU. At the time, I knew nothing about theory, but he still spoke to me with great enthusiasm about Sum-of-Squares, average-case complexity, robust statistics, and more. To this day, I am grateful that he took me as his student despite my background and lack of prior experience. As an advisor, he has taught me not only math/theory but also how to do research, while giving me the constant support and encouragement I needed. As a researcher, he is always energetic and seems to have endless ideas. Moreover, as a friend, we had a lot of fun together, for example the countless boba/food trips where we chatted about all sorts of random things.

I must also thank my other committee members, Ryan O'Donnell, Jason Li, Venkat Guruswami, and David Steurer, for being on my committee and attending my thesis proposal and defense. They provided valuable feedback and suggestions that helped shape this thesis.

I had the privilege of visiting several professors over the course of my PhD. In the summer of 2022, I visited Luca Trevisan at Bocconi University. Even before the visit, I had read several of Luca's blog posts, which had a profound impact on my research career. In one post, he wrote, "*I consider a mathematical proof to be understood if one can see it as a series of inevitable steps*", a piece of advice I continue to keep in mind. During my visit, I got to know Luca personally. He was a very inspiring person to talk to, and it's clear to me that he knew almost every part of TCS. He was also quite funny at times. Sadly, he passed away far too soon, but I will always be grateful for his mentorship and lasting impact.

I also visited Venkat Guruswami and Prasad Raghavendra at Berkeley, Sam Hopkins at MIT, Raghu Meka at UCLA, and recently David Steurer at ETH Zurich. I am deeply grateful for the opportunities I had to visit different places during my PhD. These visits allowed me to explore new areas of TCS and to make new friends. My research career would have been very different without them.

Among professors at CMU, I am especially grateful to Ryan O'Donnell for his "Theorist's Toolkit" course, which has benefited not only me but many others as well. The course was instrumental for my early career in theory, and I would not be where I am today without it. I would also like to thank Aayush Jain, Anupam Gupta, Jason Li, and David Woodruff for valuable research discussions, helpful suggestions about giving talks, and also many fun conversations.

I must also thank Li-Yang Tan at Stanford. At the time, I was doing research in computer vision, but I took his Computational Complexity course as a breadth requirement, and I absolutely loved it. That course sparked my interest in TCS. Perhaps more importantly, it was Li-Yang who recommended that I talk to Pravesh!

I would like to thank the entire CSD staff for their continuous support. In particular, I am grateful to Patricia Loring for always handling the reimbursements efficiently, and to Matthew Stewart for helping with logistical matters and sending all the email announcements. I had countless email exchanges with them throughout my PhD.

Now, among my peers, let me start with my collaborators that contributed to the results in this thesis. I would like to thank Sidhanth Mohanty for "kikuchi-fun", Jeff Xu for constantly drawing shapes on the board, Mitali Bafna for tackling the 3-coloring problem together, Peter Manohar for our struggles with planted CSPs, Theo McKenzie and Pedro Paredes for unique-neighbor expanders, and Sidhanth Mohanty and Rachel Zhang for our journey towards lossless expanders (i.e., the adventure of the three Pokémons: Scizor, Vulpix, and Psyduck).

I am deeply grateful for my first two friends in TCS: Sidhanth Mohanty and Jeff Xu. During my first year in Pravesh's group, due to COVID, Pravesh hosted virtual group meetings every week. That was when I first met Jeff (another Pravesh's student) and Sidhanth (then at Berkeley but could join remotely). Being a student with little background in theory, I am extremely lucky to have them; they taught me so many things I didn't know. In fact, my first paper in TCS was with them, just the three of us. Almost immediately, we became very good friends. I really enjoyed all the fun activities we did (and still do) together — basketball, NBA 2K, and musicals with Jeff; Pokémon Showdown, Pokémon charades, and table tennis with Sidhanth; and with both of them, countless boba trips, movies and shows, cooking, adventures, etc.

I'm lucky to have so many amazing friends at CMU. First, my officemates: Kunming (Benny) Jiang, Lingjing Kong, Jeff Xu (and Sarah Wang), and Tesla Zhang. We had so much fun together, from watching/playing various sports to seeing the solar eclipse, and much more. I also want to thank my wonderful friends among the current and former members of the CMU theory group: Omar Alrabiah, Prashanti Anderson, Ainesh Bakshi, Tolson Bell, Isaac Grosof, Daniel Hathcock, William He, Praneeth Kacham, George Li, Hoai-An Nguyen, Pedro Paredes, Madhusudhan Pittu, Kevin Pratt, Sherry Sarkar, Noah Singer, Xinyu Wu, Mik Zlatin, and many others. Many thanks as well to my friends in Machine Learning: Arundhati Banerjee, Zhili Feng, Jerry Huang, Justin Khim, Oscar Li, Yuchen Li, Yusha Liu, Tanya Marwah, Ashwini Pokle, Che-Ping Tsai, Chih-Kuan Yeh. I also want to especially thank Yusha for Lord of the Rings, Dune, hiding Bingbin's "snacks", etc, and Ashwini for introducing BTS to Bingbin and me.

Outside of CMU, I got to know many amazing friends from research visits and conferences. I am especially grateful to people from "Dirham Store 1641 Walnut": Yeshwanth Cherapanamjeri, Bingbin Liu, Sidhanth Mohanty, Amit Rajaraman, and Nived Rajaraman; so many CLUELESSes and fun memories like removing tape on the cardboard. I would also like to thank the numerous friends in TCS: Arpon Basu, Rares-Darius Buhai, Antares Chen, Sitan Chen, Tommaso d'Orsi, Louis Golowich, Aparna Gupte, Yiding Hua, Brice Huang, Nathan Ju, Jane Lange, Daniel Lee, Jerry Li, Andrew

Lin, David Lin, Allen Liu, Chih-Hung Liu, Siqi Liu, Fermi Ma, Mahbod Majid, Shivam Nadimpalli (and Stitch), Ansh Nagda, Lucas Pesenti, Louie Putterman, Goutham Rajendran, Stefan Tiegel, Thuy-Duong (June) Vuong, Yimeng (Kobe) Wang, David Wu, Rachel Zhang, and many more.

Many thanks to my friends before graduate school — from high school: Chi-Fang (Anthony) Chen, Yi-Shiou Duh, Hung-Jui (Joe) Huang, Chris Lee, and Ching-Ting Tsai; the StarCraft crew from undergrad: Justin Doong, Steven Lin, Maurice Shih, and Timothy Wu; Vision Lab: Edward Chou, Michelle Guo, De-An Huang, and Zelun Luo.

Last but not least, I am truly grateful to my family, the most important people in my life. Thanks to my parents Tsang-Jen and Chui-Hui, and my sister Jun-Shiuan for their love and support throughout my entire life. Finally, I want to thank my wife Bingbin. We have gone through many stages of life together, including Masters, PhD, and now postdoc. Bingbin has always by my side, a constant source of love, encouragement, and joy. I am so lucky to have her in my life.

Contents

Chapter 1

Introduction

Spectral methods are now widespread across computer science. Many computational problems naturally involve matrices, either as explicit inputs or as representations of the underlying structures. Spectral methods refer to those that utilize the spectra of these matrices —– their eigenvalues, eigenvectors, or singular values and vectors —– in the analysis or in an algorithm's implementation. Graphs, for example, are natural applications for spectral methods, as they can be represented by their adjacency matrices. In particular, the second eigenvalue of the adjacency matrix is closely related to the *edge expansion* due to Cheeger's inequality [Alo86], and it also gives bounds on subgraph densities via the expander mixing lemma [AC88]. These connections have led to extensive study on *spectral expanders* — graphs with small second eigenvalues — for which we now have several explicit constructions and far-reaching applications, including error correcting codes, communication and computation networks, and derandomization.[1]

Spectral methods have enjoyed a lot of success in algorithm design as well. The aforementioned Cheeger's inequality is in fact a graph partitioning algorithm that finds a sparse cut in a graph using the second eigenvector of the adjacency matrix. Another notable example is the planted clique (also known as hidden clique) problem, where the goal is to find a clique planted in a random Erdős-Rényi graph $G(n, 1/2)$. The best known polynomial-time algorithm utilizes the eigenvalues and eigenvectors of the graph's adjacency matrix to detect and recover the planted clique [AKS98].

Across the various applications, the unifying theme is to extract useful information about the graph from the spectra of these associated matrices, translating combinatorial structure to linear-algebraic information. A classic example is the connection between counting walks in a graph and the top eigenvalue of its adjacency matrix. More precisely, given a symmetric matrix $A \in \mathbb{R}^{n \times n}$, the trace power $\mathrm{tr}(A^{2\ell})$ counts weighted closed walks of length 2ℓ on n vertices, while the spectral norm λ of A satisfies $\lambda \leqslant \mathrm{tr}(A^{2\ell})^{1/2\ell} \leqslant n^{1/2\ell}\lambda$. Thus, for $\ell \gg \log n$, the spectral norm serves as a tight

[1] We refer readers to the survey of Hoory, Linial, and Wigderson [HLW06] for a thorough exposition.

proxy for combinatorial closed walks in the matrix.

Another key component in spectral methods is identifying the "correct" matrix representation and its spectral properties. In many cases, the main innovation lies in selecting the right matrix that enables the desired analysis or proof technique. For example, in Chapter 5, we analyze subgraph densities in spectral expanders using the spectral radius of the non-backtracking matrix, rather than the adjacency matrix. In Chapters 4 and 7, we will use the *Kikuchi matrix* [WAM19] constructed from the given hypergraph to prove girth-density trade-offs in hypergraphs as well as give refutation algorithms for semirandom constraint satisfaction problems (CSPs).

In this thesis, we follow this theme of studying fundamental problems in theoretical computer science through the "spectral lens". We focus on three aspects: **graph theory**, **algorithm design**, and **explicit constructions**.

1.1 Part I: Spectral methods in graph theory

We begin the thesis with some new results in graph theory. These results will be used in later chapters, and we believe they are of independent interest.

Girth-density trade-off in graphs. Alon, Hoory, and Linial [AHL02] proved a celebrated result, known as the *Moore bound*, which states that that any n-vertex graph of average degree $d > 2$ must contain a cycle of length at most $2\log_{d-1} n + 2$. While this fact is straightforward for d-regular graphs, the remarkable aspect of their result is that it also applies to irregular graphs, where the average degree d need not be an integer.

In Chapter 4, we show an improved upper bound on girth in terms of the *spectral radius* of the graph's *non-backtracking matrix* (Definition 2.2.1). This result recovers the classical Moore bound as a special case, but in many settings — including those appearing in Chapter 12 — it provides substantially stronger guarantees.

The proof relies on the relationship between non-backtracking walks in a graph and the spectrum of its non-backtracking matrix, similar to the classical connection between walks and the spectrum of the adjacency matrix.

Hypergraph Moore bound. There is a natural generalization of cycles to hypergraphs, called *even covers*. An even cover is a collection of hyperedges such that every vertex participates in an even number of hyperedges. The girth of a hypergraph is defined as the smallest size of an even cover in it. Due to the equivalence between even covers and linear dependencies over $\mathbb{F}_2$, the girth-density trade-off in hypergraphs has been previously studied in coding theory. Moreover, it is closely related to constraint satisfaction problems, which led Feige [Fei08] to conjecture that every k-uniform hypergraph with n vertices and $m \gtrsim n(\frac{n}{r})^{\frac{k}{2}-1}$ hyperedges has an even cover of size $O(r \log n)$.

Feige's conjecture was resolved by Guruswami, Kothari, and Manohar [GKM22] up to polylogarithmic factors. Their proof goes via a spectral argument applied to the

Kikuchi graph [WAM19] — a graph derived from the given hypergraph (see Definition 4.3.2).

In Chapter 4, we present a simple and short proof of the hypergraph Moore bound that is *almost tight*, up to a single logarithmic factor. Our simplified analysis also leads to refutation algorithms for semirandom constraint satisfaction problems, which we will discuss in the next section.

Subgraph density in spectral expanders. We next study the relationship between spectral expansion and vertex expansion. In particular, we focus on subgraph density and vertex expansion of small subsets in graphs with bounded non-trivial eigenvalues. The well-known expander mixing lemma [AC88] (Fact 2.1.1) already provides non-trivial bounds, which were further improved by Kahale [Kah95], who proved tight bounds for d-regular expanders. See Section 3.2 for more discussions.

In Chapter 5, we give a sharp upper bound on the density of subgraphs in *bipartite* spectral expanders, generalizing the results of Kahale [Kah95] and Asherov and Dinur [AD24]. At a high level, our approach bounds the spectral radius of the non-backtracking matrix of subgraphs, which in turn yields upper bounds on the density. This improves over the expander mixing lemma, which bounds the density via the spectrum of the adjacency matrix.

1.2 Part II: Spectral methods in algorithm design

Boolean constraint satisfaction problems (CSPs). One focus of this thesis is on algorithms for average-case Boolean CSPs, such as random k-SAT or k-XOR. A CSP instance can be represented as a hypergraph along with literal negation patterns associated with each hyperedge. The general approach we use here is to represent the hypergraph using the *Kikuchi matrix* — first introduced by Wein, Alaoui, and Moore [WAM19] and further developed by Guruswami, Kothari, and Manohar [GKM22] — and analyzing its spectral properties to design algorithms.

For fully random k-CSPs — where the hypergraph and literal patterns are both chosen randomly — the instance is highly unsatisfiable with high probability, and the task is to output a *refutation*, i.e., a certificate of unsatisfiability. A closely related setting is the *search* problem for *planted* random CSPs, where the literal patterns are chosen from some distribution that aligns with an unknown planted solution. Here, the algorithmic task is to recover the planted solution.

Random CSPs have received a lot of attention over the past decade. For both the refutation and search problem, we now have algorithms that succeed at the optimal clause density threshold [AOW15, RRS17, FPV15] (see Chapter 6 for more background). However, many such algorithms break down under minor perturbations of the instance, such as the addition of a vanishingly small fraction of extra clauses. This motivates the

study of *semirandom* models, which are hybrid models between average-case and worst-case pioneered by Blum and Spencer [BS95]. In the case of semirandom CSPs, we allow the underlying hypergraph to be adversarial, while the literal patterns are still random.

In Chapters 7 and 8, we give algorithms for both refuting semirandom CSPs as well as solving semirandom planted CSPs. The key technique is analyzing the spectral properties of the *Kikuchi matrix* (see Definition 7.0.3). At a high level, a k-XOR instance can be represented as a signed Kikuchi matrix, and its spectral norm (or a reweighted version) is an upper bound on the value of the k-XOR instance. Furthermore, the connection between closed walks in the Kikuchi matrix and its spectral norm has led to the *hypergraph Moore bound*, a result in extremal combinatorics that characterizes the girth-density trade-offs in hypergraphs [Fei08, GKM22]. See Chapter 3 for more background.

Solving problems on spectral expanders. The study of (semi)random CSPs can be viewed as a way to circumvent worst-case hardness. A complementary approach is to study problems under certain structural assumptions on the input, such as expansion. A notable success story is the line of work on Unique Games (UG). Research started with algorithms for Unique Games assuming that the underlying graph is a spectral expander or has low threshold rank [Tre08, AKK+08, MM11]. This ultimately led to the groundbreaking subexponential-time algorithm for arbitrary UG instances [ABS15].

We follow this paradigm and study the problem of finding large independent sets in graph that are either 3-colorable or promised to contain large independent sets — problems that are UG-hard without further assumptions on the input graphs. In Chapter 9, we give polynomial-time algorithms that find linear-sized independent sets in *one-sided spectral expanders* that are almost 3-colorable or are promised to contain independent sets of size $(1/2 - \varepsilon)n$.

In sharp contrast to our algorithmic result, we show that the analogous task of finding a linear-sized independent set in an almost 4-colorable one-sided expander (even when the second eigenvalue is $o_n(1)$) is NP-hard, assuming the Unique Games Conjecture. This reveals a surprising difference between 3-colorable and 4-colorable expanders. See Chapter 6 for our results and more discussions.

1.3 Part III: Spectral methods in constructing vertex expanders

Spectral expansion — more specifically, bounded second eigenvalue — is one the most well-studied notion of expansion. Many applications require graph properties such as high conductance, low subgraph densities, and rapid mixing of random walks, all of which are closely connected to the second eigenvalue of a graph. The optimal spectral expanders are called Ramanujan graphs, where all non-trivial eigenvalues are bounded

by $2\sqrt{d-1}$ for d-regular graphs. Moreover, random d-regular graphs are known to be almost Ramanujan with high probability, with all non-trivial eigenvalues bounded by $(2+o(1))\sqrt{d-1}$ [Fri08, Bor20].[2]

However, many applications require *explicit constructions* of such expanders, where random graphs do not suffice. In the weaker notion of explicitness, one requires that an n-vertex graph be generated deterministically "from scratch" in time polynomial in n. The stronger notion, called *strongly explicit*, requires that the neighborhood of any given vertex be computable in time $\mathrm{polylog}(n)$. There are now several strongly explicit constructions of Ramanujan graphs [Mar88, LPS88, Mor94]. While the analyses of these constructions are technically deep, one could attribute their success to the "analytic nature" of the second eigenvalue.

Vertex and unique-neighbor expanders. There is a wide range of problems that require different notions of expansion — *unique-neighbor expansion* or *vertex expansion*. The definitions are very intuitive: in a d-regular graph, we require every "small" set S of vertices to have $\gamma d|S|$ (unique-)neighbors, for constant $\gamma > 0$. Here, a unique-neighbor of S is a vertex v with exactly one edge to S. When γ can be made arbitrarily close to 1, i.e., $\gamma = 1-\varepsilon$, then we call such graphs *lossless expanders*. We refer readers to Chapter 10 and the survey of [HLW06] for more background and various applications of explicit vertex expanders.

Significant research has gone into explicit constructions of *bipartite* vertex expanders. In particular, in a breakthrough work, Capalbo, Reingold, Vadhan, and Wigderson [CRVW02] gave explicit constructions for *one-sided* lossless expanders, in which lossless expansion holds only for subsets on one side. Such expanders already enable numerous applications in error-correcting codes, distributed routing networks, and more (see [CRVW02] and references therein). On the other hand, constructing explicit *two-sided* lossless expanders has been a long-standing open problem.

Given the abundance of explicit Ramanujan graph constructions, a natural question is whether spectral expansion implies vertex expansion or unique-neighbor expansion. Kahale [Kah95] showed that Ramanujan graphs exhibit vertex expansion with $\gamma = 1/2$, but unfortunately it can have subsets with *zero* unique-neighbors. See Chapter 5 for our results that extend Kahale's to bipartite near-Ramanujan graphs.

Explicit lossless expanders via the tripartite line product. In a sequence of papers [HMMP24, HLMOZ25, HLMRZ25], we achieve the goal of constructing strongly explicit two-sided lossless expanders [HLMRZ25]. In all three papers, the key object is the *tripartite line product*, introduced in [HMMP24].

This product has two ingredients: a large (infinite family of) tripartite base graph and a constant-size gadget graph. The base graph G has vertex set $L \cup M \cup R$, where

[2]Recently, a breakthrough result by Huang, McKenzie, and Yau [HMY24] showed that a random d-regular graph is Ramanujan with probability approximately 0.69.

we place a (k, D_L)-biregular graph between L and M, and a (D_R, k)-biregular graph between M and R. The gadget graph H is a (d_L, d_R)-biregular graph on vertex set $[D_L] \cup [D_R]$. The tripartite line product between G and H is the (kd_L, kd_R)-biregular graph on L and R obtained as follows: for each vertex $v \in M$, place a copy of H between the D_L left neighbors of v and the D_R right neighbors of v (see Chapter 11 for more details and Figure 11.1 for an illustration).

Since the gadget graph is of constant size, we think of it as a random graph that satisfies strong expansion properties. Thus, the innovation lies in the choice of the base graph. In [HMMP24], we instantiate the base graph (i.e., the two bipartite graphs) with bipartite Ramanujan graphs and obtain two-sided γ-unique-neighbor expanders, with γ being a small universal constant. In [HLMOZ25], we instantiate it using the face-vertex incidence graphs of Ramanujan simplicial complexes and obtain two-sided 3/5-vertex expanders, beating the spectral barrier of 1/2 by Kahale. Finally, in [HLMRZ25], we instantiate it using the face-vertex incidence graphs of Ramanujan cubical complexes (see Section 13.3) and obtain two-sided lossless expanders.

1.4 Organization of the thesis

This thesis is divided into three parts. Each part begins with an introductory chapter that states the main results and provides additional background.

Part I: Graph Theory.

- Chapter 3: Introduction.
- Chapter 4: Girth-density trade-off in hypergraphs. This chapter is based on [HKM23, Sections 2, 3, and A].
- Chapter 5: Subgraph density in spectral expanders. This chapter is based on [HMMP24, Sections 5-6].

Part II: Algorithms.

- Chapter 6: Introduction.
- Chapter 7: Algorithms for strongly refuting semirandom CSPs. This chapter is based on [HKM23, Section 4].
- Chapter 8: Efficient algorithms for semirandom planted CSPs. This chapter is based on [GHKM23].
- Chapter 9: Rounding large independent sets on expanders. This chapter is based on [BHK25].

Part III: Explicit constructions.

- Chapter 10: Introduction.

- Chapter 11: Tripartite line product. This chapter is based on [HMMP24, HLMOZ25, HLMRZ25].
- Chapter 12: Unique-neighbor expanders with lossless small-set expansion. This chapter is based on [HMMP24].
- Chapter 13: Explicit lossless vertex expanders. This chapter is based on [HLMRZ25].

Chapter 2

Background and Preliminaries

2.1 Elementary graph theory

In this thesis, we will restrict to undirected graphs. Given an undirected graph $G = (V, E)$ with $|V| = n$ vertices and $|E| = m$ edges, we denote $A_G \in \{0,1\}^{n \times n}$ to be its adjacency matrix, and $D_G \in \mathbb{N}^{n \times n}$ to be its diagonal degree matrix. We also denote $\widetilde{A}_G = D_G^{-1/2} A_G D_G^{-1/2}$ to be the *normalized* adjacency matrix. We will drop the dependence on G when the graph is clear from context.

For bipartite graphs, we often write it as $G = (L \cup R, E)$, where L and R are the left and right vertex sets respectively. If G has left degree c and right degree d, then we say that G is (c, d)-biregular.

Spectrum of the graph. Unless stated otherwise, we use $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ to denote the eigenvalues of A_G.

Fact 2.1.1 (Expander mixing lemma [AC88])**.** *Let $G = (V, E)$ be a d-regular graph on n vertices, and let λ_2 be the second eigenvalue of its adjacency matrix. For any subsets $S, T \subseteq V$,*

$$\left| e(S,T) - \frac{d}{n}|S||T| \right| \leqslant \lambda_2 \sqrt{|S||T|}\,.$$

Ramanujan graphs. For any d-regular graph G, the top eigenvalue $\lambda_1 = d$. We say that G is a *Ramanujan* graph if $\max\{\lambda_2, |\lambda_n|\} \leqslant 2\sqrt{d-1}$. The quantity $2\sqrt{d-1}$ is precisely the spectral radius of the adjacency operator of the infinite d-regular tree. Therefore, d-regular Ramanujan graphs are graphs whose non-trivial eigenvalues lie in the spectrum of the infinite d-regular tree.

Ramanujan graphs are the *optimal* spectral expanders — Alon and Boppana [Nil91] showed that a d-regular graph must have $\lambda_2 \geqslant 2\sqrt{d-1} - o_n(1)$ for any fixed d.

Ramanujan bipartite graphs. The definition of Ramanujan graphs can be generalized to biregular graphs. Let $c < d$ be integers. For any (c, d)-biregular graph $G = (L \cup R, E)$,

the top eigenvalue $\lambda_1 = \sqrt{cd}$. This is witnessed by the eigenvector x with entries $x_u = \sqrt{c}$ for $u \in L$ and $x_v = \sqrt{d}$ for $v \in R$. Note that for any bipartite graph, λ is an eigenvalue if and only if $-\lambda$ is an eigenvalue, so $\lambda_n = -\sqrt{cd}$.

We say that G is a (c,d)-biregular *Ramanujan* graph if

$$|\lambda_i| \in \left[\sqrt{d-1} - \sqrt{c-1},\ \sqrt{d-1} + \sqrt{c-1}\right] \cup \{0\}, \qquad \forall i \neq 1, n.$$

The above is precisely the spectrum of the infinite (c,d)-biregular tree [GM88, LS96]. Therefore, (c,d)-biregular Ramanujan graphs are graphs whose non-trivial eigenvalues lie in the spectrum of the infinite (c,d)-biregular tree.

There is an analog of expander mixing lemma (Fact 2.1.1) for bipartite graphs (see, e.g., [Hae95]).

Fact 2.1.2 (Bipartite expander mixing lemma)**.** *Let $G = (L \cup R, E)$ be a (c,d)-biregular graph. For any subsets $S \subseteq L$ and $T \subseteq R$,*

$$\left|\frac{e(S,T)}{|E|} - \frac{|S|}{|L|} \cdot \frac{|T|}{|R|}\right| \leqslant \frac{\lambda_2}{\sqrt{cd}} \sqrt{\frac{|S|}{|L|} \cdot \frac{|T|}{|R|}}.$$

2.2 Non-backtracking matrix

Definition 2.2.1 (Non-backtracking matrix)**.** For an undirected graph $G = (V,E)$ with m edges, its non-backtracking matrix $B_G \in \{0,1\}^{2m \times 2m}$ is defined as follows: for *directed* edges $(u_1,v_1),(u_2,v_2)$ in the graph,

$$B_G[(u_1,v_1),(u_2,v_2)] = \mathbf{1}(v_1 = u_2) \cdot \mathbf{1}(u_1 \neq v_2).$$

Note that the non-backtracking matrix is *not* symmetric. Let $\lambda_1(B_G), \ldots, \lambda_{2m}(B_G) \in \mathbb{C}$ be the eigenvalues of B_G ordered such that $|\lambda_1(B_G)| \geqslant |\lambda_2(B_G)| \geqslant \cdots \geqslant |\lambda_{2m}(B_G)|$. The Perron-Frobenius theorem implies that $\lambda_1(B_G)$ is real and non-negative. We denote $\rho(B_G) = \lambda_1(B_G)$ to be the spectral radius of B_G.

A crucial identity we will need is the Ihara-Bass formula [Iha66, Bas92] which gives a translation between the eigenvalues of the adjacency matrix and the eigenvalues of the non-backtracking matrix:

Fact 2.2.2 (Ihara-Bass formula)**.** *For any graph G with n vertices and m edges, the following identity on univariate polynomials is true:*

$$\det(\mathbb{I} - B_G t) = \det(H_G(t)) \cdot (1-t^2)^{m-n}$$

where $H_G(t) := (D_G - \mathbb{I})t^2 - A_G t + \mathbb{I}$ is the Bethe Hessian *of G.*

The Ihara-Bass formula gives a direct relationship between the spectral radius of B_G and the positive definiteness of $H_G(t)$. The following is classic (e.g., [FM17, Proof of Theorem 5.1]), though we include a proof for completeness.

Lemma 2.2.3. *Let G be a graph and $0 < \alpha < 1$. Then, the spectral radius $\rho(B_G) \leqslant \frac{1}{\alpha}$ if and only if $H_G(t) \succ 0$ for all $t \in [0, \alpha)$. As a result, if $H_G(\frac{1}{\rho})$ has a non-positive eigenvalue for some $\rho > 0$, then $\rho(B_G) \geqslant \rho$.*

Proof. First observe that $H_G(0) = \mathbb{I} \succ 0$. Since $H_G(t)$ is symmetric, the eigenvalues of $H_G(t)$ are real and move continuously on the real line as t increases from 0. Note also that by the Perron-Frobenius theorem, $\rho(B_G) = \lambda_1(B_G) \geqslant 0$.

Suppose for contradiction that $\rho(B_G) \leqslant \frac{1}{\alpha}$ but $H_G(t)$ has a non-positive eigenvalue for some $t \in [0, \alpha)$. Due to $H_G(0) = \mathbb{I}$ and continuity of the eigenvalues, there must be a $t^* \in (0, t]$ such that $H_G(t^*)$ has a zero eigenvalue, meaning $\det(H_G(t^*)) = 0$. By Fact 2.2.2, this means that $\det(\mathbb{I} - B_G t^*) = 0$, i.e., B_G has an eigenvalue $\frac{1}{t^*} \geqslant \frac{1}{t} > \frac{1}{\alpha}$. This contradicts that $\rho(B_G) \leqslant \frac{1}{\alpha}$.

On the other hand, if $H_G(t) \succ 0$ for all $t \in [0, \alpha)$, then by Fact 2.2.2 $\det(\mathbb{I} - B_G t) > 0$ for all $t \in [0, \alpha)$. Since $\det(\mathbb{I} - B_G/\lambda_1) = 0$, it follows that $\frac{1}{\lambda_1} \geqslant \alpha$, i.e., $\lambda_1 \leqslant \frac{1}{\alpha}$.

Finally, suppose $H_G(\frac{1}{\rho}) \not\succ 0$ for some $\rho > 0$, then setting $\alpha = \frac{1}{\rho - \varepsilon} > \frac{1}{\rho}$ for any $\varepsilon \to 0^+$, we have that $\rho(B_G) > \rho - \varepsilon$. This implies that $\rho(B_G) \geqslant \rho$. □

2.3 Graph pruning and expander decomposition

It is a standard result that given a graph with m edges and average degree d, one can delete vertices such that the resulting graph has minimum degree εd and at least $(1 - 2\varepsilon)m$ edges. We include a short proof for completeness.

Lemma 2.3.1 (Graph pruning)**.** *Let G be an n-vertex graph with average degree d and $m = \frac{nd}{2}$ edges, and let $\varepsilon \in (0, 1/2)$. There is an algorithm that deletes vertices of G such that the resulting graph has minimum degree εd and at least $(1 - 2\varepsilon)m$ edges.*

Proof. The algorithm is simple: repeatedly remove any vertex with degree $< \varepsilon d$. First, we show by induction that each deletion cannot decrease the average degree. Suppose there are $n' \leqslant n$ vertices left and average degree $d' \geqslant d$. Then, after deleting a vertex u with degree $d_u < \varepsilon d$, the average degree becomes $\frac{n'd' - 2d_u}{n' - 1} > \frac{n'd - 2\varepsilon d}{n' - 1} = d \cdot \frac{n' - 2\varepsilon}{n' - 1}$. Thus, for $\varepsilon < 1/2$, the average degree is always at least d. Furthermore, since the algorithm can delete at most n vertices, it can delete at most $\varepsilon dn = 2\varepsilon m$ edges. □

We will also need an algorithm that partitions a graph into expanding clusters such that total number of edges across different clusters is small. Expander decomposition has been developed in a long line of work [KVV04, ST11, Wul17, SW19] and has a wide

range of applications. For our algorithm, we only require a very simple expander decomposition that recursively applies Cheeger's inequality.

Fact 2.3.2 (Expander decomposition)**.** *Given a (multi)graph $G = (V, E)$ with m edges and a parameter $\varepsilon \in (0, 1)$, there is a polynomial-time algorithm that finds a partition of V into $V_1, \ldots, V_T$ such that $\lambda_2(\widetilde{L}_{G\{V_i\}}) \geqslant \Omega(\varepsilon^2 / \log^2 m)$ for each $i \in [T]$ and the number of edges across partitions is at most εm.*

Proof. Fix $\lambda = c\varepsilon^2 / \log^2 m$ for some constant c to be chosen later. The algorithm is very simple. Given a graph $G = (V, E)$ (with potentially parallel edges and self-loops), if $\lambda_2(\widetilde{L}_G) < \lambda$, then by Cheeger's inequality we can efficiently find a subset $S \subseteq V$ with $\mathrm{vol}(S) \leqslant \mathrm{vol}(\overline{S})$ such that $\frac{|E(S,\overline{S})|}{\mathrm{vol}(S)} < \sqrt{2\lambda}$. Here $\mathrm{vol}(S) := \sum_{v \in S} \deg(v)$. Then, we cut along S, add self-loops to the induced subgraphs $G[S]$ and $G[\overline{S}]$ so that the vertex degrees remain the same (each self-loop contributes 1 to the degree). This produces two graphs $G\{S\}$ and $G\{\overline{S}\}$, and we recurse on each. By construction, in the end we will have partitions $V_1, \ldots, V_T$ where either V_i is either a single vertex or satisfies $\lambda_2(\widetilde{L}_{G\{V_i\}}) \geqslant \lambda$.

We now bound the number of edges cut via a charging argument. Consider the "half-edges" in the graph, where each edge (u, v) contributes one half-edge to u and one to v, and each self-loop counts as one half-edge. Then, $\mathrm{vol}(S)$ equals the number of half-edges attached to S. Now, imagine we have a counter for each half-edge, and every time we cut along S we add $\sqrt{2\lambda}$ to each half-edge attached to S (the smaller side). Since $E(S, \overline{S}) < \sqrt{2\lambda} \cdot \mathrm{vol}(S)$, it follows that the number of edges cut is at most the total sum of the counters. On the other hand, each half-edge can appear on the smaller side of the cut at most $\log_2 2m$ times, as each time the half-edge is on the smaller side of the cut, $\mathrm{vol}(S)$ decreases by at least a factor of 2, and $\mathrm{vol}([n]) = 2m$. So, the total sum must be $\leqslant \sqrt{2\lambda} \cdot 2m \log_2 2m \leqslant \varepsilon m$ for a small enough constant c. □

2.4 Concentration inequalities

Fact 2.4.1 (Chernoff bound)**.** *Let $X_1, \ldots, X_n$ be independent random variables taking values in $\{0, 1\}$. Let $X = \sum_{i=1}^n X_i$ and $\mu = \mathbb{E}[X]$. Then, for any $\delta \in [0, 1]$,*

$$\Pr\left[|X - \mu| \geqslant \delta\mu\right] \leqslant 2e^{-\delta^2\mu/3}.$$

Fact 2.4.2 (Matrix Chernoff [Tro15, Theorem 5.1.1])**.** *Let $X_1, \ldots, X_n \in \mathbb{R}^{d\times d}$ be independent, random, symmetric matrices such that $X_i \succeq 0$ and $\lambda_{\max}(X_i) \leqslant R$ almost surely. Let $X = \sum_{i=1}^n X_i$ and $\mu = \lambda_{\max}(\mathbb{E}[X])$. Then, for any $\delta \in [0, 1]$,*

$$\Pr\left[\lambda_{\max}(X) \geqslant (1+\delta)\mu\right] \leqslant d \cdot \exp\left(-\frac{\delta^2\mu}{3R}\right).$$

2.5 The Sum-of-Squares algorithm

We refer the reader to the monograph [FKP19] and the lecture notes [BS16] for a detailed exposition of the sum-of-squares method and its usage in algorithm design.

Pseudo-distributions. Pseudo-distributions are generalizations of probability distributions. Formally, a pseudo-distribution on $\mathbb{R}^n$ is a finitely supported *signed* measure $\mu : \mathbb{R}^n \to \mathbb{R}$ such that $\sum_x \mu(x) = 1$. The associated *pseudo-expectation* is a linear operator $\widetilde{\mathbb{E}}_\mu$ that assigns to every polynomial $f : \mathbb{R}^n \to \mathbb{R}$ the value $\widetilde{\mathbb{E}}_\mu f = \sum_x \mu(x) f(x)$, which we call the pseudo-expectation of f. We say that a pseudo-distribution μ on $\mathbb{R}^n$ has *degree* d if $\widetilde{\mathbb{E}}_\mu[f^2] \geqslant 0$ for every polynomial f on $\mathbb{R}^n$ of degree $\leqslant d/2$.

A degree-d pseudo-distribution μ is said to satisfy a constraint $\{q(x) \geqslant 0\}$ for any polynomial q of degree $\leqslant d$ if for every polynomial p such that $\deg(p^2) \leqslant d - \deg(q)$, $\widetilde{\mathbb{E}}_\mu[p^2 q] \geqslant 0$. For example, in this work we will often say that μ satisfies the Booleanity constraints $\{x_i^2 - x_i = 0,\ \forall i \in [n]\}$, which means that $\widetilde{\mathbb{E}}_\mu[p(x)(x_i^2 - x_i)] = 0$ for any i and any polynomial p of degree $d - 2$. We say that μ τ-approximately satisfies a constraint $\{q \geqslant 0\}$ if for any sum-of-squares polynomial p, $\widetilde{\mathbb{E}}_\mu[pq] \geqslant -\tau \|p\|_2$ where $\|p\|_2$ is the ℓ_2 norm of the coefficient vector of p.

We rely on the following basic connection that forms the basis of the sum-of-squares algorithm.

Fact 2.5.1 (Sum-of-Squares algorithm, [Par00, Las01])**.** *Given a system of degree $\leqslant d$ polynomial constraints $\{q_i \geqslant 0\}$ in n variables and the promise that there is a degree-d pseudo-distribution satisfying $\{q_i \geqslant 0\}$ as constraints, there is a $n^{O(d)}\,\mathrm{polylog}(1/\tau)$ time algorithm to find a pseudo-distribution of degree d on $\mathbb{R}^n$ that τ-approximately satisfies the constraints $\{q_i \geqslant 0\}$.*

Sum-of-squares proofs. Let $f_1, f_2, \ldots, f_m$ and g be multivariate polynomials in x. A *sum-of-squares* proof that the constraints $\{f_1 \geqslant 0, \ldots, f_m \geqslant 0\}$ imply $g \geqslant 0$ consists of sum-of-squares polynomials $(p_S)_{S \subseteq [m]}$ such that $g = \sum_{S \subseteq [m]} p_S \prod_{i \in S} f_i$. The *degree* of such a sum-of-squares proof equals the maximum of the degree of $p_S \prod_{i \in S} f_i$ over all S appearing in the sum above. We write $\{f_i \geqslant 0,\ \forall i \in [m]\} \left|\frac{x}{d}\right. \{g \geqslant 0\}$ where d is the degree of the sum-of-squares proof.

We will rely on the following basic connection between SoS proofs and pseudo-distributions:

Fact 2.5.2. *Let $f_1, \ldots, f_m$ and g be polynomials, and let $\mathcal{A} = \{f_i(x) \geqslant 0,\ \forall i \in [m]\}$. Suppose $\mathcal{A} \left|\frac{x}{d}\right. \{g(x) \geqslant 0\}$. Then, for any pseudo-distribution μ of degree $\geqslant d$ satisfying $\mathcal{A}$, we have $\widetilde{\mathbb{E}}_\mu[g] \geqslant 0$.*

Therefore, an SoS proof of some polynomial inequality directly implies that the same inequality holds in pseudo-expectation. We will use this repeatedly in our analysis.

2.5.1 Sum-of-Squares toolkit

The theory of univariate sum-of-squares (in particular, Lukács Theorem) says that if a univariate polynomial is non-negative on an interval, then this fact is also SoS-certifiable. The following corollary of Lukács theorem is well-known, and we will use it multiple times to convert univariate inequalities into SoS inequalities in a blackbox manner.

Fact 2.5.3 (Corollary of Lukács Theorem)**.** *Let $a \leqslant b \in \mathbb{R}$. Let $p \in \mathbb{R}[x]$ be a univariate real polynomial of degree d such that $p(x) \geqslant 0$ for all $a \leqslant x \leqslant b$. Then,*

$$\{x \geqslant a,\ x \leqslant b\} \Big|\frac{x}{d} \{p(x) \geqslant 0\}\ .$$

Similarly, true inequalities on the hypercube are also SoS-certifiable.

Fact 2.5.4. *Let p be a polynomial in n variables. Suppose $p(x) \geqslant 0$ for all $x \in \{0,1\}^n$, then*

$$\left\{x_i^2 - x_i = 0,\ \forall i \in [n]\right\} \Big|\frac{x}{\max(n,\deg(p))} \{p(x) \geqslant 0\}\ .$$

More generally, all true inequalities have SoS certificates under mild assumptions. In particular, Schmüdgen's Positivstellensatz establishes the completeness of the SoS proof system under compactness conditions (often called the Archimedean condition). Moreover, bounds on the SoS degree (given the polynomial and the constraints) were given in [PD01, Sch04].

Fact 2.5.5 (Positivstellensatz [PD01, Sch04])**.** *For all polynomials $g_1, g_2, \ldots, g_m$ over $x = (x_1, x_2, \ldots, x_n)$ defining a non-empty set*

$$S := \{x \in \mathbb{R}^n : g_1(x) \geqslant 0, \ldots, g_m(x) \geqslant 0\} \subseteq (-1,1)^n\,,$$

and for every polynomial f of degree d with coefficients bounded by R and $f^* := \min_{x \in S} f(x) > 0$, there exists an integer $D = D(n, g_1, \ldots, g_m, R, f^*) \in \mathbb{N}$ such that

$$\{g_1 \geqslant 0, \ldots, g_m \geqslant 0\} \Big|\frac{x}{D} \{f \geqslant 0\}\ .$$

Independent samples from a pseudo-distribution. Recall that a pseudo-expectation operator $\widetilde{\mathbb{E}}_\mu$ can be interpreted as the average of functions $f(x)$ over a pseudo-distribution $x \sim \mu$. We will need to be able to mimic averaging over t independently chosen samples $x^{(1)}, \ldots, x^{(t)} \sim \mu$. We define the product pseudo-distribution $\mu^{\otimes t}$ along with pseudo-expectation $\widetilde{\mathbb{E}}_{\mu^{\otimes t}}$ as follows: let $p(\boldsymbol{x}) = (x^{(1)})^{\alpha_1} \ldots (x^{(t)})^{\alpha_t}$ be a monomial in variables $\boldsymbol{x} = (x^{(1)}, \ldots, x^{(t)})$; we define

$$\widetilde{\mathbb{E}}_{\mu^{\otimes t}}[p] := \widetilde{\mathbb{E}}_\mu[x^{\alpha_1}] \cdot \widetilde{\mathbb{E}}_\mu[x^{\alpha_2}] \cdots \widetilde{\mathbb{E}}_\mu[x^{\alpha_t}]\,.$$

It is easy to check that $\widetilde{\mathbb{E}}_{\mu^{\otimes t}}$ is also a pseudo-expectation operator corresponding to t independent samples from the pseudo-distribution μ.

Fact 2.5.6. *If μ is a valid pseudo-distribution of degree D in variables x, then $\mu^{\otimes t}$ is a valid pseudo-distribution of degree D. Furthermore, if additional SoS inequalities are true for μ, they also hold for $\mu^{\otimes t}$.*

2.5.2 Conditioning pseudo-distributions

We can *reweigh* or *condition* a degree-D pseudo-distribution μ by a polynomial $s(x)$, where $s(x)$ is non-negative under the program axioms, i.e., $\mathcal{A} \left|\frac{x}{d}\right. \{s(x) \geqslant 0\}$ for $d < D$. Technically, this operation defines a new pseudo-distribution μ' of degree $D - d$ with pseudo-expectation operator $\widetilde{\mathbb{E}}_{\mu'}$ by taking

$$\widetilde{\mathbb{E}}_{\mu'}[x^\alpha] = \frac{\widetilde{\mathbb{E}}_{\mu}[x^\alpha \cdot s(x)]}{\widetilde{\mathbb{E}}_{\mu}[s(x)]},$$

for every monomial x^α of degree at most $D - d$.

It is easy to verify that μ' is a valid pseudo-distribution of degree $D - d$ and satisfies the axioms of the original μ. As an example, under the independent set axioms presented in (9.1), since $x_i \geqslant 0$ is an axiom, one can reweigh μ by x_i, essentially "conditioning" the pseudo-distribution on the event $x_i = 1$. Thus, we will also refer to this operation as conditioning and denote μ' by $\mu|s(x)$. Often times, the polynomial $s(x)$ we will "condition" on will be a polynomial approximation of the indicator function of some event E. In this case, the above operation can be interpreted as conditioning μ to satisfy some properties specified by the event E.

Reducing average correlation. An important technique we need is reducing the average correlation of random variables through iterative conditioning, which was introduced in [BRS11] (termed *global correlation reduction*) and is also applicable to pseudo-distributions of sufficiently large degree. We will use the following version from [RT12].

Lemma 2.5.7 (Lemma 4.5 of [RT12])**.** *Let $Y_1, \ldots, Y_M$ be a set of random variables each taking values in $\{1, \ldots, q\}$. Then, for any $\ell \in \mathbb{N}$, there exists $k \leqslant \ell$ such that:*

$$\mathbb{E}_{i_1,\ldots,i_k \sim [M]} \mathbb{E}_{i,j \sim [M]} [I(Y_i; Y_j \mid Y_{i_1}, \ldots, Y_{i_k})] \leqslant \frac{\log q}{\ell - 1}.$$

Note that the above lemma holds as long as there is a local collection of distributions over $(Y_1, \ldots, Y_M)$ that are valid probability distributions over all collections of $\ell + 2$ variables and are consistent with each other. Of particular interest to us would be the setting where we have a degree $\geqslant \ell + 2$-pseudo-distribution μ over the variables $(Y_1, \ldots, Y_M)$.

Lemma 2.5.7 is stated in terms of mutual information between pairs of variables. One can translate it to a bound on average correlation (in L_1) via the following:

Fact 2.5.8 (Pinsker's inequality)**.** *Given any two distributions D_1, D_2:*

$$TV(D_1, D_2) \leqslant \sqrt{\frac{1}{2} D_{KL}(D_1 \| D_2)}\,.$$

We also require a generalization of Lemma 2.5.7 to t-wise correlations.

Lemma 2.5.9 (Lemma 32 of [MR17])**.** *Let $Y_1, \ldots, Y_M$ be a set of random variables each taking values in $\{1, \ldots, q\}$. The total t-wise correlation of a distribution μ over $Y_1, \ldots, Y_M$ is defined as*

$$\mathrm{TC}_t(\mu) := \mathbb{E}_{i_1,\ldots,i_t \sim [M]} \left[\mathrm{KL}((Y_{i_1}, \ldots, Y_{i_t}) \| Y_{i_1} \times \cdots \times Y_{i_t}) \right].$$

Then, for any $\ell \in \mathbb{N}$, there exists $k \leqslant \ell$ such that:

$$\mathbb{E}_{\substack{i_1,\ldots,i_k \sim [M] \\ (y_{i_1},\ldots,y_{i_k}) \sim \mu}} \left[\mathrm{TC}_t(\mu \mid Y_{i_1} = y_{i_1}, \ldots, Y_{i_k} = y_{i_k}) \right] \leqslant \frac{t^2 \log q}{\ell}\,.$$

Similar to Lemma 2.5.7, the above holds for pseudo-distributions of degree $\geqslant \ell + t$.

Part I

Graph Theory

Chapter 3

Introduction

In Part I, we present new results in graph theory that serve both as standalone contributions and as key tools for later chapters.

(1) **Section 3.1 and Chapter 4: Girth-density trade-off in graphs and hypergraphs.** First, we generalize the classical Moore bound [AHL02] on the girth-density trade-offs in graphs. Then, we study the hypergraph analogue, known as the *hypergraph Moore bound*, which was conjectured by Feige [Fei08] and resolved up to polylogarithmic factors by Guruswami, Kothari, and Manohar [GKM22]. We present a substantially simpler and shorter proof that also improves the bound to within a single extra log factor. This chapter is based on [HKM23].

(2) **Section 3.2 and Chapter 5: Subgraph density in spectral expanders.** We give a sharp upper bound on the density of subgraphs in bipartite spectral expanders, generalizing results of Kahale [Kah95] and Asherov and Dinur [AD24]. This chapter is based on [HMMP24].

These results are proved using various tools from spectral graph theory, with a general theme of analyzing walks on graphs (or certain graphs constructed from hypergraphs) and relating them to the spectra of the adjacency or non-backtracking matrices.

These results are not only of independent interest but also play important roles in later chapters. The analysis of the hypergraph Moore bound will be extended to refutation algorithms for semirandom CSPs in Chapter 7. Similarly, the generalized Moore bound and the subgraph density bounds will be used in explicit constructions of vertex expanders in Chapter 12.

3.1 Girth-density trade-off in hypergraphs

What is the maximum girth of a graph on n vertices and average degree d? For d-regular graphs, a simple "ball growing" argument shows that the graph must have a cycle of

length at most $2\log_{d-1} n + 2$. This threshold is called the *Moore bound* [Wik22] (see also Page 180 of [Big93]), and graphs achieving it are called Moore graphs. In a classical paper that resolved a question of Bollobás [Bol78], Alon, Hoory and Linial [AHL02] proved that the same upper bound holds even for irregular graphs of average degree $d > 2$, including the case where d is not an integer.

Theorem 3.1.1 (Moore bound [AHL02])**.** *For any $d > 2$, an n-vertex graph of average degree d must contain a cycle of length at most $2\log_{d-1} n + 2$.*

Later, Hoory [Hoo02] obtained a better bound for bipartite graphs, and Babu and Radhakrishnan [BR14] found an elegant proof based on the entropy of random walks on the graph.

The Moore bound resolves a natural graph Turán problem — such problems, more generally, study the maximum number of edges that one can pack in a graph while avoiding a given forbidden structure. Turán's original work on triangle free graphs marks the birth of extremal graph theory, and its various generalizations form a centerpiece of modern extremal combinatorics.

3.1.1 Generalized Moore bound

We present an improvement of the Moore bound that bounds the maximum girth of a graph in terms of its non-backtracking matrix (Definition 2.2.1). In addition to showing the existence of short cycles, we also show the existence of short *bicycles* (also known as "tangles" in [Bor20]).

Definition 3.1.2 (Bicycles [MOP20])**.** A graph $H = (V, E)$ is cyclic if $|E| - |V| = -1$, and bicyclic if $|E| - |V| = 0$.

Theorem 3.1.3 (Generalized Moore bound)**.** *Suppose G is a graph on n vertices, and let $\rho = \rho(B_G)$ be the spectral radius of its non-backtracking matrix B_G. Suppose $\rho > 1$, then G contains a cycle of size at most $2(\lfloor \log_\rho n \rfloor + 1)$ and a bicycle of size at most $3(\lfloor \log_\rho 2n \rfloor + 1)$.*

For a graph G with average degree d, $\rho(B_G)$ is at least $d - 1$. This follows from the fact that $\vec{1}^\top H_G(\frac{1}{d-1})\vec{1} = 0$ and Lemma 2.2.3 ($H_G(t)$ is defined in Fact 2.2.2). Therefore, Theorem 3.1.3 is at least as strong as the girth guarantee of $2\log_{d-1} n$ from the classical Moore bound (Theorem 3.1.1). A simple example where this yields tighter bounds is a $(d, 2)$-biregular graph. When $d \gg 2$, the average degree is ≈ 4 and the classical Moore bound yields a cycle of length $2\log_3 n$. Nevertheless, the generalized Moore bound tells us that there is a cycle of length $\approx 4\log_{d-1} n$, which is much smaller than $2\log_3 n$ when d is large. In Chapter 12, we will also use Theorem 3.1.3 to analyze certain unbalanced bipartite graphs.

We prove Theorem 3.1.3 in Section 4.1, utilizing a key connection (Fact 4.1.1) between

non-backtracking walks in the graph G and the Bethe Hessian $H_G(t)$ (recall its definition in Fact 2.2.2).

3.1.2 Girth-density trade-off for hypergraphs

We next shift our attention to hypergraphs. A cycle[1] in a hypergraph, more descriptively called an *even cover*, is a collection of hyperedges such that every vertex participates in an even number of them. Formally,

Definition 3.1.4 (Even cover)**.** For a hypergraph $\mathcal{H}$, a set S of hyperedges in $\mathcal{H}$ is an even cover if

$$\bigoplus_{C \in S} C := \{v \in V(\mathcal{H}) : v \text{ belongs to an odd number of hyperedges in } S\} = \emptyset .$$

Equivalently, S is an even cover if $\sum_{C \in S} \mathbf{1}_C = \mathbf{0}$ over $\mathbb{F}_2$, where $\mathbf{1}_C$ denotes the characteristic vector of C. Therefore, an even cover in a k-uniform hypergraphs is exactly a linearly dependent subset of a system of k-sparse linear equations over $\mathbb{F}_2$.

The girth of a hypergraph is the smallest size of an even cover in it. When specialized to graphs, an even cover is simply a union of cycles, i.e., a subgraph with all vertices of even degree. Thus, this formulation naturally generalizes the standard notion of girth in graphs.

Like the graph Moore bound, girth-density trade-offs for hypergraphs have foundational connections to theoretical computer science. Due to the equivalence between even covers and linear dependencies, the girth-density trade-offs for k-uniform hypergraphs correspond to the rate vs. distance trade-offs for *low density parity check* (LDPC) codes. As a result, there is an extensive line of work that studies the girth-density trade-offs for hypergraphs (see e.g. [BKHL99, BMS08, AF09]).

For k-uniform hypergraphs $\mathcal{H}$ with $k > 2$, this trade-off was first studied by Naor and Verstraëte [NV08]. They showed that every $\mathcal{H}$ with $m \geqslant n^{k/2} \log^{O(1)}(n)$ hyperedges on n vertices must contain an even cover of length $O(\log n)$. The $\log^{O(1)}(n)$ factor was further improved to a $O(\log \log n)$ factor in a subsequent work of Feige [Fei08]. For $k = 2$, this recovers a coarse version of the irregular Moore bound. For $k > 2$, however, there is an interesting regime between the two extreme thresholds of $m = n + 1$ (with maximum possible girth of $n + 1$) and $m \sim n^{k/2} \log^{O(1)}(n)$ (with maximum possible girth of $O(\log n)$).

Hypergraph Moore bound. Feige [Fei08] formulated a conjecture about this regime that suggests a smooth interpolation between the two extremes noted above.

[1]There are several well-studied combinatorial notions of cycles in contrast to the more linear algebraic notion of even covers.

Conjecture 3.1.5 (Conjecture 1.2 of [Fei08]). *For every $k \in \mathbb{N}$ and $1 \leqslant r \leqslant n$, every k-uniform hypergraph with n vertices and $m \gtrsim n(\frac{n}{r})^{\frac{k}{2}-1}$ hyperedges has an even cover of size $O(r \log n)$.*

The quantitative behavior above can be verified for random hypergraphs (up to a multiplicative factor of $\log(n)$ in m). Indeed, Feige's conjecture was based on the hypothesis that random hypergraphs are approximately extremal for the purpose of avoiding short even covers.

Feige's conjecture was settled by Guruswami, Kothari and Manohar [GKM22] up to an additional $\log^{4k+1} n$ multiplicative factor in the density m. Their proof goes via a spectral argument applied to the *Kikuchi* graph [WAM19] — a graph with an appropriate algebraic structure, built from the given hypergraph (see Definition 4.3.2).

While [GKM22] begins with an elegant and simple observation, their technical analysis, especially for odd k (the "hard" case), is quite complicated and involves bucketing and pruning the Kikuchi matrix and invoking the Schudy–Sviridenko concentration inequality [SS12] for polynomials with combinatorial structure in the monomials. As a consequence, even for the simplest case of $k = 2$ (i.e., recovering the classical Moore bound), their proof incurs an additional $\log^3 n$ factor.

We give a simple and short proof of the hypergraph Moore bound that is *almost* tight up to a single logarithmic factor.

Theorem 3.1.6 (Hypergraph Moore bound). *For every $k \in \mathbb{N}$, there exists a constant $C > 0$ such that for any $1 \leqslant r \leqslant n$, every hypergraph on n vertices and $m \geqslant Cn(\frac{n}{r})^{\frac{k}{2}-1} \log n$ hyperedges has an even cover of size $O(r \log n)$.*

In a follow-up work [HKMMS25] which will not be included in this thesis, we found a simple and purely combinatorial argument that recovers Theorem 3.1.6. Moreover, for odd k, we introduced a variant of the Kikuchi graph and improved the $\log n$ factor in the density m to $(\log n)^{\frac{1}{k+1}}$. We believe that the hypergraph Moore bound is true without any extra log factors.

Our key idea is the use of a new *reweighted* Kikuchi matrix and an *edge deletion* trick (for odd k). These allow us to drop several involved steps in [GKM22]'s analysis such as combinatorial bucketing of rows of the Kikuchi matrix and the use of the Schudy–Sviridenko polynomial concentration.

As an illustration of the power of our reweighting idea, in Section 4.2 we will give a simple proof of the classical Moore bound (Theorem 3.1.1) that is tight up to an absolute constant factor (as opposed to the $\log^3 n$ loss incurred by the strategy of [GKM22]). In Section 4.3, we will generalize the reweighting idea to prove hypergraph Moore bound for all even k. Finally in Section 4.4, we combine the reweighting idea and the key new idea of *edge deletions* to prove the case of odd k.

3.2 Subgraph density in spectral expanders

Expander graphs are widely used in computer science, with applications spanning coding theory, complexity theory, cryptography, and more. Intuitively, expanders are sparse graphs that exhibit strong connectivity properties, where every (small) subset of vertices has many neighbors and contains few internal edges. We refer readers to the survey of [HLW06] for a thorough exposition on expander graphs.

While these expansion properties can be defined combinatorially, they are often studied analytically via *spectral expansion*, which concerns the non-trivial eigenvalues of the graph's adjacency matrix. Spectral expansion is closely related to *edge expansion* via Cheeger's inequality, and it is a very natural notion of expansion as it directly governs the mixing time of random walks on the graph. Moreover, several explicit constructions of Ramanujan graphs (i.e., optimal spectral expanders; see Section 2.1) are known [LPS88, Mar88, Mor94].

Therefore, a fundamental question is: to what extent does spectral expansion — i.e., bounds on the non-trivial eigenvalues — imply vertex expansion?

3.2.1 Expander mixing lemma and Kahale's improvement

Let $G = (V, E)$ be a d-regular graph on n vertices, and let $\lambda = \max\{\lambda_2, |\lambda_n|\}$. Consider a set $S \subseteq V$ with $|S| \leqslant \delta n$, where $\delta = \delta(d, \varepsilon)$ is a small constant that depends only on d and $\varepsilon \in (0, 1)$. Let $N(S) := \{v \in V : \exists u \in S,\ (u, v) \in E\}$ be the set of neighbors of S. By the expander mixing lemma (Fact 2.1.1),

$$d|S| = e(S, N(S)) \leqslant \frac{d}{n}|S||N(S)| + \lambda\sqrt{|S||N(S)|}\,.$$

Since $|N(S)| \leqslant d|S|$ trivially, for δ small enough depending on d, ε, we have

$$|N(S)| \geqslant \frac{d^2}{\lambda^2}|S|(1-\varepsilon)\,. \tag{3.1}$$

Similarly, we can upper bound the density of the subgraph $G[S]$ as follows:

$$e(S, S) \leqslant \frac{d}{n}|S|^2 + \lambda|S| \leqslant \lambda|S| \cdot (1+\varepsilon)\,.$$

Here, note that the notation $e(S, S)$ double counts each edge within S. Thus, the average degree of the induced subgraph $G[S]$ is at most

$$\frac{e(S, S)}{|S|} \leqslant \lambda \cdot (1+\varepsilon)\,. \tag{3.2}$$

For d-regular Ramanujan graphs with $\lambda \leqslant 2\sqrt{d-1}$, Eq. (3.1) and (3.2) imply that the vertex expansion $|N(S)| \geqslant \frac{d}{4}|S|$ and the average degree of $G[S]$ is at most $2\sqrt{d-1}$ (up to the $(1 \pm \varepsilon)$ factor).

Kahale's improvement. Kahale [Kah95] showed:

Theorem 3.2.1 (Expansion and subgraph density in spectral expanders [Kah95])**.** *Let $G = (V, E)$ be a d-regular n-vertex graph with eigenvalues $d = \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$, and let $\lambda = \max\{\lambda_2, |\lambda_n|, 2\sqrt{d-1}\}$. Then, for $\varepsilon \in (0,1)$, for any nonempty subset $S \subseteq V$ with $|S| \leqslant d^{-1/\varepsilon}n$,*

$$\frac{|N(S)|}{|S|} \geqslant \frac{d}{2}\left(1 - \sqrt{1 - \frac{4(d-1)}{\lambda^2}}\right)(1 - O(\varepsilon)).$$

Moreover, the average degree of the induced subgraph $G[S]$ is at most

$$\left(1 + \frac{\lambda}{2} + \sqrt{\frac{\lambda^2}{4} - (d-1)}\right)(1 + O(\varepsilon)).$$

For d-regular Ramanujan graphs where $\lambda = 2\sqrt{d-1}$, the above implies that the expansion is at least $\frac{d}{2}$ and the average degree is at most $\sqrt{d-1}+1$. These are factors of 2 better than the guarantees (Eq. (3.1) and (3.2)) from expander mixing lemma.

The $d/2$ barrier for spectral expanders. Kahale [Kah95] also showed that the $d/2$ vertex expansion for near-Ramanujan graphs is tight. [KK22] further showed that several well-known explicit algebraic constructions of d-regular Ramanujan graphs contain subsets of vertices S with exactly $d/2 \cdot |S|$ neighbors and *zero* unique-neighbors. Constructing vertex expanders that surpass the $d/2$ barrier has been a longstanding open problem, one that we will resolve in Part III.

3.2.2 Bipartite spectral expanders

We now look at biregular graphs. Recall from Section 2.1 that a (c,d)-biregular graph $G = (L \cup R, E)$ has $\lambda_1 = \sqrt{cd}$, and G is Ramanujan if $\lambda_2 \leqslant \sqrt{c-1} + \sqrt{d-1}$. Without loss of generality, we assume that $c \leqslant d$, thus $|L| \geqslant |R|$ since $|E| = c|L| = d|R|$.

For a subset $S \subseteq L$ with $|S| \leqslant \delta|L|$, where $\delta = \delta(\varepsilon, c, d) > 0$ is small enough, the bipartite expander mixing lemma (Fact 2.1.2; see [AD24, Claim 4]) gives

$$\frac{c|S|}{|N(S)|} \leqslant (1+\varepsilon)\left(1 + \frac{d-1}{c-1} + 2\sqrt{\frac{d-1}{c-1}}\right). \tag{3.3}$$

Note that by flipping the left and right side, the above actually shows *lossless expansion* from the smaller side — that is, if $c \ll d$, then for small subsets $T \subseteq R$, we have $|N(T)| \geqslant (1-\varepsilon)d|T|$.

For the expansion from the larger side, Asherov and Dinur [AD24] proved an improvement over Eq. (3.3) for bipartite Ramanujan graphs.

Theorem 3.2.2 (Expansion in bipartite Ramanujan graphs; [AD24, Theorem 2])**.** *Let $G = (L \cup R, E)$ be a (c,d)-biregular graph with second eigenvalue $\lambda_2 \leqslant \sqrt{c-1} + \sqrt{d-1}$, and let $\varepsilon \in (0,1)$. Then, there exists $\delta = \delta(\varepsilon, c, d) > 0$ such that for any nonempty subset $S \subseteq L$ with $|S| \leqslant \delta |L|$,*

$$\frac{c|S|}{|N(S)|} \leqslant 1 + (1+\varepsilon)\sqrt{\frac{d-1}{c-1}}\,.$$

Therefore, if $c \ll d$, then Theorem 3.2.2 is a significant improvement over the guarantee from expander mixing lemma (Eq. (3.3)).

Interestingly, the proof of [AD24] requires the graph to be *exactly Ramanujan*. However, many existing constructions of bipartite expanders, including those with additional useful properties, are only near-Ramanujan with $\lambda_2 \leqslant (\sqrt{c-1} + \sqrt{d-1})(1+\gamma)$ (for an arbitrarily small constant γ) [MOP20, OW20].

We prove the following general result on the subgraph density in near-Ramanujan bipartite graphs. In particular, it implies that the bound in Theorem 3.2.2 also holds (approximately) for near-Ramanujan graphs.

Theorem 3.2.3 (Subgraph density in near-Ramanujan graphs)**.** *Let $3 \leqslant c \leqslant d$ be integers, $\gamma \in [0,1]$, and $\varepsilon \in (0, 0.1)$. Let $G = (L \cup R, E)$ be a (c,d)-biregular graph such that $\lambda_2 \leqslant (\sqrt{c-1} + \sqrt{d-1})(1 + \gamma/d)$. Then, there exists $\delta = \delta(\varepsilon, c, d) > 0$ such that for every $S_1 \subseteq L$ and $S_2 \subseteq R$ with $|S_1| + |S_2| \leqslant \delta |L \cup R|$, the left and right average degrees $d_1 = \frac{|E(S_1,S_2)|}{|S_1|}$ and $d_2 = \frac{|E(S_1,S_2)|}{|S_2|}$ in the induced subgraph $G[S_1 \cup S_2]$ must satisfy*

$$(d_1 - 1)(d_2 - 1) \leqslant \sqrt{(c-1)(d-1)} \cdot (1 + O(\varepsilon + \sqrt{\gamma}))\,.$$

In fact, we will prove a more general result in Theorem 3.2.6 where the bound is stated explicitly in terms of λ_2. Theorem 3.2.3 then follows by substituting $\lambda_2 \approx \sqrt{c-1} + \sqrt{d-1}$.

Theorem 3.2.3 will be used repeatedly in Chapter 12 of Part III to analyze our construction of unique-neighbor expanders.

3.2.3 Non-backtracking matrix of subgraphs in bipartite expanders

Expander mixing lemma bounds the subgraph density by the spectrum of the adjacency matrix. In the following, we show a similar bound using the spectral radius of the non-backtracking matrix (Definition 2.2.1).

Lemma 3.2.4. *Let $G = (L \cup R, E)$ be a bipartite graph, and let the left and right average degrees be $d_1 = \frac{|E|}{|L|}$ and $d_2 = \frac{|E|}{|R|}$, respectively. Then,*

$$(d_1 - 1)(d_2 - 1) \leqslant \rho(B_G)^2\,.$$

Then, we show that the non-backtracking matrix of small subgraphs in a biregular expander must be small.

Theorem 3.2.5. *Let $\varepsilon \in (0, 0.1)$, and let $3 \leqslant c \leqslant d$ be integers. Let $G = (L \cup R, E)$ be a (c,d)-biregular graph and $S \subseteq L \cup R$ such that $|S| \leqslant d^{-1/\varepsilon}|L \cup R|$. Then, for any $t \geqslant 0$ such that*

$$\rho(B_{G[S]}) \leqslant \frac{1}{2}\left(\sqrt{\lambda^2 - (\sqrt{c-1}+\sqrt{d-1})^2} + \sqrt{\lambda^2 - (\sqrt{c-1}-\sqrt{d-1})^2}\right),$$

where $\lambda = \max(\lambda_2(A_G), \sqrt{c-1}+\sqrt{d-1}) \cdot (1 + O(\varepsilon))$.

We will prove Lemma 3.2.4 and Theorem 3.2.5 in Chapter 5. Lemma 3.2.4 and Theorem 3.2.5 directly imply the following, which is the more general version of Theorem 3.2.3:

Theorem 3.2.6. *Let $3 \leqslant c \leqslant d$ be integers, $\gamma \in [0,1]$, and $\varepsilon \in (0, 0.1)$. Let $G = (L \cup R, E)$ be a (c,d)-biregular graph, and let $\lambda = \max(\lambda_2(A_G), \sqrt{c-1}+\sqrt{d-1}) \cdot (1 + O(\varepsilon))$. Then, there exists $\delta = \delta(\varepsilon, c, d) > 0$ such that for every $S_1 \subseteq L$ and $S_2 \subseteq R$ with $|S_1| + |S_2| \leqslant \delta|L \cup R|$, the left and right average degrees $d_1 = \frac{|E(S_1,S_2)|}{|S_1|}$ and $d_2 = \frac{|E(S_1,S_2)|}{|S_2|}$ in the induced subgraph $G[S_1 \cup S_2]$ must satisfy*

$$(d_1 - 1)(d_2 - 1) \leqslant \frac{1}{4}\left(\sqrt{\lambda^2 - (\sqrt{c-1}+\sqrt{d-1})^2} + \sqrt{\lambda^2 - (\sqrt{c-1}-\sqrt{d-1})^2}\right)^2.$$

To understand Theorems 3.2.5 and 3.2.6, consider $\lambda \approx \sqrt{c-1}+\sqrt{d-1}$. Then, the bound in Theorem 3.2.5 simplifies to

$$\rho(B_{G[S]}) \lesssim ((c-1)(d-1))^{1/4},$$

and the bound in Theorem 3.2.6 implies to

$$(d_1 - 1)(d_2 - 1) \lesssim \sqrt{(c-1)(d-1)}.$$

More specifically, we will plug in $\lambda_2 = (\sqrt{c-1}+\sqrt{d-1})(1+\gamma/d)$ to prove Theorem 3.2.3:

Proof of Theorem 3.2.3. Suppose $\lambda = (\sqrt{c-1}+\sqrt{d-1})(1+\gamma/d)$ for some $\gamma \geqslant 0$. Then, denoting $\eta := (\sqrt{c-1}+\sqrt{d-1})^2(\frac{2\gamma}{d} + \frac{\gamma^2}{d^2})$, a straightforward calculation shows that Theorem 3.2.5 implies

$$\rho(B_{G[S_1 \cup S_2]}) \leqslant \sqrt{\sqrt{(c-1)(d-1)} + \eta/4} + \frac{1}{2}\sqrt{\eta}.$$

In particular, for $\gamma \leqslant 2$, the above is at most $\sqrt{(c-1)(d-1)} \cdot (1 + 3\sqrt{\gamma})$. Replacing γ with $\gamma + \varepsilon^2$ completes the proof of Theorem 3.2.3. □

Chapter 4

Girth-Density Trade-Off in Hypergraphs

In Section 4.1, we prove Theorem 3.1.3 which strengthens the classical Moore bound of Alon, Hoory and Linial [AHL02] and generalizes the result to *bicycles*.

Theorem (Restatement of Theorem 3.1.3)**.** *Suppose G is a graph on n vertices, and let $\rho = \rho(B_G)$ be the spectral radius of its non-backtracking matrix B_G. Suppose $\rho > 1$, then G contains a cycle of size at most $2(\lfloor \log_\rho n \rfloor + 1)$ and a bicycle of size at most $3(\lfloor \log_\rho 2n \rfloor + 1)$.*

Then, in Sections 4.3 and 4.4, we prove the hypergraph Moore bound.

Theorem (Restatement of Theorem 3.1.6)**.** *For every $k \in \mathbb{N}$, there exists a constant $C > 0$ such that for any $1 \leqslant r \leqslant n$, every hypergraph on n vertices and $m \geqslant Cn(\frac{n}{r})^{\frac{k}{2}-1} \log n$ hyperedges has an even cover of size $O(r \log n)$.*

As a warm-up, in Section 4.2 we present a simple, alternative proof of a weaker version of the classical Moore bound, using the strategy of "counting reweighted walks" — a key technique that we will use in the subsequent sections. In Section 4.3, we prove the hypergraph Moore bound for even k, with almost the same proof as in Section 4.2. Finally, in Section 4.4, we handle the more involved case of odd k.

4.1 Generalization of the Moore bound

The proof of Theorem 3.1.3 is based on *non-backtracking walks*, which are walks such that no edge is the inverse of its preceding edge. For a graph G on n vertices with adjacency matrix A, we define $A^{(s)}$ to be the $n \times n$ matrix whose (u, v) entry counts the number of length-s non-backtracking walks between vertices u and v in G. The following is a standard fact.

Fact 4.1.1 (Recurrence and generating function of $A^{(s)}$)**.** *The non-backtracking matrices $A^{(s)}$ satisfy the following recurrence:*

$$\begin{aligned} A^{(0)} &= \mathbb{I}\,, \\ A^{(1)} &= A\,, \\ A^{(2)} &= A^2 - D\,, \\ A^{(s)} &= A^{(s-1)}A - A^{(s-2)}(D-\mathbb{I})\,, \quad s > 2\,. \end{aligned}$$

The recurrences imply that these matrices have a generating function:

$$J(t) := \sum_{s=0}^{\infty} A^{(s)} t^s = (1-t^2)\cdot H(t)^{-1} \qquad \text{for } t \in [0,1)\,,$$

whenever the series converges. Here, we recall that $H(t) = (D-\mathbb{I})t^2 - At + \mathbb{I}$.

We first prove the following lemma,

Lemma 4.1.2. *Let $s, k \in \mathbb{N}$, $s \geqslant k$, and let q, r be the quotient and remainder of s divided by k, i.e. $s = qk + r$. Then,*

$$\operatorname{tr}(A^{(s)}) \leqslant \sqrt{n}\cdot \|A^{(k)}\|_2^q \cdot \|A^{(r)}\|_F\,.$$

Proof. $\operatorname{tr}(A^{(s)})$ counts the number of closed non-backtracking walks of length s in the graph. Now, consider the set of closed walks of length $s = qk + r$ such that after every k non-backtracking steps, we can "forget the previous step", i.e. we are allowed to backtrack at step ik for every $i = 0, \ldots, q$. The number of such walks is $\operatorname{tr}((A^{(k)})^q A^{(r)})$. The set of closed non-backtracking walk is clearly a subset of such walks, thus we have

$$\operatorname{tr}(A^{(s)}) \leqslant \operatorname{tr}((A^{(k)})^q A^{(r)}) \leqslant \left\|(A^{(k)})^q\right\|_F \cdot \left\|A^{(r)}\right\|_F\,.$$

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A^{(k)}$ and $\lambda_{\max} = \|A^{(k)}\|_2$. Then,

$$\left\|(A^{(k)})^q\right\|_F = \sqrt{\sum_{i=1}^{n} \lambda_i^{2q}} \leqslant \sqrt{n}(\lambda_{\max})^q\,.$$

This completes the proof. □

With Fact 4.1.1 and Lemma 4.1.2, we now prove Theorem 3.1.3 by analyzing the convergence of $J(t)$ as t increases from 0.

Proof of Theorem 3.1.3. Let $\rho = \lambda_1(B_G)$ be the spectral radius of the non-backtracking matrix B_G of G (recall Definition 2.2.1). We will analyze the convergence of $\operatorname{tr}(J(t)) = \sum_{s=0}^{\infty} \operatorname{tr}(A^{(s)})t^s$ as t increase from 0 to $1/\rho$. In particular, by Lemma 2.2.3 we have that $H_G(t) \succ 0$ (thus $\operatorname{tr}(J(t)) < \infty$) for all $t \in [0, 1/\rho)$, and $\operatorname{tr}(J(1/\rho))$ diverges.

Fix $k \in \mathbb{N}$. For each $s \in \mathbb{N}$ we can write $s = qk + r$, and

$$J(t) = \sum_{s=0}^{\infty} A^{(s)} t^s = \sum_{r=0}^{k-1} \sum_{q=0}^{\infty} A^{(qk+r)} t^{qk+r}.$$

By Lemma 4.1.2, we have

$$\begin{aligned} \operatorname{tr}(J(t)) &\leqslant \sum_{r=0}^{k-1} t^r \sqrt{n} \|A^{(r)}\|_F \sum_{q=0}^{\infty} \|A^{(k)}\|_2^q \cdot t^{qk} \\ &= \sum_{r=0}^{k-1} t^r \sqrt{n} \|A^{(r)}\|_F \sum_{q=0}^{\infty} \left(\|A^{(k)}\|_2 \cdot t^k \right)^q . \end{aligned} \tag{4.1}$$

Now, let $k := \lfloor \log_\rho n \rfloor + 1$ and suppose for contradiction that G contains no cycle of size $\leqslant \ell = 2k$. Observe that every entry of $A^{(k)}$ must be either 0 or 1, otherwise if $A^{(k)}[i,j] > 1$ then there are two distinct length-k paths from i to j, meaning there is a cycle of length at most $2k = \ell$, a contradiction. Therefore, the L_1 norm of each row of $A^{(k)}$ is at most n, hence $\|A^{(k)}\|_2 \leqslant n$. Then, setting $t = 1/\rho$, we have $\|A^{(k)}\|_2 \cdot (1/\rho)^k < 1$ since $k > \log_\rho n$, and Eq. (4.1) shows that $\operatorname{tr}(J(1/\rho)) < \infty$. This contradicts that $\operatorname{tr}(J(1/\rho))$ must diverge.

Similarly, let $k' := \lfloor \log_\rho 2n \rfloor + 1$, and suppose for contradiction that G hs no bicycle of size $\leqslant \ell' = 3k'$. We claim that three distinct non-backtracking walks of a given length-k' between any two vertices must form a bicycle, hence every entry of $A^{(k')}$ must be at most 2. Suppose the union of the three distinct non-backtracking walks between vertices u and v, called H_{uv}, did not give rise to a bicycle, its excess must be at most 0. Since H_{uv} is connected, it must have at most one cycle. If there are no cycles, then there is exactly one non-backtracking walk from u to v, so we assume there is exactly one cycle. Any non-backtracking walk in H_{uv} can enter and exit the cycle at most once. Further, there is a unique way to start from u and enter the cycle, and a unique way to exit the cycle and arrive at v. Between entering and exiting the cycle, there are only two choices: walking in the cycle clockwise or counterclockwise. There are at most two ways to walk between u and v in k' steps — either the shortest path between them is of length exactly k' and does not touch the cycle, or a length-k non-backtracking walk must enter the cycle, which we established gives at most 2 distinct walks.

Thus, $\|A^{(k')}\|_2 \leqslant 2n$ and $\|A^{(k')}\|_2 \cdot (1/\rho)^{k'} < 1$ since $k' > \log_\rho 2n$. Again, Eq. (4.1) shows that $\operatorname{tr}(J(1/\rho)) < \infty$, a contradiction. This completes the proof. □

4.2 Warm-up: weak Moore bound for graphs

Before giving the proof of the hypergraph Moore bound (Theorem 3.1.6), in this section, we first give a simple, alternative proof of a weaker version of the classical Moore

bound. The purpose of this is to illustrate our strategy of "counting *reweighted* walks" for the hypergraph Moore bound, which is a key ingredient in improving the analysis of [GKM22]. Indeed, our proof for the even arity case (in Section 4.3) looks almost identical to the proof in this section.

Proposition 4.2.1 (Weak Moore bound for irregular graphs)**.** *Every graph with n vertices and average degree $d > 16$ has a cycle of length at most $2\lceil \log_{(d/16)} n \rceil$.*

We note that [GKM22] also proved a weak Moore bound (see Proposition 2.3 of [GKM22]) to illustrate their "row bucketing" strategy that partitions the vertices into $O(\log n)$ buckets, each of which has vertices with degrees within a multiplicative constant factor of each other. This strategy splits the adjacency matrix into $O(\log^2 n)$ pieces and ends up requiring an average degree $d \gtrsim \log^3 n$ in order to contain a cycle of length $O(\log n)$.

Our simple proof of Proposition 4.2.1 will show how the reweighting handles different degrees automatically, avoiding the lossy row bucketing step completely.

The core of the proof of Proposition 4.2.1 is the following spectral norm bound on the reweighted adjacency matrix.

Claim 4.2.2. *Let G be a graph with n vertices and average degree $d > 1$ that has no cycle of length $\leqslant \ell$ for some even $\ell \in \mathbb{N}$. Let A be the $\{0,1\}$ adjacency matrix of G, and let $\Gamma = D + d\mathbb{I}$ be the diagonal matrix such that $D_{uu} = d_u$ where d_u is the degree of vertex u. Then, $\left\|\Gamma^{-1/2} A \Gamma^{-1/2}\right\|_2 < \frac{2n^{1/\ell}}{\sqrt{d}}$.*

We now complete the proof of Proposition 4.2.1.

Proof of Proposition 4.2.1 by Claim 4.2.2. Suppose G has no cycle of length $\leqslant \ell$, then Claim 4.2.2 implies that $A \prec \frac{2n^{1/\ell}}{\sqrt{d}}\Gamma$. Then, the quadratic form $\mathbf{1}^\top A \mathbf{1} < \frac{2n^{1/\ell}}{\sqrt{d}} \operatorname{tr}(\Gamma)$ since $\mathbf{1}^\top \Gamma \mathbf{1} = \operatorname{tr}(\Gamma)$. By definition, $\mathbf{1}^\top A \mathbf{1} = nd$ and $\operatorname{tr}(\Gamma) = \sum_{u=1}^{n} (d_u + d) = 2nd$. Thus, $n^{1/\ell} > \sqrt{d}/4$, and taking logs, we get

$$\frac{1}{\ell} \log n > \frac{1}{2} \log(d/16) \Rightarrow \frac{\ell}{2} < \log_{d/16} n\,.$$

ℓ is even, so we have $\ell < 2\lceil \log_{d/16} n \rceil$. Thus, by the contrapositive, G must contain a cycle of length $2\lceil \log_{d/16} n \rceil$. This completes the proof. □

We now prove Claim 4.2.2 using the well-known trace moment method, which reduces to counting weighted closed walks in the graph. In the analysis, we will see exactly how the choice of the reweighting matrix Γ accounts for different vertex degrees.

Proof of Claim 4.2.2. Let $\widetilde{A} = \Gamma^{-1/2} A \Gamma^{-1/2}$. For even $\ell \in \mathbb{N}$, the trace moment method states that $\|\widetilde{A}\|_2^\ell \leqslant \operatorname{tr}(\widetilde{A}^\ell) = \operatorname{tr}((\Gamma^{-1} A)^\ell)$, which is a summation of all (weighted) closed

walks of length ℓ in G. Since there is no cycle of length $\leqslant \ell$, the only closed walks are the ones that *backtrack* to the original vertex, meaning that there can be at most $\ell/2$ "new" edges and at least $\ell/2$ "old" edges in the walk. We encode each closed walk $u_1 \to u_2 \to \cdots \to u_\ell \to u_1$ as follows,

- Choose a starting vertex $u_1 \in [n]$.
- One bit $b_i \in \{0,1\}$ at each step i to encode whether this step uses a new edge or an old one.
  - If $b_i = 0$ (new edge), select one of u_i's neighbors as u_{i+1}.
  - If $b_i = 1$ (old edge), we must backtrack to the previous vertex u_{i-1}.

For $b \in \{0,1\}$ and $u \in [n]$, let $N_b(u) \subseteq [n]$ be the possible next steps in the walk from u. Then, simply expanding $\operatorname{tr}((\Gamma^{-1}A)^\ell)$, we get

$$\operatorname{tr}((\Gamma^{-1}A)^\ell) = \sum_{b\in\{0,1\}^\ell} \sum_{u_1\in[n]} \sum_{u_2\in N_{b_1}(u_1)} \Gamma^{-1}_{u_1u_1} \sum_{u_3\in N_{b_2}(u_2)} \Gamma^{-1}_{u_2u_2} \cdots \sum_{u_{\ell+1}\in N_{b_\ell}(u_\ell)} \Gamma^{-1}_{u_\ell u_\ell} \cdot \mathbf{1}(u_{\ell+1} = u_1)\,.$$

As we can see, each step $u_i \to u_{i+1}$ gets a factor $\Gamma^{-1}_{u_iu_i} = \frac{1}{d_{u_i}+d}$. We can now bound the above by observing that if $b_i = 0$ (new edge), then $|N_0(u_i)| \leqslant d_{u_i}$ and

$$\sum_{u_{i+1}\in N_0(u_i)} \Gamma^{-1}_{u_iu_i} \leqslant \frac{d_{u_i}}{d_{u_i}+d} < 1\,,$$

and if $b_i = 0$ (old edge), then $|N_1(u_i)| = 1$ (the previous step) and

$$\sum_{u_{i+1}\in N_1(u_i)} \Gamma^{-1}_{u_iu_i} \leqslant \frac{1}{d_{u_i}+d} < \frac{1}{d}\,.$$

Finally, considering $b \in \{0,1\}^\ell$, $u_1 \in [n]$, and there are at least $\ell/2$ old edges, we have

$$\operatorname{tr}((\Gamma^{-1}A)^\ell) < 2^\ell n \left(\frac{1}{d}\right)^{\ell/2},$$

and taking the ℓ-th root completes the proof. □

4.3 Hypergraph Moore bound: even arity

In this section, we prove the existence of small even covers in even arity hypergraphs.

Theorem 4.3.1 (Theorem 3.1.6, even k)**.** *For even $k \in \mathbb{N}$ and any $r \in \mathbb{N}$ with $k \leqslant r \leqslant n/8$, any k-uniform hypergraph $\mathcal{H}$ with n vertices and $m \geqslant 128n \log n \cdot (\frac{n}{r})^{k/2-1}$ hyperedges has an even cover of size at most $\lceil r \log_2 n \rceil + 1$.*

The proof is simple and almost identical to the proof of the weak Moore bound (Proposition 4.2.1), but with A being the adjacency matrix of the *Kikuchi graph*.

Definition 4.3.2 (Kikuchi graph)**.** Let $\mathcal{H}$ be a k-uniform hypergraph on vertex set $[n]$ for even k. For an integer parameter r, define the *Kikuchi graph* K_r associated to $\mathcal{H}$ is a graph on vertex set $\binom{[n]}{r}$ such that a pair of vertices $S, T \in \binom{[n]}{r}$ have an edge between them if the symmetric difference $S \oplus T \in \mathcal{H}$. For such an edge, we write $S \overset{C}{\leftrightarrow} T$ and think of the edge as "colored" by $C \in \mathcal{H}$ where $C = S \oplus T$. We call the adjacency matrix A of K_r the *Kikuchi matrix*.

The key insight of [GKM22] (and also our starting point) is relating even covers in $\mathcal{H}$ to cycles in the Kikuchi graph. For sets $R_1, R_2, \ldots, R_\ell \subseteq [n]$, let $\oplus_{i \leqslant \ell} R_i$ denote the set of elements of $[n]$ that appear in an odd number of R_is (i.e., the sum modulo 2 of the indicator vectors of R_is).

Observation 4.3.3 (Closed walks in the Kikuchi graph)**.** Let $\mathcal{H}$ be a k-uniform hypergraph on $[n]$ for even k and let $S_1 \to S_2 \to \cdots S_\ell \to S_1$ be a closed walk on vertices in K_r such that for every $i \leqslant \ell$, $S_i \overset{C_i}{\leftrightarrow} S_{i+1}$ for $C_1, C_2, \ldots, C_\ell \in \mathcal{H}$ (denoting $S_{\ell+1} = S_1$). Then, $\oplus_{i \leqslant \ell} C_i = 0$. Further, if $\mathcal{H}$ has no even cover of length ℓ, then every hyperedge in $\mathcal{H}$ appears an even number of times in the multiset $\{C_1, C_2, \ldots, C_\ell\}$. We will call such walks in K_r *trivial*.

Proof. Note that $S_i \oplus S_{i+1} = C_i$ for every $i \leqslant \ell$. If we add both sides of all ℓ such equalities then each S_i occurs in exactly two of the equations so the LHS must be 0. Thus, $\oplus_{i \leqslant \ell} C_i = 0$.

Next, we repeatedly remove hyperedges that occur an even number of times in the multiset $\{C_1, C_2, \ldots, C_\ell\}$ to obtain a collection of $\ell' \leqslant \ell$ distinct hyperedges of $\mathcal{H}$. The sum (modulo 2) of the remaining hyperedge should still be 0 as we removed hyperedges in pairs. The resulting ℓ' must be 0 as otherwise the remaining hyperedges form an even cover of length $\ell' \leqslant \ell$. □

Consider a hypergraph $\mathcal{H}$ with n vertices and m hyperedges, and its associated Kikuchi graph (V, E) with parameter r. Each $C \in \mathcal{H}$ introduces $\frac{1}{2}\binom{k}{k/2}\binom{n-k}{r-k/2}$ edges in the Kikuchi graph (select $k/2$ vertices from C and select $r - k/2$ vertices from $[n] \setminus C$ to complete S), thus the total edges $|E| = \frac{1}{2}\binom{k}{k/2}\binom{n-k}{r-k/2} \cdot m$. Let d_S be the degree of $S \in V$, and let d denote the average degree. A straightforward calculation shows that

$$\begin{aligned} d = \frac{\binom{k}{k/2}\binom{n-k}{r-k/2} m}{\binom{n}{r}} &\geqslant \left(\frac{r}{n}\right)^{k/2} m \cdot \binom{k}{k/2} \left(1 - \frac{2r}{n}\right)^{k/2} \left(1 - \frac{k}{2r}\right)^{k/2} \\ &\geqslant \frac{1}{2}\left(\frac{r}{n}\right)^{k/2} m\,, \end{aligned} \tag{4.2}$$

when $k \leqslant r \leqslant n/8$.

We will follow the reweighting strategy with $\Gamma = D + d\mathbb{I}$ to bound the spectral norm of the reweighted Kikuchi matrix. The following lemma is analogous to Claim 4.2.2.

Lemma 4.3.4. *Let $k, r, n \in \mathbb{N}$ such that $k \leqslant r \leqslant n$, and let $\ell \in \mathbb{N}$ be even. Let A be the Kikuchi matrix with parameter r of a k-uniform hypergraph $\mathcal{H}$ on n vertices, and let $\Gamma = D + d\mathbb{I}$ where D is the degree matrix and d is the average degree of the Kikuchi graph. Suppose there is no even cover of size at most ℓ in $\mathcal{H}$, then*

$$\left\|\Gamma^{-1/2} A \Gamma^{-1/2}\right\|_2 < 2n^{r/\ell}\sqrt{\frac{\ell}{d}}.$$

We can immediately complete the proof of Theorem 4.3.1.

Proof of Theorem 4.3.1 by Lemma 4.3.4. Suppose that there is no even cover of size $\leqslant \ell := \lceil r \log_2 n \rceil$ (assume this is even, otherwise add 1). Then, $n^{r/\ell} \leqslant 2$ and Lemma 4.3.4 states that the Kikuchi graph (V, E) satisfies $A \prec 4\sqrt{\ell/d} \cdot \Gamma$ where $\Gamma = D + d\mathbb{I}$. Then,

$$\mathbf{1}^\top A \mathbf{1} < 4\sqrt{\frac{\ell}{d}} \cdot \operatorname{tr}(\Gamma) = 4\sqrt{\frac{\ell}{d}} \cdot \sum_{S \in V} (d_S + d) = 8\sqrt{\frac{\ell}{d}} \cdot |V| d\,.$$

On the other hand, $\mathbf{1}^\top A \mathbf{1} = 2|E| = |V|d$. Thus, we have $d < 64\ell$. By Eq. (4.2) we have $d \geqslant \frac{1}{2}(\frac{r}{n})^{k/2} m$ when $k \leqslant r \leqslant n/8$. Thus, if there is no even cover of size $\leqslant \ell$, then $m < 128 n \log n \cdot (\frac{n}{r})^{k/2-1}$, completing the proof. □

Now, we prove Lemma 4.3.4 by counting weighted closed walks in the Kikuchi graph, essentially the same way we prove Claim 4.2.2.

Proof of Lemma 4.3.4. Let $\widetilde{A} = \Gamma^{-1/2} A \Gamma^{-1/2}$. We use the trace power method:

$$\|\widetilde{A}\|_2^\ell \leqslant \operatorname{tr}(\widetilde{A}^\ell) = \operatorname{tr}((\Gamma^{-1} A)^\ell)\,.$$

We upper bound $\operatorname{tr}((\Gamma^{-1}A)^\ell)$ by counting (weighted) closed walks of length ℓ in the Kikuchi graph. Note that each edge (S, T) of the Kikuchi graph corresponds to a hyperedge $S \oplus T \in \mathcal{H}$. Since there is no even covers of size at most ℓ, any closed walk must contain an even number of each hyperedge in $\mathcal{H}$.

We can encode a closed walk $S_1 \to S_2 \to \cdots \to S_\ell \to S_1$ as follows:

- Choose a starting vertex $S_1 \in V$.
- One bit $b_i \in \{0, 1\}$ at each step i to encode whether this step uses a new hyperedge or an old one.
  - If $b_i = 0$ (new hyperedge), select one of S_i's neighbors as S_{i+1}.
  - If $b_i = 1$ (old hyperedge), select an old hyperedge C from the previous steps, and set $S_{i+1} = S_i \oplus C$.

Note that there are at most $\ell/2$ new hyperedges and at least $\ell/2$ old hyperedges since each hyperedge must occur an even number of times. For $b \in \{0,1\}$ and $S \in V$, let $N_b(S) \subseteq V$ be the possible next steps in the walk from S (according to b). Each step $S_i \to S_{i+1}$ gets a factor $(\Gamma^{-1}A)_{S_i,S_{i+1}} = \Gamma^{-1}_{S_i,S_i} = \frac{1}{d_{S_i}+d}$. Thus,

$$\operatorname{tr}((\Gamma^{-1}A)^\ell) = \sum_{b\in\{0,1\}^\ell} \sum_{S_1\in V} \sum_{S_2\in N_{b_1}(S_1)} \frac{1}{d_{S_1}+d} \sum_{S_3\in N_{b_2}(S_2)} \frac{1}{d_{S_2}+d} \cdots \sum_{S_{\ell+1}\in N_{b_\ell}(S_\ell)} \frac{\mathbf{1}(S_{\ell+1}=S_1)}{d_{S_\ell}+d}.$$

We can upper bound the above as follows. If $b = 0$, then $|N_0(S_i)| \leqslant d_{S_i}$ and $\sum_{S_{i+1}\in N_0(S_i)} \Gamma^{-1}_{S_iS_i} \leqslant \frac{d_{S_i}}{d_{S_i}+d} < 1$. If $b = 1$, then $|N_1(S_i)| \leqslant \ell$ as there are only ℓ options to choose one of the previous steps, and $\sum_{S_{i+1}\in N_1(S_i)} \Gamma^{-1}_{S_iS_i} \leqslant \frac{\ell}{d_{S_i}+d} < \frac{\ell}{d}$. Furthermore, we can assume that $\ell \leqslant d$, otherwise we can simply treat all steps as new hyperedges.

Finally, $b \in \{0,1\}^\ell$, there are $|V| = \binom{n}{r}$ choices for the starting vertex S_1, and there are at least $\ell/2$ old hyperedges. Thus, we have

$$\operatorname{tr}((\Gamma^{-1}A)^\ell) < 2^\ell \binom{n}{r} \left(\frac{\ell}{d}\right)^{\ell/2} \leqslant 2^\ell n^r \left(\frac{\ell}{d}\right)^{\ell/2}.$$

Taking the ℓ-th root completes the proof. □

4.4 Hypergraph Moore bound: odd arity

In this section we prove the hypergraph Moore bound for k-uniform hypergraphs when k is odd.

Theorem 4.4.1 (Theorem 3.1.6, odd k)**.** *There is a universal constant B such that for any odd $k \in \mathbb{N}$, and any $r \in \mathbb{N}$ satisfying $2k \leqslant r \leqslant \frac{n}{B^k}$, any k-uniform hypergraph $\mathcal{H}$ with n vertices and $m \geqslant B^k n \log n \cdot \left(\frac{n}{r}\right)^{k/2-1}$ hyperedges has an even cover of size at most $r \log_2 n$.*

Our proof strategy broadly involves the following steps.

- **Hypergraph decomposition.** We partition $\mathcal{H}$ into subhypergraphs $\mathcal{H}^{(0)}, \mathcal{H}^{(1)}, \ldots, \mathcal{H}^{(k-1)}$ with the property that every size-$(i+1)$ set in $\mathcal{H}^{(i)}$ is contained in only a small number of clauses, and every clause in $\mathcal{H}^{(i)}$ intersects many other clauses at a size-i set. One of the $\mathcal{H}^{(i)}$ must contain at least m/k clauses, and we find an even cover in that $\mathcal{H}^{(i)}$.
- **Large i.** When $i \geqslant \frac{k+1}{2}$, we give a direct reduction to the hypergraph Moore bound for even arity hypergraphs and apply Theorem 4.3.1.
- **Kikuchi graph.** To handle the remaining values of i, we show the existence of an even cover by proving the contrapositive — a hypergraph with no small even

covers has a bounded number of hyperedges. To achieve this, we appropriately define the Kikuchi graph for odd arity hypergraphs, and show that the adjacency matrix $\widehat{A}$ of some suitably chosen subgraph (via the "edge deletion process" described below) satisfies $\widehat{A} \preceq Q$ for some diagonal matrix Q. Then the resulting inequality $\mathbf{1}^\top \widehat{A} \mathbf{1} \leqslant \mathrm{tr}(Q)$ can be rearranged to bound the number of hyperedges.

- **Trace method.** The way we prove $\widehat{A} \preceq Q$ is by using the trace moment method to show $\left\| Q^{-1/2} \widehat{A} Q^{-1/2} \right\|_2 \leqslant 1$. Bounding a high trace power of $Q^{-1/2} \widehat{A} Q^{-1/2}$ corresponds to bounding the total weight of closed walks that use every hyperedge an even number of times in the Kikuchi graph.
- **Edge deletion process.** We delete a small fraction of the edges in K_r with the guarantee that in the resulting subgraph any clause participates in only a small number of incident edges to every vertex.

Hypergraph decomposition. We describe our algorithm to partition our hypergraph.

Algorithm 4.4.2. We partition $\mathcal{H}$ into hypergraphs $\mathcal{H}^{(0)}, \ldots, \mathcal{H}^{(k-1)}$ via the following algorithm.

1. Set $t = k - 1$ and $\mathcal{H}_{\text{current}} := \mathcal{H}$.
2. Set counter $s = 1$. While there is some subset $U \subseteq [n]$ such that $|U| = t$ and $|\{C \in \mathcal{H}_{\text{current}} : U \subseteq C\}| \geqslant \max\left\{2, \left(\frac{n}{r}\right)^{\frac{k}{2}-t}\right\}$:
   (a) Choose U satisfying the condition and let $\mathcal{H}_s^{(t)}$ be a subset of $\{C \in \mathcal{H}_{\text{current}} : U \subseteq C\}$ of size $\max\left\{2, \left(\frac{n}{r}\right)^{\frac{k}{2}-t}\right\}$.
   (b) Add all clauses in $\mathcal{H}_s^{(t)}$ to $\mathcal{H}^{(t)}$.
   (c) Delete all clauses in $\mathcal{H}_s^{(t)}$ to $\mathcal{H}_{\text{current}}$.
   (d) Increment s by 1.
3. Decrement t by 1. If $t > 0$, go back to step 2; otherwise take the remaining clauses in $\mathcal{H}_{\text{current}}$ and add them to $\mathcal{H}^{(0)}$.

First, observe that the largest subhypergraph $\mathcal{H}^{(i)}$ in the partition produced by our algorithm must have at least $\frac{m}{k}$ hyperedges. Next, observe that $i \neq 0$ because if $|\mathcal{H}^{(0)}| \geqslant m/k$, then there must be a $j \in [n]$ such that $\left|\{C \in \mathcal{H}^{(0)} : j \in C\}\right| \geqslant \frac{m}{nk} \gg (\frac{n}{r})^{k/2-1}$, which would have been added to $\mathcal{H}^{(1)}$. Our goal in the rest of the proof is to find a small even cover in $\mathcal{H}^{(i)}$. The following observations articulate the properties of $\mathcal{H}^{(i)}$ we need that are guaranteed by the algorithm.

Observation 4.4.3. $\mathcal{H}^{(i)}$ can be partitioned into $\mathcal{H}_1^{(i)}, \dots, \mathcal{H}_p^{(i)}$ where for each $j \in [p]$, there is a set U_j of size i such that every $C \in \mathcal{H}_j^{(i)}$ contains U_j, and $|\mathcal{H}_j^{(i)}| \geqslant \left(\frac{n}{r}\right)^{\frac{k}{2}-i}$ and $p \leqslant m \cdot \left(\frac{r}{n}\right)^{\frac{k}{2}-i}$.

Observation 4.4.4. For $s \geqslant 1$ and any $U \subseteq [n]$ such that $|U| = i + s$, the number of hyperedges in $\mathcal{H}^{(i)}$ containing U is at most $\max\left\{1, \left(\frac{n}{r}\right)^{\frac{k}{2}-s-i}\right\}$, otherwise they would have been added to $\mathcal{H}^{(i+s)}$.

Reduction to even arity case when $i \geqslant \frac{k+1}{2}$. In this case, by Observation 4.4.4, each pair $C \neq C'$ in any $\mathcal{H}_j^{(i)}$ must satisfy $C \cap C' = U_j$. The following makes the reduction from finding even covers in $\mathcal{H}^{(i)}$ if $i \geqslant \frac{k+1}{2}$ to the even arity case concrete.

Lemma 4.4.5. *Let $\mathcal{H}$ be a k-uniform hypergraph on n vertices with no even cover of size $r \log_2 n$. Fix $1 \leqslant i \leqslant k-1$. Suppose $\mathcal{H}_1, \dots, \mathcal{H}_p$ are disjoint subsets of $\mathcal{H}$ such that for each $j \in [p]$, $|\mathcal{H}_j| \geqslant 2$ and all pairs of hyperedges $C \neq C' \in \mathcal{H}_j$ satisfy $C \cap C' = U_j$ for some $U_j \subseteq [n]$ of size i. Then,*

$$\sum_{j=1}^{p} |\mathcal{H}_j| \leqslant O(n \log n) \left(\frac{2n}{r}\right)^{k-i-1}.$$

In particular, when $i \geqslant \frac{k+1}{2}$ the above is at most $O(n \log n) \cdot \left(\frac{n}{r}\right)^{k/2-1}$.

Proof. Given such disjoint subsets $\mathcal{H}_1, \dots, \mathcal{H}_p$, we can construct a $2(k-i)$-uniform hypergraph $\widehat{\mathcal{H}}$ by the following: for each $j \in [p]$, arbitrarily order the edges: $\mathcal{H}_j = (C_1, \dots, C_{|\mathcal{H}_j|})$. Then, add the hyperedge $C_s \oplus C_{s+1}$ to $\widehat{\mathcal{H}}$ for $s = 1, \dots, |\mathcal{H}_j| - 1$. By assumption $|C_s \cap C_{s+1}| = |U_j| = i$, thus $|C_s \oplus C_{s+1}| = 2(k-i)$. The resulting $\widehat{\mathcal{H}}$ has

$$|\widehat{\mathcal{H}}| = \sum_{j=1}^{p} |\mathcal{H}_j| - 1 \geqslant \frac{1}{2} \sum_{j=1}^{p} |\mathcal{H}_j|$$

hyperedges, since $|\mathcal{H}_j| \geqslant 2$ for all $j \in [p]$.

We claim that $\widehat{\mathcal{H}}$ cannot have an even cover of size at most $\frac{r}{2} \log_2 n$. First, if $\widehat{\mathcal{H}}$ has repeated hyperedges, then there must exist $j \neq j' \in [p]$ and $C_1, C_2 \in \mathcal{H}_j$, $C_1', C_2' \in \mathcal{H}_{j'}$ such that $C_1 \oplus C_2 = C_1' \oplus C_2'$, but then $\{C_1, C_2, C_1', C_2'\}$ would be an even cover of size 4 in $\mathcal{H}$. Now, suppose $\widehat{\mathcal{H}}$ has no repeated edges but has an even cover of size ℓ. Then, for any $\widehat{C}$ in the even cover, we can uniquely identify $j \in [p]$ and $s \leqslant |\mathcal{H}_j| - 1$ such that $C_s, C_{s+1} \in \mathcal{H}_j$ and $\widehat{C} = C_s \oplus C_{s+1}$. Furthermore, by construction there must be at least two $C_s, C_{s'} \in \mathcal{H}_j$ that each occurs only once. Therefore, these edges must form an even cover of size at most 2ℓ in $\mathcal{H}$.

Since $2(k-i)$ is even and $\widehat{\mathcal{H}}$ has no even cover of size $\frac{r}{2}\log_2 n$, we can apply Theorem 4.3.1 to show that

$$|\widehat{\mathcal{H}}| \leqslant O(n \log n)\left(\frac{2n}{r}\right)^{k-i-1}.$$

This completes the proof. □

Henceforth, we assume $i \leqslant \frac{k-1}{2}$, which is the case we need an appropriate Kikuchi graph for odd arity hypergraphs.

Kikuchi matrix for odd arity hypergraphs. The following is the same Kikuchi graph defined in [GKM22, Definition 6.2].

Definition 4.4.6 (Colored Kikuchi graphs and subgraphs)**.** Fix $r \in \mathbb{N}$ and $t \in \{1, \dots, k-1\}$ such that $2k \leqslant r \leqslant n$. Let $\mathcal{H}_1, \dots, \mathcal{H}_p$ be p disjoint sets of hyperedges such that for each $i \in [p]$, all hyperedges in $\mathcal{H}_i$ have a common subset $U_i \subset [n]$ where $|U_i| = t$. For each $C \in \mathcal{H}_i$, denote $\widetilde{C} := C \setminus U_i$, and denote $\widetilde{\mathcal{H}}_i := \{\widetilde{C} : C \in \mathcal{H}_i\}$ which can be viewed as a $(k-t)$-uniform hypergraph. We define the *colored* Kikuchi graph K_r as follows.

The vertex set $V(K_r)$ consists of subsets of $[n] \times [2]$ of size r, where $S \in V$ is viewed as $(S^{(1)}, S^{(2)})$ where $S^{(1)}, S^{(2)} \subseteq [n]$ are colored *green* and *blue* respectively. For each $i \in [p]$ and each $C \neq C' \in \mathcal{H}_i$, let $\widetilde{C}^{(1)}$ be $\widetilde{C}$ colored green and $\widetilde{C}'^{(2)}$ be $\widetilde{C}'$ colored blue, and we add an edge between $S, T \in V$, denoted $S \overset{C,C'}{\longleftrightarrow} T$, if $S \oplus T = \widetilde{C}^{(1)} \oplus \widetilde{C}'^{(2)}$ and if one of the following holds,

- $|\widetilde{C} \cap S^{(1)}| = |\widetilde{C}' \cap T^{(2)}| = \left\lceil \frac{k-t}{2} \right\rceil$ and $|\widetilde{C}' \cap S^{(2)}| = |\widetilde{C} \cap T^{(1)}| = \left\lfloor \frac{k-t}{2} \right\rfloor$, or
- $|\widetilde{C} \cap S^{(1)}| = |\widetilde{C}' \cap T^{(2)}| = \left\lfloor \frac{k-t}{2} \right\rfloor$ and $|\widetilde{C}' \cap S^{(2)}| = |\widetilde{C} \cap T^{(1)}| = \left\lceil \frac{k-t}{2} \right\rceil$, or

Figure 4.1 shows an example of two edges $C, C' \in \mathcal{H}_i$ forming an edge (S, T) in the Kikuchi graph.

We say that the edge (S, T) is type-i, and for $S \in V$, we define the type-i degree as

$$d_{S,i} := \max_{C \in \mathcal{H}_i} \left| \left\{ T \in V : S \overset{C,C'}{\longleftrightarrow} T \text{ for some } C' \in \mathcal{H}_i \right\} \right|.$$

We call any subgraph of the colored Kikuchi graph as a *colored Kikuchi subgraph*.

Remark 4.4.7 (Purpose of coloring)**.** The coloring in Definition 4.4.6 is needed because $C \neq C' \in \mathcal{H}_i$ may have intersection larger than t, meaning $|C \oplus C'| = |\widetilde{C} \oplus \widetilde{C}'| < 2(k-t)$, making the analysis complicated. Coloring $\widetilde{C}, \widetilde{C}'$ with different colors automatically makes $\widetilde{C}^{(1)}, \widetilde{C}'^{(2)}$ disjoint, i.e. $|S \oplus T| = |\widetilde{C}^{(1)} \oplus \widetilde{C}'^{(2)}| = 2(k-t)$. Note also that a vertex $S \subseteq [n] \times [2]$ may contain two copies of some element in $[n]$ with different colors, as shown in Figure 4.1.

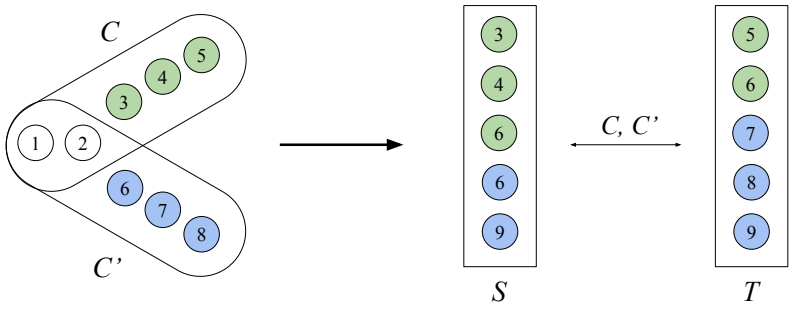


Figure 4.1: An example of a colored Kikuchi graph (Definition 4.4.6) with $k = 5$ and $t = 2$. On the left are two 5-uniform hyperedges in $\mathcal{H}_i$ with common intersection $U_i = \{1,2\}$ and $\widetilde{C} = \{3,4,5\}$, $\widetilde{C}' = \{6,7,8\}$. On the right, S and T are vertices in the Kikuchi graph where $S^{(1)} = \{3,4,6\}$, $T^{(1)} = \{5,6\}$ are colored green, and $S^{(2)} = \{6,9\}$, $T^{(2)} = \{7,8,9\}$ are colored blue. C and C' form an edge between S, T because $|\widetilde{C} \cap S^{(1)}| = 2$, $|\widetilde{C} \cap T^{(1)}| = 1$, $|\widetilde{C}' \cap S^{(2)}| = 1$, and $|\widetilde{C}' \cap T^{(2)}| = 2$.

Observation 4.4.8 (Parameters of the Kikuchi graph)**.** The Kikuchi graph (V, E) defined in Definition 4.4.6 has $|V| = \binom{2n}{r}$, and each distinct pair $C, C' \in \mathcal{H}_i$ contributes a collection of edges $E_{C,C'}$ in E, where

$$|E_{C,C'}| = \alpha_t := \binom{k-t}{\lfloor \frac{k-t}{2} \rfloor} \binom{k-t}{\lceil \frac{k-t}{2} \rceil} \binom{2n-2(k-t)}{r-(k-t)} \cdot 2^{\mathbf{1}(k-t \text{ is odd})}$$

by first choosing $\widetilde{C} \cap S^{(1)}$, $\widetilde{C}' \cap S^{(2)}$ (or $\widetilde{C} \cap S^{(2)}$, $\widetilde{C}' \cap S^{(1)}$) and completing S's remaining $r - (k-t)$ elements. Thus, $|E| = \sum_{i=1}^{p} \binom{|\mathcal{H}_i|}{2} \cdot \alpha_t$, and standard calculations show that when $2k \leqslant r \leqslant n/8$, the average degree $d = \frac{2|E|}{|V|}$ satisfies

$$\left(\frac{r}{2n}\right)^{k-t} \sum_{i=1}^{p} \binom{|\mathcal{H}_i|}{2} \leqslant d \leqslant 2^{2k} \left(\frac{r}{2n}\right)^{k-t} \sum_{i=1}^{p} \binom{|\mathcal{H}_i|}{2}.$$

Our ideal hope is that the adjacency matrix A of the Kikuchi graph, constructed from $\mathcal{H}^{(i)} = (\mathcal{H}_1^{(i)}, \ldots, \mathcal{H}_p^{(i)})$, is bounded in the PSD order by some low-trace diagonal matrix Q. To achieve this, we prove the following lemma analogous to Lemma 4.3.4, but with the additional requirement that $d_{S,i}$ is small for all $S \in V(K_r)$ and $i \in [p]$. The proof is almost identical to the proof of Lemma 4.3.4 but the encoding for an "old hyperedge" step is different.

Lemma 4.4.9. *Let $r \geqslant 2k$. Given disjoint hyperedges $\mathcal{H}_1, \ldots, \mathcal{H}_p$, let $\widehat{A}$ be the adjacency matrix of any colored Kikuchi subgraph $\widehat{K}_r$ as defined in Definition 4.4.6, and let $\Gamma = D + d\mathbb{I}$ where D is the degree matrix and d is the average degree of G. Fix $\eta \in \mathbb{R}$ and let $\ell \in \mathbb{N}$ be even. Suppose there is no even cover of size at most ℓ, and suppose $d_{S,i} \leqslant \eta$ for all $S \in V$ and $i \in [p]$. Then,*

$$\left\| \Gamma^{-1/2} \widehat{A} \Gamma^{-1/2} \right\|_2 \leqslant 2n^{r/\ell} \sqrt{\frac{2\eta\ell}{d}}.$$

Proof. Let $\widetilde{A} = \Gamma^{-1/2}\widehat{A}\Gamma^{-1/2}$. We again use the trace power method:

$$\left\|\widetilde{A}\right\|_2^\ell \leqslant \operatorname{tr}(\widetilde{A}^\ell) = \operatorname{tr}((\Gamma^{-1}A)^\ell)\,.$$

Note that each edge (S,T) in $\widehat{A}$ corresponds to two hyperedges of the same type (both from some $\mathcal{H}_i$), one green and one blue, and since there is no even covers of size at most ℓ, any closed walk must contain an even number of each hyperedge.

We encode a closed walk $S_1 \to S_2 \to \cdots \to S_\ell \to S_1$ as follows:

- Starting vertex $S_1 \in V$.
- One bit $b_i \in \{0,1\}$ at step i to encode whether this step uses two new hyperedges or one (or more) old hyperedge.
  - If $b_i = 0$ (two new hyperedges), select one of S_i's neighbors as S_{i+1}.
  - If $b_i = 1$ (old hyperedge), select an old green (or blue) hyperedge C from the previous steps, and select a blue (or green) hyperedge C' incident to S_i.

Recall that for $b \in \{0,1\}$, we write $N_b(S)$ as the possible next steps in the walk from S. Using the same analysis as the proof of Lemma 4.3.4, for $b = 0$,

$$\sum_{S_{i+1}\in N_0(S_i)} \frac{1}{d_{S_i}+d} \leqslant 1\,,$$

and for $b = 1$, suppose the old edge is of type $j \in [p]$, then $|N_1(S_i)| \leqslant 2\ell d_{S_i,j}$ (one previous step, 2 colors), thus

$$\sum_{S_{i+1}\in N_b(S_i)} \frac{1}{d_{S_i}+d} \leqslant \frac{2\ell d_{S_i,j}}{d_{S_i}+d} \leqslant \frac{2\eta\ell}{d}\,.$$

We can assume that $2\eta\ell \leqslant d$, otherwise we can simply treat all steps as new hyperedges.

There are $\binom{2n}{r} \leqslant (\frac{2en}{r})^r \leqslant n^r$ (since $r \geqslant 2k$ and $k \geqslant 3$) choices to pick the starting vertex S_1. Furthermore, there can be at most $\ell/2$ steps that use two new hyperedges, i.e. $|b| \geqslant \ell/2$, thus

$$\operatorname{tr}((\Gamma^{-1}\widehat{A})^\ell) \leqslant n^r \sum_{b\in\{0,1\}^\ell} \left(\frac{2\eta\ell}{d}\right)^{|b|} \leqslant 2^\ell n^r \left(\frac{2\eta\ell}{d}\right)^{\ell/2}\,.$$

Taking the ℓ-th root completes the proof. □

Construction of colored Kikuchi subgraph. Unfortunately, the requirement for all $d_{S,i}$ to be bounded by a small η prohibits us from obtaining a good bound on the adjacency matrix of the full colored Kikuchi graph K_r using Lemma 4.4.9. This motivates dropping a small number of edges from K_r, and bounding the adjacency matrix $\widehat{A}$ of the resulting subgraph $\widehat{K}_r$ instead. Thus, we proceed with identifying a suitable colored Kikuchi subgraph $\widehat{K}_r$ of $\mathcal{H}^{(i)}$ with adjacency matrix $\widehat{A}$ via the following *edge deletion process*:

Start with the colored Kikuchi graph K_r, and delete every edge $\{S, T\}$ caused by a pair of clauses C, C' such that S or T has strictly more than 1 edge that C or C' participates in.

To obtain a handle on the average degree of $\widehat{K}_r$, we first show that the number of edges of K_r we delete to obtain $\widehat{K}_r$ is only a small fraction of the total number of edges, and then the desired lower bound follows from a lower bound on $|E(K_r)|$.

Analyzing the edge deletion process. We find it convenient to think of the fraction of deleted edges as the *probability that a uniformly random edge in K_r is absent in $\widehat{K}_r$*. With this probabilistic interpretation in hand, observe that a uniformly random edge in K_r is the same as choosing a uniformly random pair of clauses (C, C') such that C and C' both belong to the same $\mathcal{H}_j^{(i)}$ and then choosing a random edge $\{S, T\}$ in $E_{C,C'}$, the collection of edges adorned by (C, C'). We will use the notation $C'' \to_C S$ to mean $|\widetilde{C}'' \cap S| = |\widetilde{C} \cap S|$, where we recall from Definition 4.4.6 that $\widetilde{C} := C \setminus U_j$ with U_j being the size-i common intersection of $\mathcal{H}_j^{(i)}$. We then show the following.

Claim 4.4.10 (Deletion probability)**.** *For every pair of clauses (C, C') such that C and C' belong to the same $\mathcal{H}_j^{(i)}$ for some $j \in [p]$,*

$$\Pr_{\{S,T\} \sim E_{C,C'}}[\{S, T\} \text{ deleted}] \leqslant k \cdot 4^{k+1} \sqrt{\frac{r}{n}}.$$

Proof. Recall that we defined $\widetilde{C} = C \setminus U_j$ and $\widetilde{C}' = C' \setminus U_j$. The distribution of $S = (S^{(1)}, S^{(2)})$ (the green and blue vertices) is uniform on all sets such that:

- $|\widetilde{C} \cap S^{(1)}| = \left\lceil \frac{k-i}{2} \right\rceil$, $|\widetilde{C}' \cap S^{(2)}| = \left\lfloor \frac{k-i}{2} \right\rfloor$, or
- $|\widetilde{C} \cap S^{(1)}| = \left\lceil \frac{k-i}{2} \right\rceil$, $|\widetilde{C}' \cap S^{(2)}| = \left\lfloor \frac{k-i}{2} \right\rfloor$.

Then, by union bound,

$$\begin{aligned}
\Pr_{\{S,T\} \sim E_{C,C'}}[\{S, T\} \text{ deleted}] \leqslant{}& \Pr_{\{S,T\} \sim E_{C,C'}}\left[\exists C'' \to_C S^{(1)} : C'' \in \mathcal{H}_j^{(i)}, C'' \neq C\right] \\
&+ \Pr_{\{S,T\} \sim E_{C,C'}}\left[\exists C'' \to_{C'} S^{(2)} : C'' \in \mathcal{H}_j^{(i)}, C'' \neq C'\right] \\
&+ \Pr_{\{S,T\} \sim E_{C,C'}}\left[\exists C'' \to_C T^{(1)} : C'' \in \mathcal{H}_j^{(i)}, C'' \neq C\right] \\
&+ \Pr_{\{S,T\} \sim E_{C,C'}}\left[\exists C'' \to_{C'} T^{(2)} : C'' \in \mathcal{H}_j^{(i)}, C'' \neq C'\right] \\
={}& 4\Pr_{\{S,T\} \sim E_{C,C'}}\left[\exists C'' \to_C S^{(1)} : C'' \in \mathcal{H}_j^{(i)}, C'' \neq C\right]
\end{aligned}$$

then by Markov's inequality,

$$\leqslant 4 \mathop{\mathbb{E}}_{\{S,T\}\sim E_{C,C'}} \left| C'' : C'' \to_C S^{(1)}, C'' \in \mathcal{H}_j^{(i)}, C'' \neq C \right|$$
$$= 4 \sum_{\substack{C'':C''\in\mathcal{H}_j^{(i)} \\ C''\neq C}} \Pr_{\{S,T\}\sim E_{C,C'}} \left[C'' \to_C S^{(1)} \right]. \tag{4.3}$$

Once the intersection of S with $\widetilde{C}$ and $\widetilde{C}'$ is chosen, the remaining elements are selected uniformly at random without replacement. For fixed $C'' \neq C \in \mathcal{H}_j^{(i)}$, since they contain U_j of size i, $|\widetilde{C}'' \cap \widetilde{C}| = |C'' \cap C| - i$, and S must include $\lfloor \frac{k-i}{2} \rfloor - (|C'' \cap C| - i)$ additional elements from $\widetilde{C}'' \setminus \widetilde{C}$ for $C'' \to_C S^{(1)}$ to hold. Thus,

$$\Pr_{\{S,T\}\sim E_{C,C'}} \left[C'' \to_C S^{(1)} \right] \leqslant 2^k \left(\frac{r}{n}\right)^{\lfloor \frac{k-i}{2} \rfloor - |C''\cap C| + i}.$$

Thus, we can prove:

$$\text{Eq. (4.3)} \leqslant 4 \cdot 2^k \sum_{s=i}^{k-1} \sum_{\substack{U\subseteq C \\ |U|=s}} \sum_{\substack{C'':C''\in\mathcal{H}_j^{(i)} \\ C''\neq C \\ C''\cap C = U}} \left(\frac{r}{n}\right)^{\lfloor \frac{k-i}{2} \rfloor - s + i}. \tag{4.4}$$

By Observation 4.4.4, we can bound the above as

$$\leqslant 4 \cdot 2^k \sum_{s=i}^{k-1} \sum_{\substack{U\subseteq C \\ |U|=s}} \left(\frac{n}{r}\right)^{\frac{k}{2}-s} \left(\frac{r}{n}\right)^{\frac{k-i}{2} - \frac{\mathbf{1}[k-i \text{ odd}]}{2} - s + i}$$
$$\leqslant k \cdot 4^{k+1} \sqrt{\frac{r}{n}},$$

as $\frac{i}{2} - \frac{\mathbf{1}[k-i \text{ odd}]}{2} \geqslant \frac{1}{2}$ for all $i \geqslant 1$ when k is odd. □

Lower bound on average degree in A. By choosing B large enough, the upper bound on r, and Claim 4.4.10, the fraction of edges we delete from the original colored Kikuchi graph K_r to obtain $\widehat{K}_r$ is at most .5 and hence $d(\widehat{K}_r) \geqslant .5 d(K_r)$ where $d(K_r)$ and $d(\widehat{K}_r)$ are the average degrees in K_r and $\widehat{K}_r$ respectively. Thus, we know:

$$d(K_r) \geqslant \left(\frac{r}{2n}\right)^{k-i} \sum_{j=1}^{p} \binom{|\mathcal{H}_j^{(i)}|}{2} \geqslant \left(\frac{r}{2n}\right)^{k-i} \cdot p \cdot \binom{m/kp}{2} \geqslant \left(\frac{r}{2n}\right)^{k-i} \cdot \frac{m^2}{4k^2 p},$$

where the first inequality uses Observation 4.4.8, and the second inequality is due to Jensen's inequality.

By the upper bound $p \leqslant m \cdot \left(\frac{r}{n}\right)^{\frac{k}{2}-i}$ as noted in Observation 4.4.3:

$$d(K_r) \geqslant \frac{1}{4k^2 2^k} \cdot \left(\frac{r}{n}\right)^{k-i} \cdot \left(\frac{n}{r}\right)^{\frac{k}{2}-i} \cdot m = \frac{1}{4k^2 2^k} \cdot \left(\frac{r}{n}\right)^{\frac{k}{2}} \cdot m\,.$$

As an upshot, we know:

Claim 4.4.11. $d(\widehat{K}_r) \geqslant \frac{1}{8k^2 2^k} \cdot \left(\frac{r}{n}\right)^{k/2} \cdot m.$

Spectral double counting. With a lower bound on $d(\widehat{K}_r)$ in hand, we are now ready to perform our weighted spectral double counting argument to complete the proof of Theorem 4.4.1.

Proof of Theorem 4.4.1. Recall that our goal is to prove that there is a small even cover in $\mathcal{H}^{(i)}$, the largest piece obtained from the decomposition, and also recall that if $i \geqslant \frac{k+1}{2}$, then we are done by Lemma 4.4.5. Hence, we assume $i \leqslant \frac{k-1}{2}$ for the rest of the proof.

Suppose there are no even covers in $\mathcal{H}$ of size $\ell = r \log n$, then there are also none in $\mathcal{H}^{(i)}$ from Lemma 4.4.9 we get:

$$\left\|\Gamma^{-1/2} \widehat{A} \Gamma^{-1/2}\right\|_2 \leqslant 4\sqrt{\frac{2\ell}{d(\widehat{K}_r)}}\,.$$

Thus, $\widehat{A} \preceq 4\sqrt{\frac{2\ell}{d(\widehat{K}_r)}}\Gamma$, and by taking the quadratic form with the all-ones vector, we get:

$$2|E(\widehat{K}_r)| = \mathbf{1}^\top \widehat{A} \mathbf{1} \leqslant 4\sqrt{\frac{2\ell}{d(\widehat{K}_r)}} \cdot \mathrm{tr}(\Gamma) = 16\sqrt{\frac{2\ell}{d(\widehat{K}_r)}} \cdot |E(\widehat{K}_r)|\,,$$

which implies

$$d(\widehat{K}_r) \leqslant 128\ell\,,$$

and by our lower bound on $d(\widehat{K}_r)$ from Claim 4.4.11, we get

$$\frac{1}{8k^2 2^k} \cdot \left(\frac{r}{n}\right)^{\frac{k}{2}} \cdot m \leqslant 128 r \log n\,,$$

which we can rearrange as

$$m \leqslant B^k n \log n \cdot \left(\frac{n}{r}\right)^{k/2-1}.$$

for some large enough constant B. Thus, if m is lower bounded as in the theorem statement, there must be an even cover of size $\ell \log n$. □

Chapter 5

Subgraph Density in Spectral Expanders

In this chapter, we prove Lemma 3.2.4 and Theorem 3.2.5, which we restate below.

Lemma (Restatement of Lemma 3.2.4)**.** *Let $G = (L \cup R, E)$ be a bipartite graph, and let the left and right average degrees be $d_1 = \frac{|E|}{|L|}$ and $d_2 = \frac{|E|}{|R|}$, respectively. Then,*

$$(d_1 - 1)(d_2 - 1) \leqslant \rho(B_G)^2 .$$

Theorem (Restatement of Theorem 3.2.5)**.** *Let $\varepsilon \in (0, 0.1)$, and let $3 \leqslant c \leqslant d$ be integers. Let $G = (L \cup R, E)$ be a (c,d)-biregular graph and $S \subseteq L \cup R$ such that $|S| \leqslant d^{-1/\varepsilon}|L \cup R|$. Then, for any $t \geqslant 0$ such that*

$$\rho(B_{G[S]}) \leqslant \frac{1}{2}\left(\sqrt{\lambda^2 - (\sqrt{c-1} + \sqrt{d-1})^2} + \sqrt{\lambda^2 - (\sqrt{c-1} - \sqrt{d-1})^2}\right),$$

where $\lambda = \max(\lambda_2(A_G), \sqrt{c-1} + \sqrt{d-1}) \cdot (1 + O(\varepsilon))$.

We will utilize the Ihara-Bass formula [Iha66, Bas92], which relates the spectral radius of the non-backtracking matrix to the positive definiteness of the *Bethe Hessian*. We recall the following for convenience.

Fact (Restatement of Fact 2.2.2)**.** *For any graph G with n vertices and m edges, the following identity on univariate polynomials is true:*

$$\det(\mathbb{I} - B_G t) = \det(H_G(t)) \cdot (1 - t^2)^{m-n}$$

where $H_G(t) := (D_G - \mathbb{I})t^2 - A_G t + \mathbb{I}$ is the Bethe Hessian *of G.*

Organization. We first prove Lemma 3.2.4 in Section 5.1. Then, we prove Theorem 3.2.5 in Section 5.2.

5.1 Average degree of bipartite graphs

The following is equivalent to Lemma 3.2.4.

Lemma 5.1.1 (Equivalent to Lemma 3.2.4)**.** *Let $G = (L \cup R, E)$ be a bipartite graph, and let the left and right average degrees be $d_1 = \frac{|E|}{|L|}$ and $d_2 = \frac{|E|}{|R|}$, respectively. Then, for any $t \in (-1,1) \setminus \{0\}$ such that $H_G(t) \succeq 0$, we have*

$$(d_1 - 1)(d_2 - 1) \leqslant \frac{1}{t^2}\,.$$

The equivalence to Lemma 3.2.4 is due to Lemma 2.2.3, which we restate below for convenience.

Lemma (Restatement of Lemma 2.2.3)**.** *Let G be a graph and $0 < \alpha < 1$. Then, the spectral radius $\rho(B_G) \leqslant \frac{1}{\alpha}$ if and only if $H_G(t) \succ 0$ for all $t \in [0, \alpha)$. As a result, if $H_G(\frac{1}{\rho})$ has a non-positive eigenvalue for some $\rho > 0$, then $\rho(B_G) \geqslant \rho$.*

Proof of Lemma 5.1.1. We can assume that $d_1, d_2 > 1$, otherwise the statement holds trivially with $|t| < 1$. Let x be the vector such that for $u \in L \cup R$,

$$x_u = \begin{cases} 1 & u \in L, \\ \alpha & u \in R, \end{cases}$$

where $\alpha \in \mathbb{R}$ will be determined later. Recall that $H_G(t) = (D_G - \mathbb{I})t^2 - tA_G + \mathbb{I}$. Since $|E| = d_1|L| = d_2|R|$, we have $x^\top D_G x = d_1|L| + \alpha^2 d_2|R| = (1+\alpha^2)d_1|L|$ and $x^\top A_G x = 2\alpha|E| = 2\alpha d_1|L|$, and substituting $|R| = \frac{d_1}{d_2}|L|$ we get

$$\begin{aligned} x^\top H_G(t) x &= x^\top \left((D_G - \mathbb{I})t^2 - tA_G + \mathbb{I}\right) x \\ &= t^2\left((d_1 - 1)|L| + \alpha^2(d_2 - 1)|R|\right) - t \cdot 2\alpha d_1|L| + \left(|L| + \alpha^2|R|\right) \\ &= |L|\left((d_1 - 1)t^2 - 2t\alpha d_1 + 1\right) + |R|\left(\alpha^2(d_2 - 1)t^2 + \alpha^2\right) \\ &= |L|\left((d_1 - 1)t^2 + \alpha^2 d_1 t^2 - 2t\alpha d_1 + 1 + \frac{d_1}{d_2} \cdot \alpha^2(1 - t^2)\right)\,. \end{aligned}$$

Then, $H_G(t) \succeq 0$ and $t \in (-1,1)$ imply that

$$\frac{1}{d_2} \geqslant \frac{1}{1-t^2}\left(-\frac{(d_1 - 1)t^2 + 1}{d_1\alpha^2} + \frac{2t}{\alpha} - t^2\right)\,.$$

To maximize the right-hand side, we choose $\frac{1}{\alpha} = \frac{d_1 t}{(d_1-1)t^2+1}$, which gives

$$\frac{1}{d_2} \geqslant \frac{1}{1-t^2}\left(\frac{d_1 t^2}{(d_1 - 1)t^2 + 1} - t^2\right) = \frac{1}{1-t^2} \cdot \frac{t^2(d_1 - 1)(1-t^2)}{(d_1 - 1)t^2 + 1} = \frac{1}{1 + \frac{1}{(d_1-1)t^2}}$$

$$\implies d_2 \leqslant 1 + \frac{1}{(d_1 - 1)t^2}\,.$$

Rearranging the above gives $(d_1 - 1)(d_2 - 1) \leqslant 1/t^2$. □

5.2 Non-backtracking matrix of subgraphs in bipartite expanders

The following is equivalent to Theorem 3.2.5.

Theorem 5.2.1 (Equivalent to Theorem 3.2.5)**.** *Let $\varepsilon \in (0, 0.1)$, and let $3 \leqslant c \leqslant d$ be integers. Let $G = (L \cup R, E)$ be a (c, d)-biregular graph and $S \subseteq L \cup R$ such that $|S| \leqslant d^{-1/\varepsilon}|L \cup R|$. Then, for any $t \geqslant 0$ such that*

$$\frac{1}{t} \geqslant \frac{1}{2}\left(\sqrt{\widetilde{\lambda}^2 - (\sqrt{c-1} + \sqrt{d-1})^2} + \sqrt{\widetilde{\lambda}^2 - (\sqrt{c-1} - \sqrt{d-1})^2}\right), \tag{5.1}$$

where $\widetilde{\lambda} = \max(\lambda_2(A_G), \sqrt{c-1} + \sqrt{d-1}) \cdot (1 + O(\varepsilon))$, we have

$$H_{G[S]}(t) \succ 0\,.$$

5.2.1 Proof overview

We prove $H_{G[S]}(t) \succ 0$ by showing that $\langle f, H_{G[S]}(t) f\rangle > 0$ for all $f : S \to \mathbb{R}$. The way we prove $\langle f, H_{G[S]}(t) f\rangle > 0$ is by relating it to a quadratic form of the matrix $H_G(t)$, which we can control via the spectrum of G. In particular, we consider the depth-ℓ regular tree extension T of $G[S]$, and for f we define an appropriate function extension f_t on the tree depending on t (Definition 5.2.3) such that $\langle f, H_{G[S]}(t) f\rangle = \langle f_t, H_T(t) f_t\rangle$. The function f_t additionally has the property that its ℓ_2 mass on vertices r-far from $G[S]$ decays exponentially in r. At a high level, we use the tree extension as a proxy for the ℓ-step neighborhood of S in G, and this is made precise in Section 5.2.4 as we define a natural folded function $\widetilde{f_t}$ of f_t into G (Definition 5.2.8).

This allows us to lower bound $\langle f, H_{G[S]}(t) f\rangle$ by $\langle \widetilde{f_t}, H_G(t) \widetilde{f_t}\rangle$ with some errors. The errors can be bounded using the decay of f_t from the definition, though this requires $t < ((c-1)(d-1))^{-1/4}$ (see Lemma 5.2.6). Ignoring those errors, the proof comes down to showing that

$$\frac{1}{t^4} - \frac{1}{t^2}\left(\lambda^2 - (c-1) - (d-1)\right) + (c-1)(d-1) > 0\,,$$

and we solve the quadratic formula in Lemma 5.2.11 and show that the above gives rise to Eq. (5.1). The full proof is presented in Section 5.2.5.

5.2.2 Tree extensions

We start with defining tree extensions of a graph.

Definition 5.2.2 (Tree extension)**.** For a graph $G = (V, E)$, we say that $T = (V(T), E(T))$ is a tree extension of G if T is obtained by attaching a tree T_r to each vertex $r \in V$, with r being the root. Each vertex $x \in T$ belongs to a unique tree T_r rooted at r. For any $x \in T$, we write $\mathsf{depth}(x)$ to be the distance between x and the root of the tree containing x.

Fix a tree extension T of G, for functions $f, g : V(T) \to \mathbb{R}$, define $\langle f, g \rangle = \sum_{x \in T} f(x)g(x)$ and $\|f\|_2^2 = \sum_{x \in T} f(x)^2$.

Definition 5.2.3 (Function extension)**.** Given a function $f : V(G) \to \mathbb{R}$, a tree extension T of G, and parameter $t \in \mathbb{R}$, we define $f_t : V(T) \to \mathbb{R}$ to be the *extension* of f to T such that for $x \in T$,

$$f_t(x) = f(r) \cdot t^{\mathsf{depth}(x)}, \quad \text{if } x \in T_r\,. \tag{5.2}$$

The following simple but crucial lemma establishes a relationship between H_G and H_T, which also motivates the definition of f_t.

Lemma 5.2.4. *Let G be a graph and T be any tree extension of G. Then, for any $t \in \mathbb{R}$ and $f : V(G) \to \mathbb{R}$, the extension $f_t : V(T) \to \mathbb{R}$ defined in Eq. (5.2) satisfies*

$$\left(H_T(t) \cdot f_t\right)(x) = \begin{cases} (H_G(t) \cdot f)(x) & x \in V(G)\,, \\ 0 & x \notin V(G)\,. \end{cases}$$

Proof. Recall that $H_G(t) = (D_G - \mathbb{I})t^2 - A_G t + \mathbb{I}$. For $x \notin V(G)$, let $d(x)$ be its degree and let $r \in V(G)$ be the root of the tree containing x. Observe that x has 1 parent (with value $f(r)t^{\mathsf{depth}(x)-1}$) and $d(x) - 1$ children (with value $f(r)t^{\mathsf{depth}(x)+1}$) in the tree T_r. Thus,

$$\begin{aligned} \left(H_T(t) \cdot f_t\right)(x) &= ((d(x) - 1)t^2 + 1) \cdot f(r)t^{\mathsf{depth}(x)} \\ &\qquad - t \cdot f(r)\left(t^{\mathsf{depth}(x)-1} + (d(x) - 1)t^{\mathsf{depth}(x)+1}\right) \\ &= 0\,. \end{aligned}$$

For $x \in V(G)$, let $d_G(x)$ be its degree in G and $d_T(x)$ be its degree in T. Then, x has $d_T(x) - d_G(x)$ children (with value $t \cdot f(x)$) in the tree T_x.

$$\begin{aligned} \left(H_T(t) \cdot f_t\right)(x) &= ((d_T(x) - 1)t^2 + 1) \cdot f(x) - t\Big((A_G f)(x) + (d_T(x) - d_G(x)) \cdot tf(x)\Big) \\ &= ((d_G(x) - 1)t^2 + 1) \cdot f(x) - t(A_G f)(x) \\ &= (H_G(t) \cdot f)(x)\,. \end{aligned}$$

This completes the proof. □

5.2.3 Regular tree extensions of subgraphs

For a subgraph $G[S]$ in a regular (or biregular) graph, we consider its regular tree extension.

Definition 5.2.5 (Regular tree extension)**.** Let $G = (V, E)$ be a d-regular graph, $S \subseteq V$, $\ell \in \mathbb{N}$, and consider the induced subgraph $G[S]$. We define the *depth-ℓ regular tree extension* of $G[S]$ to be the tree extension T of $G[S]$ where depth-ℓ trees are attached to vertices in S such that the resulting graph is d-regular except for the leaves. Let $\mathsf{Leaves}(T)$ denote the set of leaves.

Similarly, let $G = (L \cup R, E)$ be a (c, d)-biregular graph, $S \subseteq L \cup R$, and $\ell \in \mathbb{N}$. The depth-ℓ regular tree extension of $G[S]$ is the tree extension such that the resulting graph is (c, d)-biregular except for the leaves.

We show that given a graph $G = (V, E)$ and $S \subseteq V$, for any function $f : S \to \mathbb{R}$ and its extension f_t to the depth-ℓ regular tree extension of $G[S]$, the contribution from the leaves decays exponentially with ℓ when $t < ((c-1)(d-1))^{-1/4}$.

Lemma 5.2.6 (Decay of f_t)**.** *Let $G = (L \cup R, E)$ be a (c, d)-biregular graph with $c \leqslant d$, let $S \subseteq L \cup R$, let $\ell \in \mathbb{N}$ be even, and let T be the depth-ℓ regular tree extension of $G[S]$. Moreover, let $t \in \mathbb{R}$ such that $t^2\sqrt{(c-1)(d-1)} = 1 - \delta$ for some $\delta \in (0, 1)$. Given any function $f : S \to \mathbb{R}$, let $f_t : V(T) \to \mathbb{R}$ be the function extension (as defined in Eq. (5.2)), and let $f_t^{=\ell}$ be f_t restricted to the leaves of T. Then,*

$$\left\| f_t^{=\ell} \right\|_2^2 \leqslant \frac{2\delta}{e^{\delta\ell} - 1} \cdot \|f_t\|_2^2 \,.$$

Since the function $\frac{2x}{e^{x\ell}-1}$ is monotone decreasing, we have for any $\delta' > \delta$,

$$\left\| f_t^{=\ell} \right\|_2^2 \leqslant \frac{2\delta'}{e^{\delta'\ell} - 1} \cdot \|f_t\|_2^2 \,.$$

Proof. We will lower bound $\|f_t\|_2^2$ and upper bound the contribution from the leaves at depth ℓ. Fix a vertex $r \in R$ (with $\deg_G(r) = d$) and consider the tree T_r rooted at r. Let $\deg_{T_r}(r)$ be the degree of r in T_r. The number of children of vertices in the tree alternates between $c-1$ and $d-1$ as we go down the tree. Thus, for an even integer $k \leqslant \ell$, the number of vertices in the k-th level is

$$\deg_{T_r}(r)(c-1)\left((c-1)(d-1)\right)^{\frac{k}{2}-1} = \frac{\deg_{T_r}(r)}{\deg_G(r) - 1}\left((c-1)(d-1)\right)^{\frac{k}{2}} . \tag{5.3}$$

The same argument shows that the above also holds for $r \in L$ (with $\deg_G(r) = c$). Thus, the contribution of the tree T_r to $\|f_t\|_2^2$ can be lower bounded by the product of

the following two terms:

$$f(r)^2 \frac{\deg_{T_r}(r)}{\deg_G(r)-1}$$

$$\sum_{\substack{0\leqslant k\leqslant \ell \\ k \text{ even}}} t^{2k}((c-1)(d-1))^{\frac{k}{2}} = \sum_{i=0}^{\ell/2}(1-\delta)^{2i} = \frac{1-(1-\delta)^{\ell+2}}{1-(1-\delta)^2} \geqslant \frac{1-(1-\delta)^{\ell}}{2\delta}\,.$$

Next, the contribution from the leaves of T_r to $\left\|f_t^{=\ell}\right\|_2^2$ is also given by Eq. (5.3). Thus, we have

$$\frac{\left\|f_t^{=\ell}\right\|_2^2}{\|f_t\|_2^2} \leqslant (1-\delta)^{\ell}\frac{2\delta}{1-(1-\delta)^{\ell}} \leqslant \frac{2\delta}{e^{\delta\ell}-1}\,,$$

using $(1-\delta)^{\ell} \leqslant e^{-\delta\ell}$, finishing the proof. □

5.2.4 Folding regular tree extensions

Given a regular tree extension T of an induced subgraph $G[S]$, there is a natural folding into G via breadth-first search from S.

Definition 5.2.7 (Folding into G)**.** Let $G=(V,E)$ be a d-regular or (c,d)-biregular graph, let $S \subseteq V$, and let T be the depth-ℓ regular tree extension of $G[S]$. There is a natural homomorphism $\sigma : T \to G$ such that

- $\sigma(x) = x$ for all $x \in S$;
- $\deg_T(x) = \deg_G(\sigma(x))$ for all $x \in V(T) \setminus \mathsf{Leaves}(T)$;
- Two edges $\{x,y\}$ and $\{y,z\}$ in T sharing a vertex are not mapped to the same edge in E, i.e., all edges in T that map to the same edge in E are vertex-disjoint.

Definition 5.2.8 (Folded function)**.** Fix a map $\sigma : T \to G$. Given any $f : V(T) \to \mathbb{R}$, we associate each vertex $v \in G$ with a function $f^v : V(T) \to \mathbb{R}$ such that for $x \in T$,

$$f^v(x) = \begin{cases} f(x) & \text{if } \sigma(x) = v\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

We define the *folded* function $\widetilde{f} : V(G) \to \mathbb{R}$ to be

$$\widetilde{f}(v) = \|f^v\|_2\,.$$

Observation 5.2.9. The f^v's have disjoint support, thus $\|\widetilde{f}\|_2^2 = \sum_{v\in G}\|f^v\|_2^2 = \|f\|_2^2$. More generally, let $\Gamma, \widetilde{\Gamma}$ be diagonal operators such that $(\Gamma f)(x) = \gamma(\deg_G(\sigma(x)))f(x)$ for $x \in T$ and $(\widetilde{\Gamma} g)(v) = \gamma(\deg_G(v))g(v)$ for $v \in G$. Then, $\langle \widetilde{f}, \widetilde{\Gamma}\widetilde{f}\rangle = \langle f, \Gamma f\rangle$.

We next prove the following useful lemma that relates the quadratic forms of f and $\widetilde{f}$ with A_G.

Lemma 5.2.10. *Let $G = (V, E)$ be a d-regular or (c, d)-biregular graph, let $S \subseteq V$, and let T be a regular tree extension of $G[S]$. For any $f : V(T) \to \mathbb{R}$ and its folded function $\widetilde{f} : V(G) \to \mathbb{R}$, we have*

$$\langle f, A_T f \rangle \leqslant \left\langle \widetilde{f}, A_G \widetilde{f} \right\rangle .$$

Proof. Recall from Definition 5.2.7 that the map $\sigma : T \to G$ satisfies that if $\{x, y\}$ is an edge in T, then $\{\sigma(x), \sigma(y)\} \in E$. Then,

$$\begin{aligned}\langle f, A_T f \rangle &= 2 \sum_{\{x,y\} \in E(T)} f(x) f(y) \\ &= 2 \sum_{\{u,v\} \in E(G)} \sum_{\{x,y\} \in E(T)} \mathbf{1}(\sigma(\{x, y\}) = \{u, v\}) \cdot f^{\sigma(x)}(x) f^{\sigma(y)}(y) .\end{aligned}$$

Moreover, all edges in T that map to the same edge are vertex-disjoint. Thus, for any $\{u, v\} \in E$, $\sum_{\{x,y\} \in E(T)} \mathbf{1}(\sigma(\{x, y\}) = \{u, v\}) \cdot f^{\sigma(x)}(x) f^{\sigma(y)}(y)$ can be expressed as an inner product between some permutations of f^u and f^v, which is upper bounded by $\|f^u\|_2 \cdot \|f^v\|_2 = \widetilde{f}(u)\widetilde{f}(v)$ by Cauchy-Schwarz. Thus, we have

$$\langle f, A_T f \rangle \leqslant 2 \sum_{\{u,v\} \in E(G)} \widetilde{f}(u)\widetilde{f}(v) = \left\langle \widetilde{f}, A_G \widetilde{f} \right\rangle . \qquad \square$$

5.2.5 Proof of Theorem 3.2.5

Before we prove Theorem 3.2.5, we first prove the following lemma for convenience.

Lemma 5.2.11. *Let $3 \leqslant c \leqslant d \in \mathbb{N}$ and $\varepsilon \in (0, 1)$. Let $\lambda \geqslant \sqrt{c-1} + \sqrt{d-1}$ and $\widetilde{\lambda} = \lambda(1 + \varepsilon)$. Then, for all x such that*

$$x \geqslant \frac{1}{2}\left(\sqrt{\widetilde{\lambda}^2 - (\sqrt{c-1} + \sqrt{d-1})^2} + \sqrt{\widetilde{\lambda}^2 - (\sqrt{c-1} - \sqrt{d-1})^2}\right) ,$$

we have

$$x^4 - x^2(\lambda^2(1 + \varepsilon) - (c + d - 2)) + (c - 1)(d - 1) > 0 .$$

Proof. Denote $a := c - 1$ and $b := d - 1$ for convenience. Then, to show that $x^4 - x^2(\lambda^2(1 + \varepsilon) - a - b) + ab \geqslant 0$, it suffices to verify that

$$x^2 > \frac{1}{2}\left(\lambda^2(1 + \varepsilon) - a - b\right) + \frac{1}{2}\sqrt{(\lambda^2(1 + \varepsilon) - a - b)^2 - 4ab} .$$

Squaring both sides of $x \geqslant \frac{1}{2}\left(\sqrt{\widetilde{\lambda}^2-(\sqrt{a}+\sqrt{b})^2}+\sqrt{\widetilde{\lambda}^2-(\sqrt{a}-\sqrt{b})^2}\right)$, we get

$$\begin{aligned} x^2 &\geqslant \frac{1}{2}\left(\widetilde{\lambda}^2-a-b\right)+\frac{1}{2}\sqrt{(\widetilde{\lambda}^2-(\sqrt{a}+\sqrt{b})^2)(\widetilde{\lambda}^2-(\sqrt{a}-\sqrt{b})^2)} \\ &= \frac{1}{2}\left(\widetilde{\lambda}^2-a-b\right)+\frac{1}{2}\sqrt{\widetilde{\lambda}^4-2(a+b)\widetilde{\lambda}^2+(a-b)^2} \\ &= \frac{1}{2}\left(\widetilde{\lambda}^2-a-b\right)+\frac{1}{2}\sqrt{(\widetilde{\lambda}^2-a-b)^2-4ab}\,, \end{aligned}$$

which completes the proof with $\widetilde{\lambda}=\lambda(1+\varepsilon)$. □

Proof of Theorem 3.2.5. We first verify that the assumption on t (Eq. (5.1)) implies that

$$t^2 \leqslant \frac{1-\varepsilon}{\sqrt{(c-1)(d-1)}}\,. \tag{5.4}$$

Indeed, as $\widetilde{\lambda}=\lambda(1+O(\varepsilon)) \geqslant (\sqrt{c-1}+\sqrt{d-1})(1+\varepsilon)$, Eq. (5.1) implies that

$$\begin{aligned} \frac{1}{t^2} &\geqslant \frac{1}{4}\left((\sqrt{c-1}+\sqrt{d-1})^2(1+\varepsilon)^2-(\sqrt{c-1}-\sqrt{d-1})^2\right) \\ &\geqslant \sqrt{(c-1)(d-1)}+\frac{1}{2}(\sqrt{c-1}+\sqrt{d-1})^2\varepsilon \end{aligned}$$

which implies Eq. (5.4).

We would like to show that $\langle f, H_{G[S]}(t)f\rangle > 0$ for any function $f : S \to \mathbb{R}$. Let $\ell = \lceil \frac{1}{2\varepsilon} \rceil$ be an even integer and let T be the depth-ℓ regular tree extension of $G[S]$ (Definition 5.2.5). Let $f_t : V(T) \to \mathbb{R}$ be the function extension of f to T with parameter t. By Lemma 5.2.4, we have

$$\left\langle f, H_{G[S]}(t)f\right\rangle = \langle f_t, H_T(t)f_t\rangle = \left\langle f_t, ((D_T-\mathbb{I})t^2 - tA_T+\mathbb{I})f_t\right\rangle .$$

Note that all internal vertices $x \in T \setminus \mathsf{Leaves}(T)$ have degree c or d while the leaves have degree 1. Let D'_T be the diagonal matrix such that the leaves have the "correct" degree, i.e., for $x \in \mathsf{Leaves}(T)$ in the tree T_r rooted at $r \in S$, $D'_T[x,x] = \deg_G(r)$ (since ℓ is even). Then, by Eq. (5.4), Lemma 5.2.6 states that $f_t^{=\ell}$ decays with a factor $\frac{2\varepsilon}{e^{\varepsilon\ell}-1} \leqslant 4\varepsilon$, thus

$$\langle f_t, (D_T-\mathbb{I})f_t\rangle = \langle f_t, (D'_T-\mathbb{I})f_t\rangle - \left\langle f_t^{=\ell}, (D'_T-\mathbb{I})f_t^{=\ell}\right\rangle \geqslant \langle f_t, (D'_T-\mathbb{I})f_t\rangle\,(1-4\varepsilon)\,.$$

Consider the folded function $\widetilde{f}_t : V(G) \to \mathbb{R}$ as defined in Definition 5.2.8. By Observation 5.2.9, we have $\langle f_t, D'_T f_t\rangle = \langle \widetilde{f}_t, D_G\widetilde{f}_t\rangle$ and $\|f_t\|_2^2 = \|\widetilde{f}_t\|_2^2$. Moreover, by Lemma 5.2.10, $\langle f_t, A_T f_t\rangle \leqslant \left\langle \widetilde{f}_t, A_G\widetilde{f}_t\right\rangle$. Thus,

$$\begin{aligned} \left\langle f, H_{G[S]}(t)f\right\rangle &\geqslant t^2\left\langle \widetilde{f}_t, (D_G-\mathbb{I})\widetilde{f}_t\right\rangle(1-4\varepsilon) - t\left\langle \widetilde{f}_t, A_G\widetilde{f}_t\right\rangle + \left\|\widetilde{f}_t\right\|_2^2 \\ &\geqslant (1-4\varepsilon)\left\langle \widetilde{f}_t, \left(t^2(D_G-\mathbb{I})+\mathbb{I}\right)\widetilde{f}_t\right\rangle - t\left\langle \widetilde{f}_t, A_G\widetilde{f}_t\right\rangle . \end{aligned} \tag{5.5}$$

We would like to show that the above is non-negative. Denote $\Gamma_G := t^2(D_G - \mathbb{I}) + \mathbb{I}$, and $\gamma_1 := t^2(c-1)+1$ and $\gamma_2 := t^2(d-1)+1$. Note that $\gamma_2 \geqslant \gamma_1 > 0$ as we assume that $c \leqslant d$. Since G is a (c,d)-biregular graph, Γ_G and A_G have the following block structure,

$$\Gamma_G = \begin{pmatrix} \gamma_1 \mathbb{I} & 0 \\ 0 & \gamma_2 \mathbb{I} \end{pmatrix}, \quad A_G = \begin{pmatrix} 0 & A_{L,R} \\ A_{L,R}^\top & 0 \end{pmatrix}.$$

In particular,

$$\begin{aligned}(1-4\varepsilon)\Gamma_G - tA_G &= \Gamma_G^{1/2}\left((1-4\varepsilon)\mathbb{I} - t\cdot\Gamma_G^{-1/2}A_G\Gamma_G^{-1/2}\right)\Gamma_G^{1/2} \\ &= \Gamma_G^{1/2}\left((1-4\varepsilon)\mathbb{I} - \frac{t}{\sqrt{\gamma_1\gamma_2}}A_G\right)\Gamma_G^{1/2}.\end{aligned}$$

Then, denoting $g := \Gamma_G^{1/2}\widetilde{f}_t$, we can write Eq. (5.5) as

$$\left\langle f, H_{G[S]}(t) f\right\rangle \geqslant \left\langle \widetilde{f}_t, \left((1-4\varepsilon)\Gamma_G - tA_G\right)\widetilde{f}_t\right\rangle = (1-4\varepsilon)\left\|g\right\|_2^2 - \frac{t}{\sqrt{\gamma_1\gamma_2}}\left\langle g, A_G g\right\rangle . \quad (5.6)$$

Next, we upper bound $\langle g, A_G g\rangle$. For any (c,d)-biregular graph, the (normalized) top eigenvector of A_G is $\frac{1}{\sqrt{2|E|}}D_G^{1/2}\vec{1}$ with eigenvalue $\sqrt{cd}$. Thus,

$$\left\langle g, A_G g\right\rangle \leqslant \frac{\sqrt{cd}}{2|E|}\left\langle g, D_G^{1/2}\vec{1}\right\rangle^2 + \lambda\left\|g\right\|_2^2,$$

where $\lambda = \max(\lambda_2(A_G), \sqrt{c-1}+\sqrt{d-1})$ is the second eigenvalue.

Since T has depth ℓ, the support of $\widetilde{f}_t$ (and g) must be contained in $B := \{v \in V(G) : \mathrm{dist}(v,S) \leqslant \ell\}$. We have $|B| \leqslant |S|\sum_{i=0}^{\ell} d^i \leqslant |S|d^{\ell+1}$. Thus, by Cauchy-Schwarz,

$$\frac{\sqrt{cd}}{2|E|}\left\langle g, D_G^{1/2}\vec{1}\right\rangle^2 \leqslant \frac{\sqrt{cd}}{2|E|}\cdot d|B|\cdot\left\|g\right\|_2^2 \leqslant d^{-1/4\varepsilon}\left\|g\right\|_2^2 \leqslant \varepsilon\left\|g\right\|_2^2,$$

since $|S| \leqslant d^{-1/\varepsilon}|L\cup R|$, $\ell = \lceil\frac{1}{2\varepsilon}\rceil$, $|E| = c|L| = d|R|$, and $\varepsilon \leqslant 0.1$ (note that $d^{-1/4\varepsilon} \leqslant \varepsilon$ for all $d \geqslant 3$ and $\varepsilon \leqslant 0.1$).

Thus, $\langle g, A_G g\rangle \leqslant (\lambda+\varepsilon)\left\|g\right\|_2^2 \leqslant \lambda(1+\varepsilon)\left\|g\right\|_2^2$, and from Eq. (5.6),

$$\left\langle f, H_{G[S]}(t) f\right\rangle \geqslant \frac{1}{\sqrt{\gamma_1\gamma_2}}\left((1-4\varepsilon)\sqrt{\gamma_1\gamma_2} - t\lambda(1+\varepsilon)\right).$$

As $\frac{1+\varepsilon}{1-4\varepsilon} \leqslant 1+5\varepsilon$, to prove that the above is positive, it suffices to prove that $t^2\lambda^2(1+5\varepsilon) < \gamma_1\gamma_2 = (t^2(c-1)+1)(t^2(d-1)+1)$, or equivalently,

$$\frac{1}{t^4} - \frac{1}{t^2}\left(\lambda^2(1+5\varepsilon) - (c-1) - (d-1)\right) + (c-1)(d-1) > 0.$$

With $\widetilde{\lambda} = \lambda(1+5\varepsilon)$ and the assumption on t (Eq. (5.1)), the above holds via Lemma 5.2.11. □

Part II

Algorithms

Chapter 6

Introduction

Spectral techniques have a long history in algorithm design. Many computational problems naturally involve matrices, either as explicit inputs or as representations of useful underlying structures. Graph problems, for example, are natural applications for spectral methods, as graphs can be represented by their adjacency matrices. Similarly, constraint satisfaction problems (CSPs), such as k-SAT or k-XOR, can be formulated in terms of hypergraphs, where spectral techniques may also apply.

By now, we have a wide range of spectral techniques at our disposal — including expander decomposition, spectral sparsification, eigenspace enumeration, random matrix concentration — which serve as powerful tools in algorithmic design.

We focus on three foundational problems where spectral techniques are key ingredients in the algorithms and analyses.

(1) **Section 6.1 and Chapter 7: Algorithms for strongly refuting semirandom CSPs.** The same spectral techniques used to prove the hypergraph Moore bound (Chapter 4) extend naturally to refutation algorithms for semirandom CSPs, achieving the same improved bounds as in the hypergraph Moore bound. This chapter is based on [HKM23].

(2) **Section 6.2 and Chapter 8: Algorithms for solving semirandom planted CSPs.** We present an efficient algorithm for solving semirandom planted k-CSPs that recovers the planted assignment whenever the number of constraints exceeds $\widetilde{O}(n^k)$. This matches the threshold at which polynomial-time algorithms are known for the refutation problem. This chapter is based on [GHKM23].

(3) **Section 6.3 and Chapter 9: Finding large independent sets in expanders.** We give algorithms to find linear-sized independent sets in one-sided spectral expanders that are 3-colorable or contain independent sets of size close to $n/2$. This chapter is based on [BHK25].

6.1 Algorithms for strongly refuting semirandom CSPs

Over the past decades, complexity theory has established strong hardness results for constraint satisfaction problems (CSPs) like k-SAT in the *worst-case*. Håstad's inapproximability theorem [Hås01] shows that *sparse* instances (i.e., has $m = O(n)$ constraints on n variables) cannot be approximated better than by picking a uniformly random assignment (unless $\mathsf{P} = \mathsf{NP}$). For maximally dense instances (e.g., with $m = \Theta(n^k)$ constraints for k-SAT), there are polynomial-time approximation schemes (PTAS) due to [AKK95]. However, under the exponential time hypothesis [IP01], we can already rule out polynomial-time algorithms for $o(n^k)$ dense instances and more generally, $2^{n^{1-\delta}}$ time algorithms for any $\delta > 0$ for $o(n^{k-1})$ dense instances [FLP16].

In sharp contrast, in well-studied *average-case* settings, there appears to be significant space for new algorithms and markedly better guarantees for CSPs. In this section, we focus on the problem of *refutation* — efficiently outputting certificates of unsatisfiability for instances drawn from distributions that are almost always unsatisfiable. The related *search* problem for *planted* CSP models will be discussed in Section 6.2.

We first define the notion of a refutation algorithm formally:

Definition 6.1.1 (Refutation algorithm)**.** A refutation algorithm takes a CSP instance ψ as input and outputs a value alg-val$(\psi) \in [0,1]$ such that alg-val$(\psi) \geqslant$ val(ψ) for all instance ψ. Here, val(ψ) denotes the *value* of ψ — the maximum fraction of constraints satisfied by any assignment.

For a distribution $\mathcal{D}$ over instances, we say that an algorithm achieves

- *weak refutation* if alg-val$(\psi) < 1$ with high probability over $\psi \sim \mathcal{D}$,
- *strong refutation* if alg-val$(\psi) < 1 - \delta$ for some constant $\delta > 0$ with high probability over $\psi \sim \mathcal{D}$.

The refutation problem has been heavily investigated in the past two decades. For *fully random* k-CSPs with uniformly random clause structure (i.e., which variables appear in each clause) and "literal pattern" (i.e., which variables appear negated in each clause), there is a polynomial-time strong refutation algorithm when the number of clauses m is at least $\widetilde{O}(n^{k/2})$ [GL03, CGL07, AOW15, BM16]. Note that this threshold is far below the $\sim n^k$ hardness threshold of [FLP16]. Furthermore, lower bounds in various restricted models [Fei02, BGMT12, OW14, MW16, BCK15, KMOW17, FPV18] provide some evidence that this threshold might be tight for polynomial-time algorithms.

Building on [AOW15, BM16], Raghavendra, Rao and Schramm [RRS17] showed a smooth trade-off between m and the running time of the refutation algorithm. Specifically, for a parameter $r \leqslant n$, there is a strong refutation algorithm that runs in time $n^{O(r)}$ provided that $m \geqslant \widetilde{O}(n \cdot (\frac{n}{r})^{\frac{k}{2}-1})$. Their result offers a fairly comprehensive understanding of the refutation problem for fully random k-CSPs.

***Beyond* the average-case: semirandom instances.** Despite the phenomenal progress in average-case algorithm design, like refuting random CSPs discussed above, there is a nagging concern that the algorithms so developed rely too heavily on "brittle" properties of the specific random models. That is, our methods may have "overfitted" to the specific setting, thus yielding algorithms that only apply in a limited setting. Unfortunately, this concern turns out to be well justified — natural spectral algorithms for refuting random k-CSPs break down under minor perturbations, such as the addition of a vanishingly small fraction of extra clauses.

Motivated by such concerns, Blum and Spencer [BS95] and later Feige and Kilian [FK01, Fei07] introduced *semirandom* models for optimization problems. In semirandom models, the instances are constructed by a combination of benign average-case and adversarial worst-case choices. Algorithms that succeed for such models are naturally "robust" to perturbations of the input instance.

For CSPs, a *semirandom* instance is generated by first choosing a "worst-case" clause structure and then choosing the literal negation patterns in each clause via some sufficiently random (and thus "benign") process. Abascal, Guruswami and Kothari [AGK21] gave algorithms that succeed in refuting semirandom instances at the same $\widetilde{O}(n^{k/2})$ threshold as the *fully random* case. Later, Guruswami, Kothari and Manohar [GKM22], building on the work of [WAM19] for even k, showed that the smooth runtime vs. density trade-off established by [RRS17] holds also for the semirandom case, up to an extra $\log^{4k} n$ factor.

We improve this trade-off to just one extra $\log n$ factor with a substantially simpler and shorter proof. The following theorem is stated only for k-XOR. Combined with Feige's (now standard) "XOR principle" [Fei02, AOW15], we also obtain refutation algorithms for all *smoothed* Boolean CSPs. For refutation, we will omit such reduction in this thesis. For solving the planted problem (Section 6.2), we will need a more complicated reduction from general CSPs to XOR, which is detailed in Section 8.2.

Theorem 6.1.2 (Semirandom k-XOR refutation; informal Theorem 7.0.1)**.** *Fix $k \in \mathbb{N}$ and $r \leqslant n$, there is an $n^{O(r)}$-time algorithm such that given a semirandom k-XOR instance ψ with n variables and $m \geqslant O(n(\frac{n}{r})^{\frac{k}{2}-1} \cdot \log n)$ constraints, it certifies that ψ is not $(\frac{1}{2} + 0.01)$-satisfiable with high probability.*

A careful reader may notice that the term $n(\frac{n}{r})^{\frac{k}{2}-1}$ coincides with the number of hyperedges required in a k-uniform hypergraph to guarantee an even cover of size $O(r \log n)$, as given by the hypergraph Moore bound (Theorem 3.1.6). In fact, the original conjecture of Feige [Fei08] was directly motivated by the study of refuting k-SAT formulas, and [GKM22] made the elegant observation that the same proof for refuting semirandom k-XOR also yields the hypergraph Moore bound.

In Chapter 7, we prove Theorem 6.1.2. The proof is very similar to the proof of

Theorem 3.1.6 in Chapter 4, relying on the same core ideas — the use of a reweighted Kikuchi matrix and an edge deletion step — that allow us to remove several involved steps in [GKM22]'s analysis.

6.2 Efficient algorithms for semirandom planted CSPs

The search problem for planted models of CSPs has also received a fair bit of attention. The setting naturally arises in the investigation of *local* one-way functions and pseudorandom generators in cryptography. Indeed, the security of the well-known one-way function proposed by Goldreich [Gol00] (also conjectured to be a pseudorandom generator [MST06, App16]) is equivalent to the hardness of recovering a satisfying assignment planted (via a carefully chosen procedure) in a random CSP instance with an appropriate predicate. This has led to significant research on solving *fully random* planted CSPs [BHL+02, JMS07, BQ09, CCF10, FPV15]. Specifically, Feldman, Perkins and Vempala [FPV15] showed that for *fully random* planted k-CSPs with planted assignment x^*, there is a polynomial-time algorithm that, with high probability over the instance, recovers the planted assignment x^* *exactly*, provided that the instance has at least $\widetilde{O}(n^{k/2})$ constraints. That is, the refutation and search versions have the same clause threshold.

Given the results for refuting semirandom CSPs as discussed in Section 6.1, it is natural to wonder if the same thresholds hold also for the *search* variant of the problem. We make the first step in this direction: we give an efficient algorithm for solving semirandom planted CSPs that succeeds in finding the planted assignment whenever the number of constraints exceeds $\widetilde{O}(n^{k/2})$ — the *same* threshold at which polynomial-time algorithms exist for the refutation problem for random (and semirandom) instances.

Theorem 6.2.1 (Algorithm for planted CSPs; informal Theorem 8.0.3)**.** *There is an efficient algorithm that takes as input a k-CSP Ψ and outputs an assignment x with the following guarantee: if Ψ is a semirandom planted k-CSP with $m \geqslant \widetilde{O}(n^{k/2})$ constraints, then with high probability over Ψ, the output x satisfies $1 - o(1)$-fraction of the constraints in Ψ.*

See Definition 8.0.2 in Chapter 8 for the precise definition of a semirandom planted k-ary Boolean CSP.

We note that in the semirandom setting, it is not possible to efficiently recover an assignment that satisfies *all* of the constraints without being able to do so even when $m = O(n)$.[1] This is because it is easy to construct a semirandom instance ψ that is the "union" of two disjoint instances ψ_1 and ψ_2, where ψ_1 and ψ_2 use disjoint sets of $n/2$ variables, but ψ_1 only has $m_1 \sim O(n)$ clauses (and ψ_2, therefore, contains almost all of

[1] Achieving this would break a hardness assumption for the search problem analogous to Feige's random 3-SAT hypothesis for the refutation problem [Fei02].

the $m \sim n^{k/2}$ clauses). Thus, the guarantee in Theorem 6.2.1 of satisfying a $1 - o(1)$-fraction of constraints is qualitatively the best we can hope for.

Search vs. refutation. For average-case optimization problems, techniques for refuting random instances can often be adapted to solving the search problem in the related planted model. This can be formalized in the *proofs to algorithms* paradigm [BS14, FKP19] where spectral/SDP-based refutations can be transformed into "simple" (i.e., "captured" within the low-degree sum-of-squares proof system) efficient certificates of near-uniqueness of optimal solution — that is, every optimal solution is close to the planted assignment. Unfortunately, this intuition breaks down even in the simplest setting of semirandom 2-XOR where there can be multiple maximally far-off solutions that satisfy as many (or even more) constraints as the planted assignment. Such departure from uniqueness also breaks algorithms for recovery [FPV15] that rely on the top eigenvector of a certain matrix built from the instance being correlated with the planted assignment. In the semirandom setting, one can build instances where the top eigenspace of such matrices is the span of the multiple optimal solutions and has dimension $\omega(1)$ (searching for a Boolean vector close to the subspace is, in general, hard in super-constant dimensional subspaces).

Noisy planted k-XOR. Similar to work on random planted CSPs [FPV15] and the refutation setting [AOW15, RRS17, AGK21, GKM22, HKM23], our proof of Theorem 6.2.1 goes through a reduction to noisy k-XOR. Our algorithm achieves very strong guarantees in the noisy k-XOR case, as we now explain. We define the noisy k-XOR model below and then state our result.

Definition 6.2.2 (Noisy planted k-XOR)**.** Let $\mathcal{H} \subseteq \binom{[n]}{k}$ be a k-uniform hypergraph on n vertices, let $x^* \in \{\pm 1\}^n$, and let $\eta \in [0, 1/2)$. Let $\psi(\mathcal{H}, x^*, \eta)$ denote the distribution on k-XOR instances over n variables $x_1, \dots, x_n \in \{\pm 1\}$ obtained by, for each $C \in \mathcal{H}$, adding the constraint $\prod_{i \in C} x_i = \prod_{i \in C} x_i^*$ with probability $1 - \eta$, and otherwise adding the constraint $\prod_{i \in C} x_i = -\prod_{i \in C} x_i^*$. In the latter case, we say that the constraint C is *corrupted* or *noisy*.

We call ψ a *noisy planted* k-XOR instance if it is sampled from $\psi(\mathcal{H}, x^*, \eta)$, for some $\mathcal{H}$, x^*, and η; the hypergraph $\mathcal{H}$ is the constraint hypergraph, x^* is the planted assignment, and η is the noise parameter. Furthermore, we let $\mathcal{E}_\psi \subseteq \mathcal{H}$ denote the (unknown) set of corrupted constraints.

Theorem 6.2.3 (Algorithm for noisy k-XOR)**.** *Let $\eta \in [0, 1/2)$, let $k, n \in \mathbb{N}$, and let $\varepsilon \in (0, 1)$. Let $m \geqslant cn^{k/2} \cdot \frac{k^4 \log^3 n}{\varepsilon^5 (1-2\eta)^4}$ for a universal constant c. There is a polynomial-time algorithm $\mathcal{A}$ that takes as input a k-XOR instance ψ with constraint hypergraph $\mathcal{H}$ and outputs two disjoint sets $\mathcal{A}_1(\mathcal{H}), \mathcal{A}_2(\psi) \subseteq \mathcal{H}$ with the following guarantees: (1) for any instance ψ with m constraints, $|\mathcal{A}_1(\mathcal{H})| \leqslant \varepsilon m$ and $\mathcal{A}_1(\mathcal{H})$ only depends on $\mathcal{H}$, and (2) for any $x^* \in \{\pm 1\}^n$*

and any k-uniform hypergraph $\mathcal{H}$ with at least m hyperedges, with probability at least $1 - 1/\operatorname{poly}(n)$ over $\psi \leftarrow \psi(\mathcal{H}, x^*, \eta)$, it holds that $\mathcal{A}_2(\psi) = \mathcal{E}_\psi \cap (\mathcal{H} \setminus \mathcal{A}_1(\mathcal{H}))$.

In words, the algorithm discards a small number of constraints, and among the constraints that are not discarded, correctly identifies all (and only) the corrupted constraints. In particular, the subinstance obtained by discarding the $\lesssim (\varepsilon + \eta)m$ constraints $\mathcal{A}_1(\mathcal{H}) \cup \mathcal{A}_2(\psi)$ is satisfiable (and a solution can be found by Gaussian elimination). Thus, Theorem 6.2.3 immediately implies that for k-XOR, the NP-hard task of deciding if ψ has value $\geqslant 1 - \eta$ or $\leqslant \frac{1}{2} + \eta$ is actually *easy* if ψ has $\sim n^{k/2}$ constraints (far below the $\sim n^k$-hardness of [FLP16]), provided that the η-fraction of corrupted constraints in the "yes" case are a *randomly chosen subset* of the otherwise arbitrary constraints.

Exact vs. approximate recovery. As alluded to above, the guarantees of Theorem 6.2.3 are much stronger: not only can we find a good assignment to ψ, we can break the constraints into two parts, a small fraction, $\mathcal{A}_1(\mathcal{H})$, where we are unable to determine the corrupted constraints,[2] and a large fraction, $\mathcal{H} \setminus \mathcal{A}_1(\mathcal{H})$, where we can determine *exactly* all of the corrupted constraints, $\mathcal{A}_2(\psi)$. Moreover, this partition depends only on the hypergraph $\mathcal{H}$ and is *independent of the noise*. We remark that it is not immediately obvious that this guarantee is achievable even for exponential-time algorithms, as x^* may not be the globally optimal assignment with constant probability. This strong guarantee of Theorem 6.2.3 is in fact required for the reduction from Theorem 6.2.1 to Theorem 6.2.3; the weaker (and more intuitive) guarantee of approximate recovery — obtaining an assignment of value $1 - \eta - o(1)$ for the noisy XOR instance — is insufficient for the reduction.

One can view Theorem 6.2.3 as an algorithm that extracts almost all the information about the planted assignment x^* encoded by the instance ψ. Indeed, notice that even if $\eta = 0$, the instance ψ only determines x^* "up to a linear subspace."[3] Namely, if we let $y \in \{\pm 1\}^n$ be any solution to the system of constraints $\prod_{i \in C} y_i = 1$ for $C \in \mathcal{H}$, then $y \odot x^*$ is also a planted assignment for ψ: formally, $\psi(\mathcal{H}, x^*, \eta) = \psi(\mathcal{H}, y \odot x^*, \eta)$ as distributions. So, aside from the εm constraints that are discarded, with high probability over ψ the algorithm determines the uncorrupted right-hand sides $\prod_{i \in C} x^*_i$ for every remaining constraint, which is all the information about the planted assignment x^* encoded in the remaining constraints.

The importance of relative spectral approximation. As a key technical ingredient in the algorithm, we uncover a *deterministic* condition — relative spectral approximation of the Laplacian of a graph (associated with the input instance) by a certain correlated

[2]Note that discarding a small fraction of constraints is necessary in the semirandom setting, as ψ may contain many disconnected constant-size subinstances where it is not possible, even information-theoretically, to exactly identify the corrupted constraints with $1 - o(1)$ probability.

[3]A k-XOR constraint $x_{C_1} \cdots x_{C_k} = b_C \in \{\pm 1\}$ can be equivalently written as a linear equation $x'_{C_1} + \cdots + x'_{C_k} = b'_C$ over $\mathbb{F}_2$, where we map $+1$ to 0 and -1 to 1.

random sample from it — which when satisfied implies uniqueness of the SDP solution (Lemma 8.1.4). In Lemma 8.1.5 and Lemma 8.4.7, we establish such spectral approximation guarantees.

This spectral approximation property is the key ingredient in our certificate of unique identifiability of the planted assignment in a noisy k-XOR instance (see Section 8.1.4 for details). This property allows us to *exactly* recover the planted assignment for 2-XOR instances where the constraint graph G is a weak spectral expander (i.e., spectral gap $\gg 1/\operatorname{polylog} n$) (Lemma 8.1.4), and it forms the backbone of our final algorithm. We note that our spectral approximation condition can be seen as an analog of (and is, in fact, stronger than) the related spectral norm upper bound property that underlie the refutation algorithm of [AGK21].

6.3 Finding large independent sets in expanders

Finding large independent sets is a notoriously hard problem in the worst case. The best known algorithms can only find independent sets of size $\widetilde{O}(\log^2 n)$ in n-vertex graphs with independent sets of near-linear size [Fei04]. In this paper, we are interested in the important setting when the input graph contains an independent set of size cn for a large constant $c < 1/2$.[4] In this setting, the problem remains challenging and has served as a benchmark for developing new techniques in approximation algorithm design over the years. When $c = 1/2 - \varepsilon$ for tiny enough $\varepsilon > 0$, a generalization of the SDP rounding of Karger, Motwani, and Sudan [KMS98] finds an independent set of size $n^{1-O(\varepsilon)}$ (see Section 9.5). When $c \ll 1/2$, all known efficient algorithms [BH92, AK98] can only find independent sets of size $n^{\delta(c)}$ for some $\delta(c) < 1$, and this is true even when the graph is k-colorable (thus $c \geqslant 1/k$). Decades of research on coloring k-colorable graph has progressively improved the constant appearing in the exponent, with the most recent improvement being in 2024 [Wig83, Blu94, BK97, KMS98, ACC06, Chl09, KT17, KTY24]. To summarize, in the worst-case, even when c approaches $1/2$, our best known efficient algorithms can only find independent sets of size that is a polynomial factor smaller than n.

There is evidence that the difficulties in improving the above algorithms might be inherent. Assuming the Unique Games Conjecture (UGC), for any constant $\varepsilon > 0$, it is NP-hard to find an independent set of size εn even when the input graph contains an independent set of size $(1/2 - \varepsilon)n$ [KR08, BK09]. Similar hardness results suggest that it may be hard to color 3-colorable graphs with any constant number of colors [DS05, DMR06, DKPS10, KS12, GS20].

Given the above worst-case picture, a substantial effort over the past three decades

[4]The classical 2-approximation algorithm for vertex cover (complement of an independent set) implies an algorithm to find an independent set of size $2\varepsilon n$ when the input graph has one of size $(1/2 + \varepsilon)n$.

has explored algorithms that work under natural structural assumptions on the input graphs. One line of work studies *planted average-case* models for independent set [Kar72, Jer92, Kuč95] and coloring [BS95, AK97]. A related body of research has focused on graphs that satisfy natural, deterministic assumptions, such as expansion, with the goal of isolating simple and concrete properties of random instances that enable efficient algorithms. This approach has been explored for Unique Games [Tre08, AKK$^+$08, MM11, ABS15, BBKSS21] and UG-hard problems like Max-Cut and Sparsest Cut [DHV16, RV17], and has been instrumental in making progress even for worst-case instances; for example, the works on unique games on expanders eventually led to a subexponential algorithm for arbitrary UG instances [ABS15]. Over the past decade, such assumptions have also been investigated for independent set and coloring [AG11, DF16, KLT18]. In particular, a recent work of David and Feige [DF16] gave polynomial-time algorithms for finding large independent sets in *planted* k-colorable expander graphs. We discuss these works in more detail next.

Prior works and one-sided vs two-sided expansion. There is a crucial difference between the expansion assumptions in prior works on coloring vs other problems. A d-regular graph whose normalized adjacency matrix $\frac{1}{d}A$ (a.k.a., the uniform random walk matrix) has eigenvalues $1 = \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ is called a *one-sided* spectral expander if for some $\lambda < 1$ (λ is called the *spectral gap*), $\lambda_2 \leqslant \lambda$, and a *two-sided* spectral expander, $\max\{\lambda_2, |\lambda_n|\} \leqslant \lambda$. Most algorithms for problems (e.g., Unique Games and other constraint satisfaction problems) on expanders only need one-sided spectral expansion, as they primarily rely on the *conductance*, or the fraction of edges leaving any subset of vertices in the graph, a combinatorial property closely related to λ_2 via Cheeger's inequality. In contrast, previous algorithms for finding independent sets in expanders with a planted k-coloring rely on two-sided spectral expansion (i.e., control of even the negative end of the spectrum).

This is *not* just a technical quirk; the main observation underlying such algorithms (due to Alon and Kahale [AK97], following Hoffman [Hof70]) is that a random graph is a two-sided spectral expander (thus, has no large negative eigenvalues) and that planting a k-coloring in it introduces *negative* eigenvalues of large magnitude, whose corresponding eigenvectors are correlated with indicator vectors of the color classes. This allows using the bottom eigenvectors of the graph to obtain a coarse *spectral clustering*. All the works above, including those on deterministic expander graphs [DF16], build on this basic observation for their algorithmic guarantees.

This basic idea becomes inapplicable if we are working with *one-sided* spectral expanders that behave markedly differently in the context of graph coloring. To illustrate this point, we observe the following proposition with a simple proof (see Section 9.4) which implies that there is likely no efficient algorithm to find an $\Omega(n)$-sized independent set in an ε-almost 4-colorable graph (i.e., 4-colorable if one removes ε fraction of

vertices), even when promised to have nearly perfect one-sided spectral expansion with $\lambda_2 \leqslant o_n(1)$!

Proposition 6.3.1 (See Proposition 9.4.2)**.** *Assuming the Unique Games Conjecture, for any constants $\varepsilon, \gamma > 0$, it is NP-hard to find an independent set of size γn in an n-vertex regular graph that is ε-almost 4-colorable and has normalized 2nd eigenvalue $\lambda_2 \leqslant o_n(1)$.*

This is in sharp contrast to David and Feige's algorithm [DF16] which shows how to find a planted k-coloring in a sufficiently strong two-sided spectral expander for any constant k.[5]

We prove Proposition 6.3.1 by a reduction from the UG-hardness of finding linear-sized independent sets in ε-almost 2-colorable graphs [BK09] and guaranteeing one-sided expansion in addition at the cost of obtaining an almost 4-colorable graph. A similar reduction allows us to show hardness of finding linear-sized independent sets in exactly 6-colorable ($\varepsilon = 0$) one-sided spectral expanders (see Proposition 9.4.6).

We are thus led to the main question:

Can polynomial-time algorithms find a large independent set in a 3-colorable one-sided spectral expander?

Proposition 6.3.1 injects a fair amount of intrigue into this question, but our motivations for studying it go further. In light of the above discussion, an affirmative answer would necessarily require developing a new algorithmic approach that departs from previous spectral clustering methods based on bottom eigenvectors since such techniques do not distinguish between 3 vs 4-colorable graphs.

Let us spoil the intrigue: we develop new algorithms for finding large independent sets via rounding sum-of-squares (SoS) relaxations. Our polynomial-time algorithms succeed in finding linear-sized independent sets in almost 3-colorable graphs that satisfy one-sided spectral expansion. Given the UG-hardness (i.e., Proposition 6.3.1) of finding linear-sized independent sets in an almost 4-colorable one-sided expander, we obtain a stark and surprising difference between almost 3-colorable and almost 4-colorable one-sided expander graphs.

Theorem 6.3.2 (Informal Theorem 9.3.1)**.** *There is a polynomial-time algorithm that, given an n-vertex regular 10^{-4}-almost 3-colorable graph with normalized 2nd eigenvalue $\lambda_2 \leqslant 10^{-4}$, finds an independent set of size $\geqslant 10^{-4}n$.*

Our techniques succeed without the 3-colorability assumption if the input graph has an independent set of size $(1/2 - \varepsilon)n$, and satisfies a weaker quantitative one-sided

[5] [DF16] focused on finding a partial or full coloring, which requires the planted coloring to be roughly balanced. Their spectral clustering technique can likely find a large independent set even when the coloring is not balanced.

spectral expansion.

Theorem 6.3.3. *For every positive $\varepsilon \leqslant 0.001$, there is a polynomial-time algorithm that, given an n-vertex regular graph that contains an independent set of size $(\frac{1}{2} - \varepsilon)n$ and has normalized 2nd eigenvalue $\lambda_2 \leqslant 1 - 40\varepsilon$, outputs an independent set of size at least $10^{-3}n$.*

Note that we get an algorithm for ε-almost 2-colorable one-sided expanders as an immediate corollary. Before this work, no algorithm that beat the worst-case guarantee of outputting a $n^{1-O(\varepsilon)}$-sized independent set [KMS98] was known in this setting.

Follow-up work. In a follow-up work [Hsi25] that is not included in this thesis, we improve Theorem 6.3.2 to 3-colorable graphs with small (one-sided) threshold rank. Specifically, given an n-vertex 3-colorable graph whose uniform random walk matrix has at most r eigenvalues larger than ε, our algorithm finds a proper 3-coloring on at least $(\frac{1}{2} - O(\varepsilon))n$ vertices in time $n^{O(r/\varepsilon^2)}$.

Furthermore, in another work [BHSV25], Buhai, Hua, Steurer, and Vári-Kakas showed that it is UG-hard to properly 3-color more than $(\frac{1}{2} + \varepsilon)n$ vertices even assuming one-sided expansion, thus establishing the tightness of the result in [Hsi25]. On the flip side, [BHSV25] also showed an algorithm that properly 3-colors almost all vertices if all color classes have size bounded away from $1/2$.

Chapter 7

Algorithms for Strongly Refuting Semirandom CSPs

In this chapter, we prove Theorem 6.1.2. As discussed in Section 6.1, our improved proof for the hypergraph Moore bound (Theorem 3.1.6) extends to strong refutation algorithms for semirandom k-XOR, losing only a single $\log n$ factor in the density.

Theorem 7.0.1 (Semirandom k-XOR refutation; formal statement of Theorem 6.1.2)**.** *Fix $k \in \mathbb{N}$. There is an algorithm with parameter $r \in \mathbb{N}$, $2k \leqslant r \leqslant n/8$ that takes as input a semirandom k-XOR instance*

$$\psi(x) = \frac{1}{m} \sum_{C \in \mathcal{H}} b_C x_C$$

where $\mathcal{H}$ is a k-uniform hypergraph with n vertices and m hyperedges, and each $b_C \in \{\pm 1\}$ is chosen uniformly at random. The algorithm has the following guarantee: there is a universal constant C such that if $m \geqslant C^k n \log n \cdot (\frac{n}{r})^{\frac{k}{2}-1} \varepsilon^{-4}$ for $\varepsilon \in (0, 1/2)$, then with probability over $1 - \frac{1}{\mathrm{poly}(n)}$ over $\{b_C\}_{C \in \mathcal{H}}$, the algorithm runs in time $n^{O(r)}$ and certifies that $\psi(x) \leqslant \varepsilon$.

Combined with Feige's "XOR principle" [Fei02, AOW15], we also obtain refutation algorithms for all *smoothed* Boolean CSPs. We will omit such reduction in this work and direct the reader to [GKM22] for a detailed exposition.

Remark 7.0.2 (Refutation strength: dependence on ε)**.** For the even arity case, we actually obtain a stronger guarantee (weaker requirement) of $m \geqslant O(n \log n) \cdot (\frac{n}{r})^{\frac{k}{2}-1} \varepsilon^{-2}$. For the odd arity case however, our analysis incurs a (likely suboptimal) dependence of $1/\varepsilon^4$ on the refutation strength (i.e., the upper bound on the value of the input k-XOR instance), though improving the $1/\varepsilon^5$ dependence of [GKM22, Theorem 5.1]. In contrast, a $1/\varepsilon^2$ dependence is known to hold for fully random k-XOR instances [RRS17]. Apart from a somewhat unsatisfying deficiency, this suboptimality turns out to be consequential – in particular, it changes the threshold at which efficient FKO refutation witnesses

exist for semirandom k-SAT (and other CSPs) by a polynomial factor in n. Finding the "right" dependence of $1/\varepsilon^2$ (for the odd case) is an interesting open problem.

Our refutation algorithm will utilize the same Kikuchi graphs from Definition 4.3.2 and Definition 4.4.6 but with signs added to the edges in the natural way.

Definition 7.0.3 (Signed Kikuchi graph)**.** Let $\mathcal{H}$ be a k-uniform hypergraph associated with $\{\pm 1\}$ signs $\{b_C\}_{C\in\mathcal{H}}$. For the even arity case, let A_b be the signed adjacency matrix of the Kikuchi graph from Definition 4.3.2 where each edge $S \overset{C}{\leftrightarrow} T$ has a sign b_C. For the odd arity case, let A_b be the signed adjacency matrix of the Kikuchi graph from Definition 4.4.6 where each edge $S \overset{C,C'}{\longleftrightarrow} T$ has a sign $b_C b_{C'}$.

7.1 Refuting semirandom even arity XOR

In this section, we prove Theorem 7.0.1 when k is even. As we will see in the short proof, our idea of the reweighted Kikuchi matrix from the hypergraph Moore bound naturally applies here, and in fact, we obtain the "right" $1/\varepsilon^2$ dependence in this case, i.e., we can certify that $\psi(x) \leqslant \varepsilon$ when $m \geqslant O(n\log n)\cdot(\frac{n}{r})^{\frac{k}{2}-1}\varepsilon^{-2}$.

Recall that in the Kikuchi graph (V,E), each $C\in\mathcal{H}$ contributes $\alpha := \frac{1}{2}\binom{k}{k/2}\binom{n-k}{r-k/2}$ edges in E, hence $|E| = \frac{1}{2}|V|d = m\alpha$. Thus, it is clear that

$$\psi(x) = \frac{1}{m}\cdot\frac{1}{\alpha}\sum_{(S,T)\in E} b_{S\oplus T}x_{S\oplus T} = \frac{1}{\binom{n}{r}d}(x^{\odot r})^\top A_b x^{\odot r} \tag{7.1}$$

where $x^{\odot r}\in\{\pm1\}^{\binom{n}{r}}$ and the S-entry of $x^{\odot r}$ is x_S for $S\subseteq[n]$, $|S|=r$.

We now follow the same reweighting strategy: with $\Gamma = D + d\mathbb{I}$, we bound the spectral norm of the reweighted Kikuchi matrix $\left\|\Gamma^{-1/2}A_b\Gamma^{-1/2}\right\|_2$ with an almost identical proof as Lemma 4.3.4.

Lemma 7.1.1. *Let k be even and $r\in\mathbb{N}$. Let A_b be the signed Kikuchi graph with random $\{\pm1\}$ coefficients $\{b_C\}_{C\in\mathcal{H}}$, and let $\Gamma = D + d\mathbb{I}$ where D is the degree matrix and d is the average degree of the Kikuchi graph. Then, with probability at least $1-\frac{1}{\mathrm{poly}(n)}$ over the randomness of $\{b_C\}_{C\in\mathcal{H}}$,*

$$\left\|\Gamma^{-1/2}A_b\Gamma^{-1/2}\right\|_2 \leqslant O\left(\sqrt{\frac{r\log n}{d}}\right).$$

Proof. Let $\widetilde{A}_b = \Gamma^{-1/2}A_b\Gamma^{-1/2}$. We again use the trace power method $\|\widetilde{A}_b\|_2^\ell \leqslant \mathrm{tr}((\Gamma^{-1}A_b)^\ell)$ where we choose an even $\ell = 2\lceil r\log_2 n\rceil$. Observe that in expectation, $\mathbb{E}_b\,\mathrm{tr}((\Gamma^{-1}A_b)^\ell)$ counts the closed walks that use each hyperedge an even number of times. This is exactly the same as Lemma 4.3.4 where we count closed walks in an unsigned Kikuchi

graph assuming there is no even cover of size $\leqslant \ell$. Thus, Lemma 4.3.4 shows that

$$\mathbb{E}_b \operatorname{tr}((\Gamma^{-1}A_b)^\ell) \leqslant 2^\ell n^r \left(\frac{\ell}{d}\right)^{\ell/2} \leqslant O\left(\frac{\ell}{d}\right)^{\ell/2}$$

when $\ell \geqslant r \log_2 n$. Then, by Markov's inequality, for any $\lambda > 0$,

$$\Pr_b\left[\|\widetilde{A}_b\|_2 \geqslant \lambda\right] = \Pr_b\left[\|\widetilde{A}_b\|_2^\ell \geqslant \lambda^\ell\right] \leqslant \lambda^{-\ell} \cdot \mathbb{E}_b \operatorname{tr}((\Gamma^{-1}A_b)^\ell) \leqslant O\left(\frac{\ell}{\lambda^2 d}\right)^{\ell/2}$$

Choosing $\lambda = O(\sqrt{\ell/d})$ completes the proof. □

We can complete the proof of Theorem 7.0.1 for even k.

Proof of Theorem 7.0.1 for even k. Let A_b be the signed Kikuchi graph with signs $\{b_C\}_{C\in\mathcal{H}}$, let $\Gamma = D + d\mathbb{I}$ where D is the degree matrix and d is the average degree of the Kikuchi graph, and let $\widetilde{A}_b = \Gamma^{-1/2}A_b\Gamma^{-1/2}$. The certification algorithm is simply to compute $\|\widetilde{A}_b\|_2$. Since $A_b \preceq \|\widetilde{A}_b\|_2 \cdot \Gamma$, and $\operatorname{tr}(\Gamma) = 2\binom{n}{r}d$, by Lemma 7.1.1,

$$\psi(x) = \text{Eq. (7.1)} \leqslant \frac{1}{\binom{n}{r}d}\|\widetilde{A}_b\|_2 \cdot \operatorname{tr}(\Gamma) \leqslant O\left(\sqrt{\frac{r\log n}{d}}\right)$$

using the fact that $x^{\odot r} \in \{\pm 1\}^{\binom{n}{r}}$ and $(x^{\odot r})^\top \Gamma x^{\odot r} = \operatorname{tr}(\Gamma)$. There is some constant C such that when $m \geqslant Cn\log n \cdot (\frac{n}{r})^{\frac{k}{2}-1}\varepsilon^{-2}$, by Eq. (4.2) the average degree $d \geqslant \frac{1}{2}(\frac{r}{n})^{k/2}m = \frac{C}{2}r\log n \cdot \varepsilon^{-2}$, thus giving us $\psi(x) \leqslant \varepsilon$. This completes the proof. □

7.2 Refuting semirandom odd arity XOR

Our proof of Theorem 7.0.1 for the odd arity case closely mimics the steps taken in proving the hypergraph Moore bound for odd arity hypergraphs (Theorem 4.4.1). Given a semirandom k-XOR instance ψ on hypergraph $\mathcal{H}$ with random signs $\{b_C\}_{C\in\mathcal{H}}$, we first apply the following hypergraph decomposition algorithm (a variant of Algorithm 4.4.2) to decompose the hypergraph into subhypergraphs $\mathcal{H}^{(1)}, \dots, \mathcal{H}^{(k-1)}$. The main difference compared to Algorithm 4.4.2 is that in the final step, we add the "leftover" hyperedges to $\mathcal{H}^{(1)}$ instead of an extra $\mathcal{H}^{(0)}$.

Algorithm 7.2.1 (Hypergraph decomposition)**.** Given a k-uniform hypergraph $\mathcal{H}$ on n vertices and m hyperedges, and thresholds $\tau_1, \dots, \tau_{k-1} \geqslant 2$, we partition $\mathcal{H}$ into hypergraphs $\mathcal{H}^{(1)}, \dots, \mathcal{H}^{(k-1)}$ via the following algorithm.

1. Set $t = k - 1$ and $\mathcal{H}_{\text{current}} := \mathcal{H}$.

2. Set counter $s = 1$. While there is $T \subseteq [n]$ such that $|T| = t$ and $|\{C \in \mathcal{H}_{\text{current}} : T \subseteq C\}| \geqslant \tau_t$:
   (a) Choose T satisfying the condition and let $\mathcal{H}_s^{(t)}$ be a subset of $\{C \in \mathcal{H}_{\text{current}} : T \subseteq C\}$ of size τ_t.
   (b) Add all clauses in $\mathcal{H}_s^{(t)}$ to $\mathcal{H}^{(t)}$.
   (c) Delete all clauses in $\mathcal{H}_s^{(t)}$ to $\mathcal{H}_{\text{current}}$.
   (d) Increment s by 1.
3. Decrement t by 1. If $t > 0$, go back to step 2; otherwise take the remaining clauses in $\mathcal{H}_{\text{current}}$ and partition them into n parts $F_1, \ldots, F_n$ where each clause C goes to some F_i such that $i \in C$. Add $F_1, \ldots, F_n$ to $\mathcal{H}^{(1)}$ and terminate.

Notations and parameters. Throughout this section we will use the following notations.

- In Algorithm 7.2.1, we set thresholds $\tau_t = \max\left\{1, \left(\frac{n}{r}\right)^{\frac{k}{2}-t}\right\} \cdot 4k\varepsilon^{-2}$.
- In the decomposition, each $\mathcal{H}^{(t)}$ contains p_t groups $\mathcal{H}_1^{(t)}, \ldots, \mathcal{H}_{p_t}^{(t)}$ where group $\mathcal{H}_i^{(t)}$ has a center $T_i^{(t)}$ of size t, and for each $C \in \mathcal{H}_i^{(t)}$, we write $\widetilde{C} = C \setminus T_i^{(t)}$.
- Each $|\mathcal{H}_i^{(t)}| = \tau_t$, with the exception that $|\mathcal{H}_i^{(1)}| \leqslant \tau_1$ may have different sizes (the leftover hyperedges in Algorithm 7.2.1). Let $m_t := \sum_{i=1}^{p_t} |\mathcal{H}_i^{(t)}|$ be the total number of hyperedges in $\mathcal{H}^{(t)}$.
- When $t = 1$ and $m \geqslant C^k n \log n \cdot (\frac{n}{r})^{\frac{k}{2}-1}\varepsilon^{-4}$ for a large enough constant C, we have $m \geqslant n\tau_1$, hence $p_1 \leqslant \frac{m}{\tau_1} + n \leqslant \frac{2m}{\tau_1}$. Thus, we will use $p_t\tau_t \leqslant 2m$ for all $t \in [k-1]$.
- For each $t \in [k-1]$, the colored Kikuchi graph (V, E) obtained from $\mathcal{H}^{(t)} = (\mathcal{H}_1^{(t)}, \ldots, \mathcal{H}_{p_t}^{(t)})$ (from Definition 4.4.6) has edges $|E| = \alpha_t \sum_{i=1}^{p_t} \binom{|\mathcal{H}_i^{(t)}|}{2} \leqslant \frac{1}{2}\alpha_t m_t \tau_t$, where $\alpha_t \approx (\frac{2n}{r})^{r-(k-t)}$ is the number of edges contributed by each distinct pair $C, C' \in \mathcal{H}_i$ (see Observation 4.4.8).

With these notations and parameters in mind, we can write $\psi(x)$ as

$$\psi(x) = \frac{1}{m}\sum_{t=1}^{k-1}\sum_{C\in\mathcal{H}^{(t)}} b_C x_C = \frac{1}{k}\sum_{t=1}^{k-1}\psi_t(x)$$

$$\text{where} \quad \psi_t(x) := \frac{k}{m}\sum_{i=1}^{p_t}\sum_{C\in\mathcal{H}_i^{(t)}} b_C x_C = \frac{k}{m}\sum_{i=1}^{p_t} x_{T_i}\sum_{C\in\mathcal{H}_i^{(t)}} b_C x_{\widetilde{C}}\,. \tag{7.2}$$

Essentially, each ψ_t is the sub-instance of ψ restricted to the partition $\mathcal{H}^{(t)}$. Recall that for the purpose of showing existence of even covers, we only need to focus on one $\mathcal{H}^{(t)}$. For refutation however, we need to certify a bound on $\psi_t(x)$ for all $t \in [k-1]$.

Lemma 7.2.2 (Refuting each ψ_t)**.** *Fix an odd $k \in \mathbb{N}$, $t \in [k-1]$, and let $2k \leqslant r \leqslant n/8$. There is a constant C such that given a semirandom k-XOR instance ψ with n variables and $m \geqslant C^k n \log n (\frac{n}{r})^{\frac{k}{2}-1} \varepsilon^{-4}$ clauses for $\varepsilon \in (0, 1/2)$, and suppose ψ_t is the subinstance from Eq. (7.2) obtained by the hypergraph decomposition algorithm (Algorithm 7.2.1), then with probability $1 - \frac{1}{\mathrm{poly}(n)}$ over the random signs, we can certify that $\psi_t(x) \leqslant \varepsilon$ in $n^{O(r)}$ time.*

Lemma 7.2.2 immediately completes the proof of Theorem 7.0.1 for odd k.

Proof of Theorem 7.0.1 by Lemma 7.2.2. Given the hypergraph $\mathcal{H}$, we apply the hypergraph decomposition algorithm (Algorithm 7.2.1) with thresholds $\tau_1, \dots, \tau_{k-1}$ and obtain subinstances $\psi_1, \dots, \psi_{k-1}$ as in Eq. (7.2). For each $t \in [k-1]$, we can certify that $\psi_t(x) \leqslant \varepsilon$ by Lemma 7.2.2 with high probability, which immediately implies the desired bound $\psi(x) \leqslant \varepsilon$. □

Edge deletion process. The proof of Lemma 7.2.2 requires deleting the "bad" edges from the signed Kikuchi matrix $A_b^{(t)}$ via a similar deletion process as the one used in the proof of Theorem 4.4.1, but with some parameter $\eta > 1$ instead of 1 and an additional *equalizing* step:

> Start with the colored Kikuchi graph, and delete every edge $\{S, T\}$ caused by a pair of clauses $C, C' \in \mathcal{H}_i^{(t)}$ such that S or T has more than η edges that C or C' participates in.
>
> Suppose $\rho < 1$ is the maximum fraction of edges deleted among all pairs of clauses. Then, for every $i \in [p_t]$ and every distinct pair $C, C' \in \mathcal{H}_i^{(t)}$, we delete (additional) edges caused by C, C' arbitrarily such that exactly ρ fraction of edges are deleted.

Observation 7.2.3 (Uniform deletion)**.** The final step in the above edge deletion process ensures that every pair C, C' contributes the *same* number of edges ($(1-\rho)\alpha_t$ to be exact) in the Kikuchi graph.

Mirroring the proof of Claim 4.4.10 yields the following generalization.

Lemma 7.2.4 (Deletion rate)**.** *Suppose a subhypergraph $\mathcal{H}^{(i)}$ satisfies that for any $s \geqslant i$ and any $T \subseteq [n]$ with $|T| = s$, the number of hyperedges in $\mathcal{H}^{(i)}$ containing T is at most τ_s, then the deletion process with parameter $\eta \geqslant 1$ satisfies*

$$\Pr_{\{S,T\} \sim E_{C,C'}}[\{S, T\} \text{ deleted}] \leqslant \frac{4^k}{\eta} \cdot \sum_{s=i}^{\lfloor \frac{k+i}{2} \rfloor} \tau_s \left(\frac{r}{n}\right)^{\lfloor \frac{k+i}{2} \rfloor - s}.$$

Proof. The proof is identical to the proof of Claim 4.4.10. Eq. (4.3) holds with an additional $1/\eta$ factor due to Markov's inequality. The lemma statement then follows immediately from Eq. (4.4). □

Proof of Lemma 7.2.2 via the Cauchy-Schwarz trick and the deletion process.

Proof of Lemma 7.2.2. We apply the Cauchy-Schwarz trick to ψ_t from Eq. (7.2):

$$\psi_t(x)^2 \leqslant \frac{k}{m^2}\sum_{i=1}^{p_t} x_{T_i}^2 \cdot \sum_{i=1}^{p_t}\left(\sum_{C\in\mathcal{H}_i^{(t)}} b_C x_{\widetilde{C}}\right)^2 \leqslant \frac{kp_t}{m^2}\sum_{i=1}^{p_t}\sum_{C,C'\in\mathcal{H}_i^{(t)}} b_C b_{C'} x_{\widetilde{C}} x_{\widetilde{C}'}$$
$$\leqslant \frac{kp_t m_t}{m^2} + \frac{kp_t}{m^2}\sum_{i=1}^{p_t}\sum_{C\neq C'\in\mathcal{H}_i^{(t)}} b_C b_{C'} x_{C\oplus C'} \tag{7.3}$$

since $x \in \{\pm 1\}^n$, $b_C \in \{\pm 1\}$ and $\sum_{i=1}^{p_t}|\mathcal{H}_i^{(t)}| = m_t$. For the first term, since for all $t \in [k-1]$, we set $\tau_t \geqslant 4k\varepsilon^{-2}$ and $p_t \leqslant 2m/\tau_t \leqslant \frac{m\varepsilon^2}{2k}$, thus

$$\frac{kp_t m_t}{m^2} \leqslant \frac{\varepsilon^2}{2}. \tag{7.4}$$

We can now focus our attention on the second term in Eq. (7.3).

Given $\mathcal{H}^{(t)}$ and its partitions $\mathcal{H}_1^{(t)},\dots,\mathcal{H}_{p_t}^{(t)}$ of size τ_t, and signs $\{b_C\}_{C\in\mathcal{H}^{(t)}}$, let $A_b^{(t)}$ be the signed Kikuchi matrix defined in Definition 7.0.3, which is the signed version of the colored Kikuchi graph (V,E) from Definition 4.4.6. Recall from Observation 4.4.8 that each distinct pair $C,C' \in \mathcal{H}_i^{(t)}$ contributes $\alpha_t \approx (\frac{2n}{r})^{r-(k-t)}$ edges in the graph. Thus, similar to Eq. (7.1) in the even case, we can write the second term of Eq. (7.3) as a quadratic form:

$$f_t(x) := \frac{kp_t}{m^2}\sum_{i=1}^{p_t}\sum_{C\neq C'\in\mathcal{H}_i^{(t)}} b_C b_{C'} x_{C\oplus C'} = \frac{kp_t}{2\alpha_t m^2}(x^{\odot r})^\top A_b^{(t)} x^{\odot r} \tag{7.5}$$

where $x^{\odot r} \in \{\pm 1\}^{\binom{2n}{r}}$ such that for $S \in [n]\times[2]$ with $S = (S^{(1)},S^{(2)})$ (green and blue elements), the S-entry of $x^{\odot r}$ is $x_{S^{(1)}\oplus S^{(2)}}$.

We proceed to certify an upper bound on $f_t(x)$. Given the signed Kikuchi matrix $A_b^{(t)}$, we first apply the deletion process with parameter $\eta = B^k\varepsilon^{-2}$ for some large enough constant B. With the chosen thresholds τ_s, Lemma 7.2.4 states that the deletion probability ρ is at most

$$\rho \leqslant \frac{4^k}{\eta}\cdot\sum_{s=t}^{\lfloor\frac{k+t}{2}\rfloor} 4k\varepsilon^{-2}\cdot\max\left\{1,\left(\frac{n}{r}\right)^{\frac{k}{2}-s}\right\}\cdot\left(\frac{r}{n}\right)^{\lfloor\frac{k+t}{2}\rfloor-s} \leqslant \frac{1}{2},$$

since $s \leqslant \lfloor\frac{k+t}{2}\rfloor$ in the summation and $\lfloor\frac{k+t}{2}\rfloor \geqslant \frac{k+1}{2}$ for all $t \geqslant 1$.

Let $\widehat{A}_b^{(t)}$ be the Kikuchi matrix after the deletion process. By Observation 7.2.3, each distinct pair $C, C' \in \mathcal{H}_i^{(t)}$ contributes exactly $(1-\rho)$ fraction of the original edges. Thus, we have

$$(x^{\odot r})^\top \widehat{A}_b^{(t)} x^{\odot r} = (1-\rho) \cdot (x^{\odot r})^\top A_b^{(t)} x^{\odot r}. \tag{7.6}$$

Next, we follow the same argument as the proof of Lemma 7.1.1 to analyze $\widehat{A}_b^{(t)}$, using the norm bound of Lemma 4.4.9. Let $\Gamma = D + d\mathbb{I}$ where D is the degree matrix and d is the average degree, and let $\widetilde{A}_b = \Gamma^{-1/2}\widehat{A}_b^{(t)}\Gamma^{-1/2}$. To bound $\|\widetilde{A}_b\|_2$, we again use the trace power method $\|\widetilde{A}_b\|_2^\ell \leqslant \operatorname{tr}((\Gamma^{-1}\widehat{A}_b^{(t)})^\ell)$ where we choose an even $\ell = 2\lceil r \log_2 n \rceil$. Observe that in expectation, $\mathbb{E}_b \operatorname{tr}((\Gamma^{-1}A_b)^\ell)$ counts the closed walks that use each hyperedge an even number of times. This is exactly the same as Lemma 4.4.9 where we count closed walks in an unsigned Kikuchi graph assuming there is no even cover of size $\leqslant \ell$. Furthermore, $d_{S,i} \leqslant \eta$ is automatically satisfied after the deletion process. Thus, we can directly apply Lemma 4.4.9 and show that

$$\mathbb{E}_b \operatorname{tr}\left((\Gamma^{-1}\widehat{A}_b^{(t)})^\ell\right) \leqslant 2^\ell n^r \left(\frac{2\eta\ell}{d}\right)^{\ell/2} \leqslant O\left(\frac{\eta\ell}{d}\right)^{\ell/2}$$

when $\ell \geqslant r \log_2 n$. Then, by Markov's inequality, we have that $\Pr_b\left[\|\widetilde{A}_b\|_2 \geqslant O\left(\sqrt{\frac{\eta\ell}{d}}\right)\right] \leqslant \frac{1}{\operatorname{poly}(n)}$.

Thus, with high probability we have $\widehat{A}_b^{(t)} \preceq O\left(\sqrt{\frac{\eta\ell}{d}}\right) \cdot \Gamma$, then since $\operatorname{tr}(\Gamma) = 4|E|$,

$$(x^{\odot r})^\top \widehat{A}_b^{(t)} x^{\odot r} \leqslant O\left(\sqrt{\frac{\eta\ell}{d}}\right) \cdot \operatorname{tr}(\Gamma) = O\left(\sqrt{\frac{\eta\ell}{d}}\right) \cdot |E|.$$

Next, let $\widehat{f}_t(x) = \frac{kp_t}{2\alpha_t m^2}(x^{\odot r})^\top \widehat{A}_b^{(t)} x^{\odot r}$. By Observation 4.4.8, we have

$$d \geqslant (\frac{r}{2n})^{k-t} \sum_{i=1}^{p_t} \binom{|\mathcal{H}_i^{(t)}|}{2}$$

when $2k \leqslant r \leqslant n/8$. Plugging in parameters $|E| = \alpha_t \sum_{i=1}^{p_t} \binom{|\mathcal{H}_i^{(t)}|}{2}$, $p_t\tau_t \leqslant 2m$, $\eta = B^k\varepsilon^{-2}$, and $\ell = 2\lceil r \log_2 n \rceil$, standard calculations show that

$$\begin{aligned}\widehat{f}_t(x) &\leqslant O(1)\frac{kp_t}{\alpha_t m^2}\sqrt{\frac{\eta\ell}{d}}|E| \leqslant O(1)\frac{kp_t}{m^2}\sqrt{\eta\ell\left(\frac{2n}{r}\right)^{k-t}\sum_{i=1}^{p_t}\binom{|\mathcal{H}_i^{(t)}|}{2}} \\ &\leqslant O(1)\sqrt{\frac{\eta r \log n}{m\tau_t}\left(\frac{2n}{r}\right)^{k-t}}.\end{aligned}$$

Suppose $m \geqslant C^k n \log n \cdot (\frac{n}{r})^{\frac{k}{2}-1}\varepsilon^{-4}$ for some large enough constant C. We split into cases:

1. For $t \leqslant \frac{k-1}{2}$, we set $\tau_t = (\frac{n}{r})^{\frac{k}{2}-t} \cdot 4k\varepsilon^{-2}$, thus $\widehat{f}_t(x) \leqslant \frac{\varepsilon^2}{4}$.
2. For $t \geqslant \frac{k+1}{2}$, we set $\tau_t = 4k\varepsilon^{-2}$, thus $\widehat{f}_t(x) \leqslant \frac{\varepsilon^2}{4}(\frac{n}{r})^{\frac{k}{4}-\frac{t}{2}} < \frac{\varepsilon^2}{4}$.

Therefore, by calculating $\|\widetilde{A}_b\|_2$, which can be done in $n^{O(r)}$ time, we can certify that $\widehat{f}_t(x) \leqslant \frac{\varepsilon^2}{4}$. Combined with Eq. (7.6) and the bound of $\rho \leqslant 1/2$, we can certify that

$$f_t(x) \leqslant \frac{1}{1-\rho} \cdot \widehat{f}_t(x) \leqslant \frac{\varepsilon^2}{2},$$

and with Eq. (7.4), we can certify an upper bound on Eq. (7.3):

$$\psi_t(x)^2 \leqslant (7.4) + (7.5) \leqslant \frac{\varepsilon^2}{2} + f_t(x) \leqslant \varepsilon^2,$$

completing the proof. □

Chapter 8

Efficient Algorithms for Semirandom Planted CSPs

In this chapter, we prove Theorem 6.2.1. We first define the semirandom planted CSP model that we work with and explain some of the subtleties in the definition. Our model is the natural one that arises if we wish to enforce independent randomness (for each clause) in the literal negations, while still fixing a particular satisfying assignment.

Definition 8.0.1 (k-ary Boolean CSPs)**.** A CSP instance Ψ with a k-ary predicate $P\colon \{\pm 1\}^k \to \{0,1\}$ is a set of m constraints on variables $x_1, \ldots, x_n$ of the form

$$P\Big(\ell(\vec{C})_1 x_{\vec{C}_1},\ \ell(\vec{C})_2 x_{\vec{C}_2},\ \ldots,\ \ell(\vec{C})_k x_{\vec{C}_k}\Big) = 1\,.$$

Here, $\vec{C}$ ranges over a collection $\vec{\mathcal{H}}$ of *scopes*[1] (a.k.a. clause structure) of k-tuples of n variables and $\ell(\vec{C}) \in \{\pm 1\}^k$ are "literal negations", one for each $\vec{C}$ in $\vec{\mathcal{H}}$. We let $\mathrm{val}_\Psi(x)$ denote the fraction of constraints satisfied by an assignment $x \in \{\pm 1\}^n$, and we define the *value* of Ψ, $\mathrm{val}(\Psi)$, to be $\max_{x\in\{\pm 1\}^n} \mathrm{val}_\Psi(x)$.

Definition 8.0.2 (Semirandom planted k-ary Boolean CSPs)**.** Let $P\colon \{\pm 1\}^k \to \{0,1\}$ be a predicate. We say that a distribution Q over $\{\pm 1\}^k$ is a *planting distribution for* P if $\Pr_{y\leftarrow Q}[P(y) = 1] = 1$.

We say that an instance Ψ with predicate P is a *semirandom planted instance* with *planting distribution* Q if it is sampled from a distribution $\Psi(\vec{\mathcal{H}}, x^*, Q)$ where

(1) the scopes $\vec{\mathcal{H}} \subseteq [n]^k$ and planted assignment $x^* \in \{\pm 1\}^n$ are arbitrary, and

(2) $\Psi(\vec{\mathcal{H}}, x^*, Q)$ is defined as follows: for each $\vec{C} \in \vec{\mathcal{H}}$, sample literal negations $\ell(\vec{C}) \leftarrow Q(\ell(\vec{C}) \odot x^*_{\vec{C}})$, where "$\odot$" denotes the element-wise product of two vectors. That

[1] We additionally allow $\vec{\mathcal{H}}$ to be a multiset, i.e., that multiple clauses can contain the same ordered set of variables.

is, $\Pr[\ell(\vec{C}) = \ell] = Q(\ell \odot x^*_{\vec{C}})$ for each $\ell \in \{\pm 1\}^k$. Then, add the constraint $P(\ell(\vec{C})_1 x_{\vec{C}_1}, \ell(\vec{C})_2 x_{\vec{C}_2}, \ldots, \ell(\vec{C})_k x_{\vec{C}_k}) = 1$ to Ψ.

Notice that because Q is supported only on satisfying assignments to P, it follows that if $\Psi \leftarrow \Psi(\vec{\mathcal{H}}, x^*, Q)$, then x^* satisfies Ψ with probability 1.

A (fully) random planted CSP, e.g., as defined in [FPV15], is generated by first sampling $\vec{\mathcal{H}} \leftarrow [n]^k$ uniformly at random, and then sampling $\Psi \leftarrow \Psi(\vec{\mathcal{H}}, x^*, Q)$. The difference in the semirandom planted model is that we allow $\vec{\mathcal{H}}$ to be *worst case*.

Notice that in Definition 8.0.2, there are some choices of Q for which the planted instance becomes easy to solve. In the case of, e.g., 3-SAT, one could set the planting distribution Q to be uniform over all 7 satisfying assignments, which results in the literal negations in each clause being chosen uniformly conditioned on x^* satisfying the clause. However, by simply counting how many times the variable x_i appears negated versus not negated and taking the majority vote, we recover x^* with high probability [BHL+02, JMS07].

Unlike in the case of random planted CSPs, we cannot hope to recover the planted assignment x^* exactly in the semirandom setting. Indeed, the scopes $\vec{\mathcal{H}}$ may not use some variable x_i at all, and so we cannot hope to recover x^*_i! Thus, our goal is instead to recover an assignment x that has nontrivially large value, ideally value $1 - \varepsilon$ for arbitrarily small ε. Our main result, stated formally below, gives an algorithm to accomplish this task.

Theorem 8.0.3 (Formal Theorem 6.2.1)**.** *Let $k \in \mathbb{N}$ be constant. There is a polynomial-time algorithm that takes as input a k-CSP Ψ and outputs an assignment x with the following guarantee. If Ψ is a semirandom planted k-CSP with $m \geqslant c^k n^{k/2} \cdot \frac{\log^3 n}{\varepsilon^9}$ constraints drawn from $\Psi(\vec{\mathcal{H}}, x^*, Q)$, then with probability $1 - 1/\operatorname{poly}(n)$ over Ψ, the output x of the algorithm has $\operatorname{val}_\Psi(x) \geqslant 1 - \varepsilon$. Here, c is a universal constant.*

In particular, setting $\varepsilon = 1/\operatorname{polylog}(n)$, if $m \geqslant \widetilde{O}(n^{k/2})$, then with high probability over $\Psi \leftarrow \Psi(\vec{\mathcal{H}}, x^*, Q)$, the algorithm outputs x with $\operatorname{val}_\Psi(x) \geqslant 1 - o(1)$.

Similar to random planted CSPs [FPV15] and the refutation setting [AOW15, RRS17, AGK21, GKM22, HKM23], our proof of Theorem 8.0.3 goes through a reduction to noisy k-XOR. Here, we restate our result for solving noisy k-XOR.

Theorem (Restatement of Theorem 6.2.3)**.** *Let $\eta \in [0, 1/2)$, let $k, n \in \mathbb{N}$, and let $\varepsilon \in (0, 1)$. Let $m \geqslant cn^{k/2} \cdot \frac{k^4 \log^3 n}{\varepsilon^5 (1-2\eta)^4}$ for a universal constant c. There is a polynomial-time algorithm $\mathcal{A}$ that takes as input a k-XOR instance ψ with constraint hypergraph $\mathcal{H}$ and outputs two disjoint sets $\mathcal{A}_1(\mathcal{H}), \mathcal{A}_2(\psi) \subseteq \mathcal{H}$ with the following guarantees: (1) for any instance ψ with m constraints, $|\mathcal{A}_1(\mathcal{H})| \leqslant \varepsilon m$ and $\mathcal{A}_1(\mathcal{H})$ only depends on $\mathcal{H}$, and (2) for any $x^* \in \{\pm 1\}^n$ and any k-uniform hypergraph $\mathcal{H}$ with at least m hyperedges, with probability at least $1 - 1/\operatorname{poly}(n)$*

over $\psi \leftarrow \psi(\mathcal{H}, x^*, \eta)$*, it holds that* $\mathcal{A}_2(\psi) = \mathcal{E}_\psi \cap (\mathcal{H} \setminus \mathcal{A}_1(\mathcal{H}))$.

Organization. The rest of the chapter is organized as follows. First, in Section 8.1, we give an overview of our algorithm for noisy planted k-XOR. In Section 8.2, we prove Theorem 8.0.3 from Theorem 6.2.3 by reducing semirandom planted CSPs to noisy XOR. In Sections 8.3 and 8.4, we prove Theorem 6.2.3; Section 8.3 handles the reduction from k-XOR to "bipartite k-XOR", and then Section 8.4 gives the algorithm for the bipartite k-XOR case.

As we will see in Section 8.1, we will encounter various notions of relative graph approximations, including cut, SDP, and spectral approximation. These are discussed in Section 8.5, and we also show a separation between SDP and spectral approximation.

8.1 Technical overview

In this section, we give an overview of the proof of Theorem 6.2.3 and our algorithm for noisy planted k-XOR. We defer discussion of the reduction from general k-CSPs to k-XOR used to obtain Theorem 8.0.3 to Section 8.2. There, we explain the additional challenges encountered in the semirandom case as compared to the random case [FPV15, Section 4]. Somewhat surprisingly, the reduction is complicated and quite different from the random planted case or even the semirandom refutation setting, where the reduction to XOR is straightforward.

We now explain Theorem 6.2.3. As is typical in algorithm design for k-XOR, the case when k is even is considerably simpler than when k is odd. For the purpose of this overview, we will focus mostly on the even case, and only briefly discuss the additional techniques for odd k in Section 8.1.5.

Notation. Given a k-XOR instance ψ on hypergraph $\mathcal{H} \subseteq \binom{[n]}{k}$ with $m = |\mathcal{H}|$ and right-hand sides $\{b_C\}_{C \in \mathcal{H}}$, we define $\psi(x) := \sum_{C \in \mathcal{H}} b_C \prod_{i \in C} x_i$ to be a degree-k polynomial mapping $\{\pm 1\}^n \to [-m, m]$. We note that $\mathrm{val}_\psi(x) = \frac{1}{2} + \frac{1}{2m}\psi(x) \in [0, 1]$ is the fraction of constraints in ψ satisfied by x. Moreover, we will write $x_C := \prod_{i \in C} x_i$.

Unless otherwise stated, we will use ϕ to denote a 2-XOR instance and ψ to denote a k-XOR instance for any $k \geqslant 2$.

We note that for even arity k-XOR, we have $\mathrm{val}_\psi(x) = \mathrm{val}_\psi(-x)$, and so it is only possible for the optimal solution to be unique *up to a global sign*. We will abuse terminology and say that x^* is the unique optimal assignment if $\pm x^*$ are the only optimal assignments, and we will say that we have recovered x^* exactly if we obtain one of $\pm x^*$.

8.1.1 Approximate recovery for 2-XOR from refutation

First, let us focus on the case of $k = 2$, the simplest case, and let us furthermore suppose that we only want to achieve the weaker goal of recovering an assignment of value $1 - \eta - o(1)$. (Note that we do need the stronger guarantee of Theorem 6.2.3 to solve general planted CSPs in Theorem 8.0.3.)

For 2-XOR, this goal is actually quite straightforward to achieve using 2-XOR refutation as a blackbox. Let us represent the 2-XOR instance ϕ as a graph G on n vertices, along with right-hand sides b_{ij} for each edge $(i, j) \in E$. Recall that we have $b_{ij} = x^*_i x^*_j$ with probability $1 - \eta$, and $b_{ij} = -x^*_i x^*_j$ otherwise. Note that by concentration, $\mathrm{val}_\phi(x^*) = 1 - \eta \pm o(1)$ with high probability.

We now make the following observation. Let us suppose that we sample the noise in two steps: first, we add each $(i, j) \in E$ to a set E' with probability 2η independently; then for each $(i, j) \in E'$ we set b_{ij} to be uniformly random from $\{\pm 1\}$. Using known results for semirandom 2-XOR refutation, it is possible to certify, via an SDP relaxation, that no assignment x can satisfy (or violate) more than $\frac{1}{2} + o(1)$ fraction of the constraints in E'.

Thus, we can simply solve the SDP relaxation for ϕ and obtain a degree-2 pseudo-expectation $\widetilde{\mathbb{E}}$ in the variables $x_1, \dots, x_n$ over $\{\pm 1\}^n$ that maximizes $\phi(x)$. Let $\phi_{E'}$ be the subinstance containing only the constraints in E', and let $\phi_{E \setminus E'}$ be the subinstance containing only the constraints in $E \setminus E'$, which are uncorrupted. We have $\widetilde{\mathbb{E}}[\mathrm{val}_\phi(x)] \geqslant 1 - \eta - o(1)$, and the guarantee of refutation implies that $\widetilde{\mathbb{E}}[\mathrm{val}_{\phi_{E'}}(x)] \leqslant \frac{1}{2} + o(1)$. As $\mathrm{val}_\phi(x) = (1 - 2\eta) \cdot \mathrm{val}_{\phi_{E \setminus E'}}(x) + 2\eta \cdot \mathrm{val}_{\phi_{E'}}(x)$, we therefore have that $\widetilde{\mathbb{E}}[\mathrm{val}_{\phi_{E \setminus E'}}(x)] \geqslant 1 - o(1)$, i.e., $\widetilde{\mathbb{E}}$ satisfies $1 - o(1)$ fraction of the constraints in $E \setminus E'$. Then, applying the standard Gaussian rounding, we obtain an x that satisfies $1 - \sqrt{o(1)}$ fraction of the constraints in $E \setminus E'$ and thus has value $\mathrm{val}_\phi(x) \geqslant 1 - \eta - o(1)$ (as any x must satisfy at least $\frac{1}{2} - o(1)$ fraction of the constraints in E', with high probability over the noise).

One interesting observation is that in the above discussion, we can additionally allow E' to be an *arbitrary* subset of E of size $2\eta m$. Indeed, this is because the rounding only "remembers" that $\widetilde{\mathbb{E}}[\mathrm{val}_{\phi_{E \setminus E'}}(x)]$ has value $1 - o(1)$. As we shall see shortly, this is the key reason that the reduction breaks down for k-XOR.

8.1.2 The challenges for k-XOR and our strategy

Unfortunately, the natural blackbox reduction to refutation given in Section 8.1.1 does not generalize to k-XOR for $k \geqslant 3$. Following the approach described in the previous section, given a k-XOR instance ψ, one can solve a sum-of-squares SDP and obtain a pseudo-expectation $\widetilde{\mathbb{E}}$ where $\widetilde{\mathbb{E}}[\mathrm{val}_\psi(x)] \geqslant 1 - \eta - \delta$ and $\widetilde{\mathbb{E}}[\mathrm{val}_{\psi_{E \setminus E'}}(x)] \geqslant 1 - \delta$ as before, where $\delta \sim 1/\operatorname{polylog}(n)$ when $m \gtrsim n^{k/2}$, due to the guarantees of refutation algorithms [AGK21]. However, unlike 2-XOR where we have Gaussian rounding, for k-XOR there is no known rounding algorithm that takes a pseudo-expectation $\widetilde{\mathbb{E}}$ with

$\widetilde{\mathbb{E}}[\mathrm{val}_{\psi_{E\setminus E'}}(x)] \geqslant 1-\delta$ and outputs an assignment x such that $\mathrm{val}_{\psi_{E\setminus E'}}(x) \geqslant 1-f(\delta)$, for some $f(\cdot)$ such that $f(\delta) \to 0$ as $\delta \to 0$. In fact, if we only "remember" that $\psi_{E\setminus E'}$ has value $1-\delta$, then it is NP-hard to find an x with value $> 1/2+\delta$ even when $\delta = n^{-c}$ for some constant $c > 0$, assuming a variant of the Sliding Scale Conjecture [BGLR93][2] (see e.g. [MR10, Mos15] for more details).

As we have seen, while semirandom k-XOR refutation allows us to efficiently approximate and certify the *value* of the planted instance, the challenge lies in the *rounding* of the SDP, where the goal is to recover an assignment x. This is a technical challenge that does not arise in the context of CSP refutation, as there we are merely trying to bound the value of the instance. As a result, new ideas are required to address this challenge.

Reduction from k-XOR to 2-XOR for even k. One could still consider the following natural approach. For simplicity, let $k = 4$. Given a 4-XOR instance ψ, we can write down a natural and related 2-XOR instance ϕ, as follows.

Definition 8.1.1 (Reduction to 2-XOR)**.** Let ψ be a 4-XOR instance, and let ϕ be the 2-XOR defined as follows. The variables of ϕ are $y_{\{i,j\}}$ and correspond to *pairs* of variables $\{x_i, x_j\}$, and for each constraint $x_i x_j x_{i'} x_{j'} = b_{i,j,i',j'}$ in ψ, we split $\{i,j,i',j'\}$ into $\{i,j\}$ and $\{i',j'\}$ arbitrarily and add a constraint $y_{\{i,j\}} y_{\{i',j'\}} = b_{i,j,i',j'}$ to ϕ. See Fig. 8.1 for an example. This reduction easily generalizes to k-XOR for any even k.

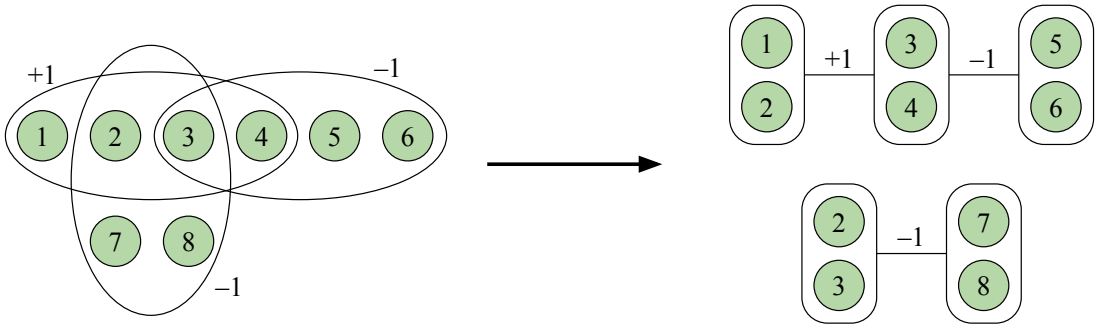


Figure 8.1: An example of the 2-XOR instance ϕ from a 4-XOR instance ψ.

By following the approach for 2-XOR described in Section 8.1.1, we can recover an assignment y that satisfies $1-\eta-o(1)$ fraction of the constraints in ϕ. However, we need to recover an assignment x to the original k-XOR ψ, and it is quite possible that while y is a good assignment to ϕ, *it is not* close to $x^{\otimes 2}$ for *any* $x \in \{\pm 1\}^n$. If this happens, we will be unable to recover a good assignment to the 4-XOR instance ψ.

[2]Note that we do need the Sliding Scale Conjecture, as the hardness shown in [MR10] is not strong enough; it only proves hardness for $\delta \geqslant (\log\log n)^{-c}$, whereas we have $\delta \sim 1/\operatorname{polylog}(n)$.

The key reason that this simple idea fails is because, unlike for random noisy XOR, the assignment y recovered is *not* necessarily unique, and we cannot hope for it to be in the semirandom setting! For random noisy XOR, one can argue that with high probability, y will be equal to $x^{*\otimes 2}$, and then we can immediately decode and recover x^* up to a global sign, i.e., we recover $\pm x^*$. But for semirandom instances, the situation can be far more complex.

Approximate 2-XOR recovery does not suffice for 4-XOR. When constructing the 2-XOR instance ϕ from the 4-XOR ψ (Definition 8.1.1), it may be the case that ϕ can be partitioned into multiple disconnected clusters (or have very few edges across different clusters), even when the hypergraph $\mathcal{H}$ of ψ is connected; see Fig. 8.1 for example. By the algorithm described in Section 8.1.1, we can get an assignment y that satisfies $1 - \eta - o(1)$ fraction of the constraints within each cluster.

The main challenge is to combine the information gathered from each cluster to recover an assignment x for the original 4-XOR ψ. Unfortunately, we do not know of a way to obtain a good assignment x based solely on the guarantee that y satisfies $1 - \eta - o(1)$ fraction of constraints in each cluster. The issue occurs because the same variable $i \in [n]$ can appear in different clusters, e.g., $y_{\{1,2\}}$ and $y_{\{2,3\}}$ lie in different clusters in Fig. 8.1, and the recovered assignments in each cluster may implicitly choose different values for x_i because of the noise. Indeed, even if the local optimum is consistent with x^*, there can still be multiple "good" assignments that achieve $1 - \eta - o(1)$ value on the subinstance restricted to a cluster. So, unless the SDP can certify unique optimality of x^*, standard rounding techniques such as Gaussian rounding will merely output a "good" y, which may be inconsistent with x^* and thus can choose inconsistent values of x_i across the different clusters.

Exact 2-XOR recovery implies exact 4-XOR recovery. This leads to our main insight: if the subinstance of ϕ admits a *unique* local optimal assignment y^* (restricted to the cluster) that matches the planted assignment up to a sign, i.e., $y^*_{\{i,j\}} = \pm x^*_i x^*_j$, then for each edge in the cluster we know $y^*_{\{i,j\}} y^*_{\{i',j'\}} = x^*_i x^*_j x^*_{i'} x^*_{j'}$, and so the local constraints that are violated must be exactly the corrupted ones. Moreover, if the SDP can certify the uniqueness of the local optimal assignment for a cluster, then the SDP solution will be a *rank* 1 *matrix* $y^* y^{*\top}$, and so we can precisely identify which constraints in ϕ are corrupted. By repeating this for every cluster, we can identify all corrupted constraints in the original 4-XOR ψ (except for the small number of "cross cluster" edges), and thus achieve the guarantee stated in Theorem 6.2.3.

The general algorithmic strategy. The above discussion suggests that given a k-XOR instance ψ, we should first construct the 2-XOR ϕ, and then decompose the constraint graph G of ϕ into pieces in some particular way so that the induced local instances have unique solutions. Namely, the examples suggest the following algorithmic strategy.

Strategy 1 (Algorithm Blueprint for even k)**.** Given a noisy k-XOR instance ψ with planted assignment x^* and m constraints, we do the following:

(1) Construct the 2-XOR instance ϕ described in Definition 8.1.1, which is a noisy 2-XOR on $n^{k/2}$ variables with planted assignment y^*. Moreover, there is a one-to-one mapping between constraints in ϕ and ψ.

(2) Let G be the constraint graph of ϕ. Decompose G into subgraphs $G_1, \dots, G_T$ while only discarding a $o(1)$-fraction of edges such that each subgraph G_i satisfies "some property". For each subgraph G_i, we define ϕ_i to be the subinstance of ϕ corresponding to the constraints in G_i. The goal is to identify a local property that the G_i's satisfy so that (1) we can perform the decomposition efficiently, and (2) for each subinstance ϕ_i, we can "recover y^* locally", i.e., we can find an assignment $y^{(i)}$ to the 2-XOR instance ϕ_i that is consistent with the planted assignment y^*.

(3) As each $y^{(i)}$ is consistent with y^*, the constraints in ϕ_i violated by $y^{(i)}$ must be precisely the corrupted constraints in ϕ_i. Hence, for the constraints that appear in one of the ϕ_i's, we have determined exactly which ones are corrupted.

(4) We have thus determined, for all but $o(m)$ constraints, precisely which ones are corrupted in the original k-XOR instance ψ. (Note that this is the *stronger* guarantee that we achieve in Theorem 6.2.3.) By discarding the corrupted constraints along with the $o(m)$ constraints where we "give up", we thus obtain a system of k-sparse linear equations with $m(1 - \eta - o(1))$ equations that has at least one solution (namely x^*), and so by solving it we obtain an x with $\mathrm{val}_\psi(x) \geqslant 1 - \eta - o(1)$.

8.1.3 Information-theoretic exact recovery from relative cut approximation

Following Strategy 1, the first technical question to now ask is: given a noisy 2-XOR instance ϕ with n variables, $m \gg n$ constraints, and planted assignment x^*, what conditions do we need to impose on the constraint graph G so that we can recover x^* (up to a sign) exactly? As a natural first step, we investigate what conditions are required so that we can accomplish this *information-theoretically*.

Fact 8.1.2. *Let $G = (V, E_G)$ be an n-vertex graph, and let $H = (V, E_H)$ be a subgraph of G where $E_H \subseteq E_G$. Let L_G, L_H be the* unnormalized *Laplacians of G and H. Consider a noisy planted 2-XOR instance ϕ on G with planted assignment $x^* \in \{\pm 1\}^n$ (Definition 6.2.2), and suppose E_H is the set of corrupted edges. Suppose that for every $x \in \{\pm 1\}^n \setminus \{\vec{1}, -\vec{1}\}$, it holds that $x^\top L_H x < \frac{1}{2} x^\top L_G x$. Then, x^* and $-x^*$ are the only two optimal assignments to ϕ.*

Note that the condition $x^\top L_H x < \frac{1}{2} x^\top L_G x$ for $x \notin \{\vec{1}, -\vec{1}\}$ implies that G is connected, as otherwise L_G has a kernel of dimension $\geqslant 2$, which would contradict this assumption.

Proof. Let $x \in \{\pm 1\}^n$ be any assignment. We wish to show that $\phi(x)$ is uniquely maximized when $x = x^*, -x^*$. We observe that

$$\phi(x) = \sum_{(i,j)\in E_G} x_i x_j b_{ij} = \sum_{(i,j)\in E_G} x_i x_j x_i^* x_j^* - 2 \sum_{(i,j)\in E_H} x_i x_j x_i^* x_j^* \ .$$

Hence, by replacing x with $x \odot x^*$, without loss of generality we can assume that $x^* = \vec{1}$. Now, let D_G, D_H and A_G, A_H be the degree and adjacency matrices of G and H, so that $L_G = D_G - A_G$ and $L_H = D_H - A_H$. We thus have that

$$\begin{aligned} 2\phi(x) &= x^\top A_G x - 2x^\top A_H x = x^\top (D_G - 2D_H)x - x^\top (L_G - 2L_H)x \\ &= 2(|E_G| - 2|E_H|) - x^\top (L_G - 2L_H)x \ . \end{aligned}$$

By assumption, if $x \in \{\pm 1\}^n$ and $x \neq \vec{1}, -\vec{1}$, then we have that $x^\top (L_G - 2L_H)x > 0$, which implies that $\phi(x) < \phi(\vec{1})$, and finishes the proof. □

Fact 8.1.2 shows that if we can argue that $x^\top L_H x < \frac{1}{2} x^\top L_G x$ for every $x \in \{\pm 1\}^n \setminus \{\vec{1}, -\vec{1}\}$, then at least information-theoretically we can uniquely determine x^*. Observe that if we view x as the signed indicator vector of a subset $S \subseteq [n]$, then $x^\top L_G x = E_G(S, \bar{S})$, the number of edges in G crossing the cut defined by S, and similarly for $x^\top L_H x$. So, one can view the condition in Fact 8.1.2 as saying that the subgraph H needs to be a (one-sided) cut sparsifier of G, i.e., it needs to roughly preserve the size of all cuts in G. The following relative cut approximation result of Karger [Kar94] shows that this will hold with high probability when H is a randomly chosen subset of G, provided that the minimum cut in G is not too small.

Lemma 8.1.3 (Relative cut approximation [Kar94])**.** *Let $\eta \in (0, 1)$. Suppose an n-vertex graph G has min-cut $c_{\min} \geqslant \frac{12 \log n}{\eta}$, and suppose H is a subgraph of G by selecting each edge with probability η. Then, with probability $1 - o(1)$,*

$$(1 - \delta) x^\top L_G x \leqslant \frac{1}{\eta} \cdot x^\top L_H x \leqslant (1 + \delta) x^\top L_G x \,, \quad \textit{for all } x \in \{\pm 1\}^n$$

for $\delta = \sqrt{\frac{12 \log n}{\eta c_{\min}}}$.

With Lemma 8.1.3 and Fact 8.1.2 in hand, we now have at least an information-theoretic algorithm with the same guarantees as in Theorem 6.2.3. We follow the strategy highlighted in Strategy 1. To decompose the graph G, we recursively find a min cut and split if it is below the threshold in Lemma 8.1.3. Notice that this discards at most

$O(n \log n) = o(m)$ constraints (for $m \gg n \log n$), and these are precisely the constraints that we "give up" on and do not determine which ones are corrupted. Then, with high probability the local optimal assignment is consistent with x^*, and so locally we have learned *exactly* which constraints are corrupted. Hence, we have produced two sets of constraints: E_1, the $o(1)$-fraction of edges discarded during the decomposition, and $E_2 = (G \setminus E_1) \cap \mathcal{E}_\phi$, which is exactly the set of corrupted constraints after discarding E_1. We note that it is a priori not obvious that this is achievable even for an *exponential-time* algorithm, as even though the 2^n-time brute force algorithm will find the best assignment x to ϕ, it may not necessarily be x^*, and so the set of constraints violated by the globally optimal assignment might not be $\mathcal{E}_\phi$.

8.1.4 Efficient exact recovery from relative spectral approximation

Information-theoretic uniqueness implies that the planted assignment x^* is the unique optimal assignment. But can we efficiently recover x^*? One natural approach is to simply solve the basic SDP relaxation of ϕ: for $X \in \mathbb{R}^{n \times n}$, maximize $\phi(X) := \sum_{(i,j) \in G} X_{ij} b_{ij}$ subject to $X \succeq 0$, $X = X^\top$, and $\mathrm{diag}(X) = \mathbb{I}$. If the optimal SDP solution is simply $X = x^* x^{*\top}$, then we trivially recover x^* from the SDP solution. We thus ask: does the min cut condition of Fact 8.1.2 and Lemma 8.1.3 imply that $x^* x^{*\top}$ is the unique optimal solution to the SDP? Namely, is the min cut condition sufficient for the SDP to certify that x^* is the unique optimal assignment?

Unfortunately, it turns out that this is not the case, and we give a counterexample in Section 8.5. We thus require a stronger condition than the min cut one in order to obtain efficient algorithms. Nonetheless, an analogue of Fact 8.1.2 continues to hold, although now we require a stronger version that holds for all SDP solutions X, not just $x \in \{\pm 1\}^n$. This stronger statement shows the SDP can *certify* that x^* is the unique optimal assignment if and only if a certain relative spectral approximation guarantee holds for the corrupted edges.

Lemma 8.1.4 (SDP-certified uniqueness from relative spectral approximation)**.** *Let $G = (V, E_G)$ be an n-vertex connected graph, and let $H = (V, E_H)$ be a subgraph of G where $E_H \subseteq E_G$. Let L_G, L_H be the unnormalized Laplacians of G and H. Consider a noisy planted 2-XOR instance ϕ on G with planted assignment $x^* \in \{\pm 1\}^n$ (Definition 6.2.2), and suppose E_H is the set of corrupted edges.*

The SDP relaxation of ϕ satisfies

$$\max_{X \succeq 0,\ X = X^\top,\ \mathrm{diag}(X) = \mathbb{I}} \phi(X) = \phi(x^*) = |E_G| - 2|E_H|\,,$$

where $X = x^* x^{*\top}$ is the unique *optimum if and only if G and H satisfy*

$$\langle X, L_H \rangle < \frac{1}{2} \langle X, L_G \rangle\,, \quad \forall X \succeq 0,\ X = X^\top,\ \mathrm{diag}(X) = \mathbb{I},\ X \neq \vec{1}\vec{1}^\top\,.$$

Proof. Recall that each $e = \{i, j\} \in E$ corresponds to a constraint $x_i x_j = b_e$ where $b_e = x_i^* x_j^*$ if $e \in E_G \setminus E_H$ and $b_e = -x_i^* x_j^*$ if $e \in E_H$, meaning that $\phi(X) = \sum_{\{i,j\} \in G \setminus E} X_{ij} x_i^* x_j^* - \sum_{\{i,j\} \in E} X_{ij} x_i^* x_j^*$. Without loss of generality, we can assume that $x^* = \vec{1}$ and that $\phi(X) = \frac{1}{2}\langle X, A_G - 2A_H \rangle$, where A_G, A_H are the adjacency matrices of G and H.

Note that $L_G = D_G - A_G$ and $L_H = D_H - A_H$, and $\mathrm{tr}(D_G) = 2|E_G|$, $\mathrm{tr}(D_H) = 2|E_H|$. For any $X \succeq 0$ with $\mathrm{diag}(X) = \mathbb{I}$,

$$\langle X, A_G - 2A_H \rangle = \langle X, (D_G - L_G) - 2(D_H - L_H) \rangle = 2(|E_G| - 2|E_H|) + \langle X, 2L_H - L_G \rangle \ .$$

Suppose $\langle X, L_H \rangle < \frac{1}{2}\langle X, L_G \rangle$ for all $X \neq \vec{1}\vec{1}^\top$. Since $\langle \vec{1}\vec{1}^\top, L_G \rangle = \langle \vec{1}\vec{1}^\top, L_H \rangle = 0$, we have that the maximum of $\frac{1}{2}\langle X, A_G - 2A_H \rangle$ is $|E_G| - 2|E_H|$ and $X = \vec{1}\vec{1}^\top$ is the unique maximum.

For the other direction, suppose there is an $X \neq \vec{1}\vec{1}^\top$ such that $\langle X, L_H \rangle \geqslant \frac{1}{2}\langle X, L_G \rangle$. Then, $\phi(X) \geqslant |E_G| - 2|E_H| = \phi(\vec{1}\vec{1}^\top)$, meaning that $\vec{1}\vec{1}^\top$ is not the unique optimum. □

Relative spectral approximation from uniform subsamples. We now come to a key technical observation. Suppose that H is a *spectral sparsifier* of G, so that $v^\top(\frac{1}{\eta}L_H)v$ is $(1 \pm \delta)v^\top L_G v$ for any $v \in \mathbb{R}^n$. Then clearly $\langle X, L_H \rangle < \frac{1}{2}\langle X, L_G \rangle$ if $\eta < 1/2$ and $\delta = o(1)$, as we can write $X = \sum_{i=1}^n \lambda_i v_i v_i^\top$, and

$$\langle X, L_H \rangle = \sum_{i=1}^n \lambda_i v_i^\top L_H v_i \leqslant \eta(1+\delta) \sum_{i=1}^n \lambda_i v_i^\top L_G v_i = \eta(1+\delta) \cdot \langle X, L_G \rangle < \frac{1}{2}\langle X, L_G \rangle \, .$$

Furthermore, note that above we only required that $L_H \preceq \eta(1+\delta)L_G$, i.e., we only use the upper part of the spectral approximation.

We are now ready to state the key relative spectral approximation lemma. We observe that when H is a uniformly random subsample of G and G has a *spectral gap* and minimum degree polylog(n), then with high probability $L_H \preceq \eta(1+\delta)L_G$. We note that, while we do not provide a formal proof, the same argument using the lower tail of Matrix Chernoff can also establish a lower bound on L_H, which proves that H is indeed a spectral sparsifier of G.

Lemma 8.1.5 (Relative spectral approximation from uniform subsamples)**.** *Let $\eta \in (0, 1)$. Suppose $G = (V, E)$ is an n-vertex graph with minimum degree $d_{\min}$ (self-loops allowed) and spectral gap $\lambda_2(\widetilde{L}_G) = \lambda$ such that $d_{\min}\lambda > \frac{18}{\eta}\log n$, where $\widetilde{L}_G := D_G^{-1/2} L_G D_G^{-1/2}$ is the* normalized *Laplacian. Let H be a subgraph of G obtained by selecting each edge with probability η. Then, with probability at least $1 - O(n^{-2})$,*

$$L_H \preceq \eta(1+\delta) \cdot L_G$$

for $\delta = \sqrt{\frac{18 \log n}{\eta d_{\min} \lambda}}$.

Proof. First, note that $\vec{1}$ lies in the kernel of both L_G and L_H, and because of the spectral gap of G, $\dim(\ker(L_G)) = 1$. Therefore, recalling that $L_G = D_G^{1/2} \widetilde{L}_G D_G^{1/2}$, it suffices to prove that

$$\left\| (\widetilde{L}_G^\dagger)^{1/2} D_G^{-1/2} L_H D_G^{-1/2} (\widetilde{L}_G^\dagger)^{1/2} \right\|_2 \leqslant \eta(1+\delta)\,.$$

Here $\widetilde{L}_G^\dagger$ is the pseudo-inverse of $\widetilde{L}_G$, and $\|\widetilde{L}_G^\dagger\|_2 \leqslant 1/\lambda$ because G has spectral gap λ. We will write $X := (\widetilde{L}_G^\dagger)^{1/2} D_G^{-1/2} L_H D_G^{-1/2} (\widetilde{L}_G^\dagger)^{1/2}$ for convenience.

Note that $L_G = \sum_{e \in E} L_e$, where $L_e \succeq 0$ is the Laplacian of a single edge e and $\|L_e\|_2 = 2$. Let $X_e = (\widetilde{L}_G^\dagger)^{1/2} D_G^{-1/2} L_e D_G^{-1/2} (\widetilde{L}_G^\dagger)^{1/2}$ if e is chosen in H and 0 otherwise. Then, $X = \sum_{e \in E} X_e$ and $\|\mathbb{E}[X]\|_2 = \eta$. Moreover, each X_e satisfies $X_e \succeq 0$ and $\|X_e\|_2 \leqslant \|\widetilde{L}_G^\dagger\|_2 \cdot \|D_G^{-1}\|_2 \cdot \|L_e\|_2 \leqslant \frac{2}{d_{\min}\lambda}$. Thus, by Matrix Chernoff (Fact 2.4.2),

$$\Pr\left[\|X\|_2 \geqslant \eta(1+\delta)\right] \leqslant n \cdot \exp\left(-\frac{\delta^2 \eta}{3} \cdot \frac{d_{\min}\lambda}{2}\right) \leqslant O(n^{-2})$$

as long as $\frac{18 \log n}{\eta d_{\min}\lambda} \leqslant \delta^2 \leqslant 1$. □

Finishing the algorithm. By Lemmas 8.1.4 and 8.1.5, we can thus recover x^* exactly if the constraint graph G of ϕ has a nontrivial spectral gap and minimum degree $d_{\min} \geqslant \mathrm{polylog}(n)$. To finish the implementation of Strategy 1, we thus need to explain how to algorithmically decompose any graph G into subgraphs $G_1, \ldots, G_T$, each with reasonable min degree and nontrivial spectral gap, while only discarding a $o(1)$-fraction of the edges in G. This is the well-studied task of expander decomposition, for which we appeal to known results [KVV04, ST11, Wul17, SW19].

This completes the high-level description of the algorithm in the even k case. Below, we summarize the steps of the final algorithm.

Algorithm 8.1.6 (Algorithm for k-XOR for even k)**.**

Input: k-XOR instance ψ on n variables with m constraints and constraint hypergraph $\mathcal{H}$.

Output: Disjoint sets of constraints $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathcal{H}$ such that $|\mathcal{A}_1| \leqslant o(m)$ and only depends on $\mathcal{H}$, and $\mathcal{A}_2 = (\mathcal{H} \setminus \mathcal{A}_1) \cap \mathcal{E}_\psi$.

Operation:

1. Construct the 2-XOR instance ϕ with constraint graph G, as described in Definition 8.1.1.
2. Remove small-degree vertices and run expander decomposition on G to produce expanders $G_1, \ldots, G_T$. Set $\mathcal{A}_1$ to be the set of discarded constraints of size $o(m)$.
3. For each $i \in [T]$, solve the basic SDP on the subinstance ϕ_i defined by

the constraints G_i. Let $\mathcal{A}_2^{(i)}$ denote the set of constraints violated by the optimal local SDP solution.

4. Output $\mathcal{A}_1$ and $\mathcal{A}_2 = \bigcup_{i=1}^{T} \mathcal{A}_2^{(i)}$.

8.1.5 The case of odd k

We are now ready to briefly explain the differences in the case when k is odd. For the purposes of this overview, we will focus only on the case of $k = 3$. Recall that we are given a 3-XOR instance ψ, specified by a 3-uniform hypergraph $\mathcal{H} \subseteq \binom{[n]}{3}$, as well as the right-hand sides $b_C \in \{\pm 1\}$ for $C \in \mathcal{H}$, where $b_C = x^*_C$ with probability $1 - \eta$ and $b_C = -x^*_C$ otherwise and $x^* \in \{\pm 1\}^n$ is the planted assignment.

We now produce a 4-XOR instance using the well-known "Cauchy-Schwarz trick" from CSP refutation [CGL07]. The general idea is to, for any pair of clauses (C, C') that intersect, add the "derived constraint" $x_C x_{C'} = b_C b_{C'}$ to the 4-XOR instance. Notice that if, e.g., $C = \{u, i, j\}$ and $C' = \{u, i', j'\}$, then x_u appears twice on the left-hand side, and thus the constraint is $x_i x_j x_{i'} x_{j'} = b_C b_{C'}$. Given this 4-XOR, we produce a 2-XOR following a similar strategy as in Definition 8.1.1. The above description omits many technical details, which we handle in Sections 8.3 and 8.4; we remark here that these are the same issues that arise in the CSP refutation case, and we handle them using the techniques in [GKM22].

We have thus produced a 2-XOR instance ϕ that is noisy but not in the sense of Definition 6.2.2. Indeed, each edge e in ϕ is "labeled" by a pair (C, C') of constraints in ψ, and e is noisy if and only if *exactly* one of (C, C') is, and so the noise is not independent across constraints. Nonetheless, we can still follow the general strategy as in Algorithm 8.1.6. The main technical challenge is to argue that the relative spectral approximation guarantee of Lemma 8.1.5 holds even when the noise has the aforementioned correlations, and we do this in Lemma 8.4.7. This allows us to recover, for most intersecting pairs (C, C'), the quantity $\xi(C)\xi(C')$, where $\xi(C) = -1$ if C is corrupted, and is 1 otherwise, i.e., $b_C = x^*_C \xi(C)$; we do not determine $\xi(C)\xi(C')$ if and only if the pair (C, C') corresponds to an edge e that was discarded during the expander decomposition.

However, we are not quite done, as we would like to recover $\xi(C)$ for most C, but we only know $\xi(C)\xi(C')$ for most intersecting pairs (C, C'). Let us proceed by assuming that we know $\xi(C)\xi(C')$ for all intersecting pairs (C, C'), and then we will explain how to do a similar decoding process when we only know most pairs. Let us fix a vertex u, and let $\mathcal{H}_u$ denote the set of $C \in \mathcal{H}$ containing u. Now, we know $\xi(C)\xi(C')$ for all $C, C' \in \mathcal{H}_u$, and so by Gaussian elimination we can determine $\xi(C)$ for all $C \in \mathcal{H}_u$ up to a global sign. Now, we know that the vector $\{\xi(C)\}_{C \in \mathcal{H}_u}$ should have roughly $\eta|\mathcal{H}_u|$ entries that are -1. So, choosing the global sign that results in fewer -1's, we thus correctly determine $\xi(C)$ for all $C \in \mathcal{H}_u$. We can then repeat this process for each choice

of u to decode $\xi(C)$ for all C.

Of course, we only actually know $\xi(C)\xi(C')$ for most intersecting pairs (C, C'). This implies that for most choices of u, the graph G_u with vertices $\mathcal{H}_u$ and edges (C, C') if we know $\xi(C)\xi(C')$ is obtained from the complete graph on vertices $\mathcal{H}_u$ and deleting some $o(1)$-fraction of edges. This implies that G_u has a connected component of size $(1 - o(1))|\mathcal{H}_u|$, and again via Gaussian elimination and picking the proper global sign, we can determine $\xi(C)$ on this large connected component. By repeating this process for each choice of u, we thus recover $\xi(C)$ for most u.

8.2 From planted CSPs to noisy XOR

In this section, we show how to use Theorem 6.2.3 to prove Theorem 8.0.3. Before we delve into the formal proof, we will first explain the reduction given in [FPV15]. We begin with some definitions.

Setup. Let Ψ be sampled from $\Psi(\vec{\mathcal{H}}, x^*, Q)$, where $x^* \in \{\pm 1\}^n$, $\vec{\mathcal{H}} \subseteq [n]^k$, and Q is a planting distribution for the predicate P. Let $Q(y) = \sum_{S \subseteq [k]} \hat{Q}(S) \prod_{i \in S} y_i$ be the Fourier decomposition of Q, where $\hat{Q}(S) = \frac{1}{2^k} \sum_{y \in \{\pm 1\}^k} Q(y) \prod_{i \in S} y_i \in [-2^{-k}, 2^{-k}]$. Recall (Definition 8.0.2) that Ψ is specified by a collection $\vec{\mathcal{H}} \subseteq [n]^k$ of scopes, along with a vector $\ell(\vec{C}) \in \{\pm 1\}^k$ for each $\vec{C} \in \vec{\mathcal{H}}$ of literal negations.

Definition 8.2.1. Let $S \subseteq [k]$ be nonempty. Let $\psi^{(S,+)}$ be the $|S|$-XOR instance obtained by, for each constraint $\vec{C}$ in Ψ, adding the constraint $\prod_{i \in S} x_{\vec{C}_i} = \prod_{i \in S} \ell(\vec{C})_i$. Similarly, let $\psi^{(S,-)}$ have constraints $\prod_{i \in S} x_{\vec{C}_i} = -\prod_{i \in S} \ell(\vec{C})_i$.

We make use of the following simple claim.

Claim 8.2.2. *For each nonempty $S \subseteq [k]$, $\psi^{(S,+)}$ is a noisy $|S|$-XOR instance (Definition 6.2.2) with planted assignment x^* and noise $\eta = \frac{1}{2}(1 - 2^k \hat{Q}(S))$. Similarly, $\psi^{(S,-)}$ is a noisy $|S|$-XOR instance with planted assignment x^* and noise $\eta = \frac{1}{2}(1 + 2^k \hat{Q}(S))$.*

Proof. For each $\vec{C}$, the literal negation $\ell(\vec{C})$ is sampled such that $\Pr[\ell(\vec{C}) = \ell] = Q(\ell \odot x^*_{\vec{C}})$, where $\odot$ denotes the element-wise product. This is equivalent to sampling $y \leftarrow Q$ and setting $\ell(\vec{C}) = y \odot x^*_{\vec{C}}$. It thus follows that the probability that the constraint $\vec{C}$ produces a corrupted constraint in $\psi^{(S,+)}$ is

$$\Pr_{y \leftarrow Q}\left[\prod_{i \in S} y_i = -1\right] = \frac{1}{2}\left(1 - \mathbb{E}_{y \leftarrow Q}\left[\prod_{i \in S} y_i\right]\right) = \frac{1}{2}(1 - 2^k \hat{Q}(S)) \ ,$$

and is independent for each $\vec{C}$. A similar calculation handles the case of $\psi^{(S,-)}$. □

With the above observations in hand, we can now describe the reduction in [FPV15]. First, their reduction requires the algorithm to have a description of the distribution Q. Given Q, the algorithm then finds the smallest S such that $\hat{Q}(S)$ is nonzero. Since they know the exact value of $\hat{Q}(S)$, they can determine its sign correctly. Suppose that $\hat{Q}(S) > 0$ (the other case is similar). Then, by solving the $|S|$-XOR instance $\psi^{(S,+)}$, they recover the planted assignment of $\psi^{(S,+)}$ *exactly*.[3] But this planted assignment is precisely x^*, and so they have also recovered the planted assignment of ψ.

The aforementioned reduction clearly does not generalize to the semirandom setting, as in general the subinstances $\psi^{(S,\pm)}$ will not uniquely determine x^*. Furthermore, their reduction additionally requires knowing Q, and while it is not too unreasonable to assume this for random planted CSPs (as it is perhaps natural for the algorithm to know the distribution), in the semirandom setting this assumption is a bit strange because we want to view semirandom CSPs as "moving towards" worst case ones.

We now prove Theorem 8.0.3 from Theorem 6.2.3.

Proof of Theorem 8.0.3 from Theorem 6.2.3. We will present the proof in three steps. First, like [FPV15], we will assume that the algorithm is given a description of Q and we will assume that each $|\hat{Q}(S)|$ is either 0 or at least $2^{-k}\varepsilon > 0$.[4] Then, we will remove this assumption provided that $Q(y) > 2\varepsilon$ for all y with $Q(y) > 0$, i.e., the every y in the support of Q has some minimum probability. Finally, we will remove the last assumption.

Step 1: the proof when we are given Q. For each S where $\hat{Q}(S) \neq 0$, we construct the instance $\psi^{(S,+)}$ (if $\hat{Q}(S) > 0$) or $\psi^{(S,-)}$ (if $\hat{Q}(S) < 0$). We then apply[5] Theorem 6.2.3 to each such instance. Note that by Claim 8.2.2, the instance has noise $\eta = \frac{1}{2}(1 - 2^k|\hat{Q}(S)|) \leqslant \frac{1}{2}(1-\varepsilon)$ (because we picked the correct sign when choosing between $\psi^{(S,+)}$ and $\psi^{(S,-)}$, and we assume $|\hat{Q}(S)| \geqslant 2^{-k}\varepsilon$). Then, since $m \geqslant c^k n^{k/2} \cdot \frac{\log^3 n}{\varepsilon^9}$ and $|S| \leqslant k$, by applying Theorem 6.2.3 with noise η and parameter $\varepsilon' := 2^{-k}\varepsilon$, we obtain sets $\vec{\mathcal{H}}^{(S,1)}$ (the discarded set) and $\vec{\mathcal{H}}^{(S,2)}$ (the corrupted constraints) where $|\vec{\mathcal{H}}^{(S,1)}| \leqslant \varepsilon' m$ and $\vec{\mathcal{H}}^{(S,2)} = (\vec{\mathcal{H}} \setminus \vec{\mathcal{H}}^{(S,1)}) \cap \mathcal{E}_{\psi^{(S)}}$. Hence, for every constraint $\vec{C} \in \vec{\mathcal{H}} \setminus \vec{\mathcal{H}}^{(S,1)}$, it follows that we have learned $\prod_{i\in S} x^*_{\vec{C}_i}$, where x^* is the planted assignment for Ψ. By setting $\vec{\mathcal{H}}' := \vec{\mathcal{H}} \setminus \cup_{S:\hat{Q}(S)\neq 0} \vec{\mathcal{H}}^{(S,1)}$, it follows that we know $\prod_{i\in S} x^*_{\vec{C}_i}$ for all $\vec{C} \in \vec{\mathcal{H}}'$ and S with $\hat{Q}(S) \neq 0$, where $|\vec{\mathcal{H}}'| \geqslant (1 - 2^k\varepsilon')m = (1-\varepsilon)m$.

We now solve the system of linear equations given by $\prod_{i\in S} x^*_{\vec{C}_i}$ for all $\vec{C} \in \vec{\mathcal{H}}'$ and S with $\hat{Q}(S) \neq 0$ to obtain some assignment $x \in \{\pm 1\}^n$. As x^* is a valid solution to these equations, such an x exists, although it may not be x^*.

[3] Here, they also treat $|\hat{Q}(S)|$ as constant, since if $|\hat{Q}(S)| \ll 1/n$, say, then their algorithm would not succeed in recovering the planted assignment on the XOR instance.

[4] This assumption is implicit in [FPV15]; see the previous footnote.

[5] Note that Theorem 6.2.3 only applies when $|S| \geqslant 2$. When $|S| = 1$, there is a trivial algorithm.

The final step is to argue that for every $\vec{C} \in \vec{\mathcal{H}}'$, x satisfies the constraint $\vec{C}$, namely that $P(\ell(\vec{C})_1 x_{\vec{C}_1}, \ell(\vec{C})_2 x_{\vec{C}_2}, \ldots, \ell(\vec{C})_k x_{\vec{C}_k}) = 1$. Indeed, if this is true then we are done, as x satisfies at least $(1-\varepsilon)m$ constraints in Ψ, and so we have obtained the desired assignment.

Let $\vec{C} \in \vec{\mathcal{H}}'$. We know that for every S with $\hat{Q}(S) \neq 0$, we have that $\prod_{i \in S} x_{\vec{C}_i} = \prod_{i \in S} x^*_{\vec{C}_i}$. Hence, it follows that

$$Q(\ell(\vec{C}) \odot x) = \sum_{S \subseteq [k]} \hat{Q}(S) \prod_{i \in S} \ell(\vec{C})_i x_{\vec{C}_i} = \sum_{S \subseteq [k]} \hat{Q}(S) \prod_{i \in S} \ell(\vec{C})_i x^*_{\vec{C}_i} = Q(\ell(\vec{C}) \odot x^*) > 0\,,$$

where the last inequality is because $\ell(\vec{C})$ was sampled from the distribution $Q(\ell(\vec{C}) \odot x^*)$, and so it must be sampled with nonzero probability. As Q is supported only on satisfying assignments to the predicate P, it follows that $\ell(\vec{C}) \odot x^*$ must also satisfy P.

Step 2: removing the dependence on Q assuming a lower bound on $Q(y)$. First, we observe that because k is constant, we can, for each S, guess a symbol $\{0, +, -\}$, where 0 denotes, informally, the belief that $|\hat{Q}(S)| < 2^{-k}\varepsilon$, $+$ denotes that $\hat{Q}(S) \geqslant 2^{-k}\varepsilon$, and $-$ denotes that $\hat{Q}(S) \leqslant -2^{-k}\varepsilon$. For each of the 3^{2^k} choices of guesses, i.e., functions $f \colon \{S \subseteq [k]\} \to \{0, +, -\}$, we run algorithm mentioned in the previous step. Namely, for each S: (1) if $f(S) = 0$, then we ignore S, (2) if $f(S) = +$, then we run Theorem 6.2.3 on $\psi^{(S,+)}$ to obtain $\vec{\mathcal{H}}^{(S,1)}$ and $\vec{\mathcal{H}}^{(S,2)}$, and (3) if $f(S) = -$, then we run Theorem 6.2.3 on $\psi^{(S,+)}$ to obtain $\vec{\mathcal{H}}^{(S,1)}$ and $\vec{\mathcal{H}}^{(S,2)}$. As before, we solve the system of linear equations to obtain some assignment $x^{(f)} \in \{\pm 1\}^n$. By enumerating over all possible choices of f, we obtain a list of at most $3^{2^k} = O(1)$ assignments. We then try all of them and output the best one.

It thus remains to show that at least one of the assignments in the list has high value. As one may expect, this will be the assignment $x^{(f^*)}$, where f^* is the correct label function. Indeed, when $f = f^*$, then we are precisely running the algorithm in Step 1, and as observed, after solving the linear system of equations we obtain an assignment $x := x^{(f^*)}$ with the following property. For every $\vec{C} \in \vec{\mathcal{H}}'$ and every S with $|\hat{Q}(S)| \geqslant 2^{-k}\varepsilon$, we have that $\prod_{i \in S} x_{\vec{C}_i} = \prod_{i \in S} x^*_{\vec{C}_i}$, where $\vec{\mathcal{H}}' \subseteq \vec{\mathcal{H}}$ has size $\geqslant (1-\varepsilon)m$.

Finally, we show that for every $\vec{C} \in \vec{\mathcal{H}}'$, x satisfies the constraint $\vec{C}$. Namely, we have $P(\ell(\vec{C})_1 x_{\vec{C}_1}, \ell(\vec{C})_2 x_{\vec{C}_2}, \ldots, \ell(\vec{C})_k x_{\vec{C}_k}) = 1$. Let $\vec{C} \in \vec{\mathcal{H}}'$. We know that for every S with

$|\hat{Q}(S)| \geqslant 2^{-k}\varepsilon$, we have that $\prod_{i\in S} x_{\vec{C}_i} = \prod_{i\in S} x^*_{\vec{C}_i}$. Hence, it follows that

$$\begin{aligned}\left|Q(\ell(\vec{C})\odot x) - Q(\ell(\vec{C})\odot x^*)\right| &= \left|\sum_{S\subseteq[k]} \hat{Q}(S)\prod_{i\in S}\ell(\vec{C})_i x_{\vec{C}_i} - \sum_{S\subseteq[k]}\hat{Q}(S)\prod_{i\in S}\ell(\vec{C})_i x^*_{\vec{C}_i}\right| \\ &= \left|\sum_{S\subseteq[k]:|\hat{Q}(S)|<2^{-k}\varepsilon} \hat{Q}(S)\left(\prod_{i\in S}\ell(\vec{C})_i x_{\vec{C}_i} - \prod_{i\in S}\ell(\vec{C})_i x^*_{\vec{C}_i}\right)\right| \\ &\leqslant 2^k\cdot 2^{-k+1}\varepsilon\,.\end{aligned}$$

Now, if we assume that $Q(y) > 2\varepsilon$ for every $y \in \{\pm1\}^k$ with $Q(y) > 0$, then we have $Q(\ell(\vec{C})\odot x) > 0$, and so x satisfies the constraint $P(\ell(\vec{C})_1 x_{\vec{C}_1}, \ldots, \ell(\vec{C})_k x_{\vec{C}_k}) = 1$.

Step 3: removing the lower bound on $Q(y)$. In Step 2, we assumed that $Q(y) > 2\varepsilon$ for all $y \in \{\pm1\}^k$ with $Q(y) > 0$. However, we only used this fact in the final step, when we argue that $Q(\ell(\vec{C})\odot x) > 0$ by observing that $Q(\ell(\vec{C})\odot x) \geqslant Q(\ell(\vec{C})\odot x^*) - 2\varepsilon > 0$. To remove the assumption, we will show that for at most $2^{k+2}\varepsilon$ constraints $\vec{C} \in \vec{\mathcal{H}}$, it holds that $Q(\ell(\vec{C})\odot x^*) \leqslant 2\varepsilon$. This then implies that x satisfies at least $(1-\varepsilon-2^{k+2}\varepsilon)m = (1-O(\varepsilon))m$ constraints, which finishes the proof.

Let $\mathcal{S}$ denote the set of $\vec{C} \in \vec{\mathcal{H}}$ where $Q(\ell(\vec{C})\odot x^*) \leqslant 2\varepsilon$. Observe that the probability, over the choice of $\ell(\vec{C})$, that $\vec{C} \in \mathcal{S}$ is at most $2^k\cdot 2\varepsilon = 2^{k+1}\varepsilon$, and moreover this is independent for each $\vec{C} \in \vec{\mathcal{H}}$. Thus, by a Chernoff bound, it follows that with probability $\geqslant 1-\exp(-O(\varepsilon m)) \geqslant 1 - 1/\operatorname{poly}(n)$, we have $|\mathcal{S}| \leqslant 2\cdot 2^{k+1}\varepsilon$, and so we are done. □

Remark 8.2.3 (Tolerating fewer constraints for structured Q's)**.** We have shown that the above algorithm succeeds in finding an assignment x that satisfies at least $(1-O(\varepsilon))m$ constraints when $m \geqslant n^{k/2}\cdot\operatorname{poly}(\log n, 1/\varepsilon)$. However, if the distribution Q has $|\hat{Q}(S)| < 2^{-k}\varepsilon$ for all S with $|S| > r$, then we only need $n^{r/2}\cdot\operatorname{poly}(\log n, 1/\varepsilon)$ constraints. (If $r = 0$, then for small enough constant ε, Q will be supported on all of $\{\pm1\}^k$, and so any assignment satisfies all constraints. If $r = 1$, we require $O(n\cdot\frac{\log n}{\varepsilon})$ constraints; see Section C of [GHKM23].) Indeed, this follows because for such Q, the true label function f^* will have $f^*(S) = 0$ for any S with $|S| > r$. Hence, for this choice of f^*, we only call Theorem 6.2.3 on noisy t-XOR instances for $t \leqslant r$, and so we have enough constraints. It therefore follows that the assignment $x^{(f^*)}$ that we obtain for the label function f^* will be, with high probability an assignment that satisfies at least $(1-O(\varepsilon))m$ constraints.

An example where this gives an improvement is the well-studied NAE-3-SAT (not-all-equal-3SAT) predicate [AE98, ACIM01, DSS14]. Suppose Q is the uniform distribution over satisfying assignments to NAE-3-SAT: $Q(x_1,x_2,x_3) = \frac{1}{6}\cdot\frac{1}{4}(3 - x_1x_2 - x_2x_3 - x_1x_3)$. Then, we only need $m \geqslant \widetilde{O}(n)$ constraints, even though it is a 3-CSP ($k = 3$).

8.3 From k-XOR to spread bipartite k-XOR

In this section, we begin the proof of Theorem 6.2.3. See Definition 6.2.2 for a reminder of our semirandom planted k-XOR model $\psi(\mathcal{H}, x^*, \eta)$ given a k-uniform hypergraph $\mathcal{H}$, assignment $x^* \in \{\pm 1\}^n$, and noise parameter $\eta \in (0, 1/2)$. Recall also that $\mathcal{E}_\psi$ denotes the set of corrupted hyperedges.

We think of $\mathcal{A}_1(\mathcal{H})$ as the small set of edges that we discard (or give up on), and this will only depend on the hypergraph $\mathcal{H}$. For the rest of the graph, the algorithm will correctly identify which edges are corrupted.

Our proof of Theorem 6.2.3 goes via a reduction to *spread bipartite t-XOR* instances for $t = 2, \dots, k$, which are t-XOR instances with some additional desired structure. Such instances were introduced in [GKM22] to study the refutation of semirandom k-XOR instances. The reduction here is nearly identical to the corresponding reduction in [GKM22, Section 4].

Definition 8.3.1 (Spread bipartite k-XOR)**.** A p-bipartite k-XOR instance ψ on n variables with m constraints is defined by a collection of $(k-1)$-uniform hypergraphs $\mathcal{H} = \{\mathcal{H}_u\}_{u \in [p]}$ on the vertex set $[n]$, as well as "right-hand sides" $b_{u,C}$ for each $u \in [p]$ and $C \in \mathcal{H}_u$. There are two sets of variables of ψ: the "normal" variables $x_1, \dots, x_n$, and the "special" variables $y_1, \dots, y_p$. The constraints of ψ are $y_u \prod_{i \in C} x_i = b_{u,C}$ for each $u \in [p]$, $C \in \mathcal{H}_u$.

We furthermore say that ψ is *τ-spread* if it has the following additional properties:

(1) $|\mathcal{H}_u| = \frac{m}{p} \geqslant 2\lfloor \frac{1}{2\tau^2} \rfloor$ and $\frac{m}{p}$ is even for each $u \in [p]$,

(2) For each $u \in [p]$ and set $Q \subseteq [n]$, $\deg_u(Q) \leqslant \frac{1}{\tau^2} \max(1, n^{\frac{k}{2}-1-|Q|})$.

Analogously to Definition 6.2.2, we call ψ a *semirandom planted* instance with planted assignment (x^*, y^*) and noise parameter η if the right-hand sides $b_{u,C}$ are generated by setting $b_{u,C} = y_u^* \prod_{i \in C} x_i^*$ with probability $1 - \eta$ and $b_{u,C} = -y_u^* \prod_{i \in C} x_i^*$ otherwise, independently for each choice of u, C. For a choice of x^*, y^*, $\mathcal{H} = \{\mathcal{H}_u\}_{u \in [p]}$, and η, we call this distribution $\psi(\{\mathcal{H}_u\}_{u \in [p]}, x^*, y^*, \eta)$. As before, if an edge (u, C) has $b_{u,C} = -y_u^* \prod_{i \in C} x_i^*$, we call (u, C) a *corrupted* hyperedge, and we denote the set of corrupted hyperedges in ψ by $\mathcal{E}_\psi$.

The main technical result of the paper is the following lemma, which gives an algorithm to find the noisy constraints in a semirandom planted τ-spread bipartite k-XOR instance.

Lemma 8.3.2 (Algorithm for τ-spread bipartite k-XOR)**.** *Let $k \geqslant 2$, $n, p \in \mathbb{N}$, $\varepsilon \in (0, 1)$, $\eta \in [0, 1/2)$, and let $\gamma := 1 - 2\eta > 0$. Let $\tau \leqslant \frac{c\gamma}{\sqrt{k \log n}}$, and let $m \geqslant C n^{\frac{k-1}{2}} \sqrt{p} \cdot \frac{(k \log n)^{3/2}}{\tau \gamma^2 \varepsilon^{3/2}}$ for some universal constants c, C. There is a polynomial-time algorithm $\mathcal{A}$ that takes as input an τ-spread p-bipartite k-XOR instance ψ with constraint hypergraph $\mathcal{H} = \{\mathcal{H}_u\}_{u \in [p]}$ and outputs*

two disjoint sets $\mathcal{A}_1(\mathcal{H}), \mathcal{A}_2(\psi) \subseteq \mathcal{H}$ with the following guarantee: (1) for any instance ψ with m constraints, $|\mathcal{A}_1(\mathcal{H})| \leqslant \varepsilon m$ and $\mathcal{A}_1(\mathcal{H})$ only depends on $\mathcal{H}$, and (2) for any $x^* \in \{\pm 1\}^n, y^* \in \{\pm 1\}^p$ and any $\mathcal{H} = \{\mathcal{H}_u\}_{u \in [p]}$ with $|\mathcal{H}| := \sum_{u \in [p]} |\mathcal{H}_u| \geqslant m$, with probability $1 - \frac{1}{\mathrm{poly}(n)}$ over $\psi \leftarrow \psi(\{\mathcal{H}_u\}_{u \in [p]}, x^*, y^*, \eta)$, it holds that $\mathcal{A}_2(\psi) = \mathcal{E}_\psi \cap (\mathcal{H} \setminus \mathcal{A}_1(\mathcal{H}))$.

Note that as $\eta \to \frac{1}{2}$, $\gamma = 1 - 2\eta \to 0$ and $\tau \to 0$, which blows up m. This is the expected behavior since when $\eta = \frac{1}{2}$, it is impossible to recover the planted assignment since the signs of the constraints are uniformly random.

8.3.1 Proof of Theorem 6.2.3 from Lemma 8.3.2

With Lemma 8.3.2, we can finish the proof of Theorem 6.2.3. The high-level idea of this proof is very simple. First, we decompose the k-XOR instance ψ into subinstances $\psi^{(t)}$ for each $t = 2, \dots, k$, using a hypergraph decomposition algorithm very similar to the one used to prove the hypergraph Moore bound (Algorithm 4.4.2).

Algorithm 8.3.3.

Given: A semirandom (with noise η) k-XOR instance ψ with constraint hypergraph $\mathcal{H}$ over n vertices, and a spread parameter $\tau \in (0, 1)$.

Output: For each $t = 2, \dots, k$, a semirandom (with noise η) planted τ-spread $p^{(t)}$-bipartite t-XOR instance $\psi^{(t)}$ with constraint hypergraph $\{\mathcal{H}_u^{(t)}\}_{u \in [p^{(t)}]}$, along with "discarded" hyperedges $\mathcal{H}^{(1)}$.

Operation:

1. **Initialize:** $\psi^{(t)}$ to the empty instance, and $p^{(t)} = 0$ for $t = 2, \dots, k$.
2. **Fix violations greedily**:
   (a) Find a maximal nonempty violating Q. That is, find $Q \subseteq [n]$ of size $1 \leqslant |Q| \leqslant k - 1$ such that $\deg(Q) = |\{C \in \mathcal{H} : Q \subseteq C\}| > \frac{1}{\tau^2} \max(1, n^{\frac{k}{2} - |Q|})$, and $\deg(Q') \leqslant \frac{1}{\tau^2} \max(1, n^{\frac{k}{2} - |Q'|})$ for all $Q' \supsetneq Q$.
   (b) Let $q = |Q|$. Let $u = 1 + p^{(k+1-q)}$ be a new "label", and define $\mathcal{H}_u^{(k+1-q)}$ to be an arbitrary subset of $\{C \setminus Q : C \in \mathcal{H}, Q \subseteq C\}$ of size exactly $2 \cdot \lfloor \frac{1}{2\tau^2} \max(1, n^{\frac{k}{2} - q}) \rfloor$.
   (c) Set $p^{(k+1-q)} \leftarrow 1 + p^{(k+1-q)}$, and $\mathcal{H} \leftarrow \mathcal{H} \setminus \mathcal{H}_u^{(k+1-q)}$.
3. If no such Q exists, then put the remaining hyperedges in $\mathcal{H}^{(1)}$.

Then, we run the algorithm in Lemma 8.3.2 to identify a set of corrupted constraints and a small set of discarded constraints within each subinstance $\psi^{(t)}$. We then take the union of these outputs to be the final output of the algorithm.

Proof of Theorem 6.2.3. We begin with the decomposition of ψ into $\psi^{(2)}, \dots, \psi^{(k)}$ along with a set of "discarded" hyperedges $\mathcal{H}^{(1)}$, which is done using Algorithm 8.3.3 with

spread parameter $\tau := \frac{c(1-2\eta)}{\sqrt{k \log n}}$ where c is the constant in Lemma 8.3.2. For each $t = 2, \dots, k$, $\psi^{(t)}$ is a semirandom (with noise η) planted τ-spread $p^{(t)}$-bipartite t-XOR instance specified by $(t-1)$-uniform hypergraphs $\{\mathcal{H}_u^{(t)}\}_{u \in [p^{(t)}]}$.

Let $m^{(t)} := \sum_{u \in [p^{(t)}]} |\mathcal{H}_u^{(t)}|$. Algorithm 8.3.3 has the following guarantees:

(1) The runtime is $n^{O(k)}$,

(2) For each $t \in \{2, \dots, k\}$ and $u \in [p^{(t)}]$, $|\mathcal{H}_u^{(t)}| = \frac{m^{(t)}}{p^{(t)}} = 2\lfloor \frac{1}{2\tau^2} \max(1, n^{t-\frac{k}{2}-1}) \rfloor$; in particular, $|\mathcal{H}_u^{(t)}|$ is even and is at least $2\lfloor \frac{1}{2\tau^2} \rfloor$,

(3) For each $t = 2, \dots, k$, the instance $\psi^{(t)}$ is τ-spread,

(4) The number of "discarded" hyperedges is $m^{(1)} := |\mathcal{H}^{(1)}| \leqslant \frac{1}{k\tau^2} n^{\frac{k}{2}}$,

(5) For $t \in \{2, \dots, k\}$, each $C \in \mathcal{H}_u^{(t)}$ is obtained by removing $k - (t-1)$ vertices from an edge in the original hypergraph $\mathcal{H}$. Thus, there is a one-to-one map $\mathsf{Decomp} \colon \mathcal{H} \to \mathcal{H}^{(1)} \cup \bigcup_{t=2}^{k} \{\mathcal{H}_u^{(t)}\}_{u \in [p^{(t)}]}$, such that an edge $C \in \mathcal{H}$ is corrupted if and only if the edge $\mathsf{Decomp}(C)$ is corrupted in the instance $\psi^{(t)}$ that it lies in.

For convenience, we denote $\gamma := 1 - 2\eta$ and $\beta := 4C \cdot \frac{(k \log n)^{3/2}}{\tau \gamma^2 \varepsilon^{3/2}} = \frac{4C}{c} \cdot \frac{k^2 \log^2 n}{\gamma^3 \varepsilon^{3/2}}$ where C, c are the constants in Lemma 8.3.2. The algorithm in Theorem 6.2.3 works as follows. First, it runs Algorithm 8.3.3 to produce the instances $\psi^{(2)}, \dots, \psi^{(k)}$. Then, for each $t = 2, \dots, k$, if $m^{(t)} \geqslant n^{\frac{t-1}{2}} \sqrt{p^{(t)}} \cdot \beta$, we run Lemma 8.3.2 on $\psi^{(t)}$ and obtain, with probability $1 - 1/\operatorname{poly}(n)$, a set $A_1^{(t)}$ where $|A_1^{(t)}| \leqslant \frac{\varepsilon}{2} m^{(t)}$ and $A_2^{(t)} = \mathcal{E}_{\psi^{(t)}} \setminus A_1^{(t)}$. Otherwise, if $m^{(t)} < n^{\frac{t-1}{2}} \sqrt{p^{(t)}} \cdot \beta$, we set $A_1^{(t)} = \mathcal{H}^{(t)}$ and $A_2^{(t)} = \emptyset$. Finally, we output $\mathcal{A}_1 := \mathcal{H}^{(1)} \cup \bigcup_{t=2}^{k} \mathsf{Decomp}^{-1}(A_1^{(t)})$ and $\mathcal{A}_2 := \bigcup_{t=2}^{k} \mathsf{Decomp}^{-1}(A_2^{(t)})$, where Decomp is the mapping in property (5) of Algorithm 8.3.3.

Note that $m^{(t)} = p^{(t)} |\mathcal{H}_u^{(t)}| \geqslant p^{(t)} \cdot \frac{1}{2\tau^2} n^{t-\frac{k}{2}-1}$, which means $p^{(t)} \leqslant 2\tau^2 n^{\frac{k}{2}-t+1} m^{(t)}$, and since $\sum_t \sqrt{m^{(t)}} \leqslant \sqrt{k \sum_t m^{(t)}} \leqslant \sqrt{km}$ by Cauchy-Schwarz, we have

$$\sum_{t=2}^{k} n^{\frac{t-1}{2}} \sqrt{p^{(t)}} \cdot \beta \leqslant O(\tau) \cdot n^{\frac{k}{4}} \sqrt{km} \cdot \beta \leqslant o(\varepsilon) m$$

as long as $m \gg n^{\frac{k}{2}} \cdot k\tau^2 \beta^2 / \varepsilon^2$. Moreover, $m^{(1)} \leqslant \frac{1}{k\tau^2} n^{\frac{k}{2}} = \frac{\log n}{c^2 \gamma^2} n^{\frac{k}{2}} \leqslant o(\varepsilon) m$. One can verify, by plugging in β, that the lower bound on m in Theorem 6.2.3 suffices.

By union bound over t, it thus follows that

$$|\mathcal{A}_1| \leqslant m^{(1)} + \sum_{t=2}^{k} \frac{\varepsilon}{2} m^{(t)} + \sum_{t=2}^{k} n^{\frac{t-1}{2}} \sqrt{p^{(t)}} \beta \leqslant \varepsilon m \,,$$

and $\mathcal{A}_2 = \mathcal{E}_\psi \setminus \mathcal{A}_1$. Moreover, by Lemma 8.3.2, $\mathcal{A}_1$ only depends on the hypergraph $\mathcal{H}$. This completes the proof. □

8.4 Identifying noisy constraints in spread bipartite k-XOR

In this section, we prove Lemma 8.3.2. The proof will be decomposed into the following steps. First, we take the semirandom planted bipartite k-XOR instance ψ and transform it into a 2-XOR instance ϕ. Second, we decompose the constraint graph of ϕ into expanders. For each expander in the decomposition, we argue that the SDP solution to this subinstance is rank 1, and moreover agrees *exactly* with the planted assignment. This allows us to identify, for each expanding subinstance, *exactly* which edges in ϕ are errors. Finally, we use this information to identify the set of corrupted constraints in the original instance ψ, which finishes the proof.

8.4.1 Setup and key notation

We now introduce the key notation that shall be used throughout this section. Let ψ be the semirandom τ-spread p-bipartite k-XOR instance (recall Definition 8.3.1) with m constraints given as the input to the algorithm. Recall that the instance ψ is specified by a collection of p hypergraphs $\{\mathcal{H}_u\}_{u\in[p]}$, where each $\mathcal{H}_u$ is a $(k-1)$-uniform hypergraph on n vertices and $|\mathcal{H}_u| = m/p$. Each constraint in ψ is specified by a pair (u, C) where $u \in [p]$, $C \in \mathcal{H}_u$, and has a right-hand side $b_{u,C} \in \{\pm 1\}$, and the constraints are $y_u \prod_{i\in C} x_i = b_{u,C}$, where $\{y_u\}_{u\in[p]}$ and $\{x_i\}_{i\in[n]}$ are variables. Because the instance ψ is semirandom with noise parameter η and planted assignment (x^*, y^*), for each constraint (u, C) we have, with probability $1-\eta$ independently, $b_{u,C} = y^*_u \prod_{i\in C} x^*_i$, and otherwise $b_{u,C} = -y^*_u \prod_{i\in C} x^*_i$. Our goal is to output, in $n^{O(k)}$-time, a set $\mathcal{A}_1(\mathcal{H})$ of size $\leqslant \tau m$ to discard, and then for the rest of the instance, identify exactly the corrupted constraints, i.e., those for which $b_{u,C} = -y^*_u \prod_{i\in C} x^*_i$.

We now define the 2-XOR instance ϕ from ψ. An example is shown in Fig. 8.2.

Definition 8.4.1 (2-XOR instance ϕ from bipartite k-XOR ψ)**.** For every $u \in [p]$ and $\mathcal{H}_u$, we partition $\mathcal{H}_u$ arbitrarily into two sets $\mathcal{H}^{(L)}_u$ and $\mathcal{H}^{(R)}_u$ of equal size.

- If k is odd, then there are $\binom{n}{\frac{k-1}{2}}^2$ variables in ϕ, one variable $z_{(S_1,S_2)}$ for each pair of sets $S_1, S_2 \subseteq [n]$ where $|S_1| = |S_2| = \frac{k-1}{2}$.
- If k is even, then there are $2\binom{n}{\lceil\frac{k-1}{2}\rceil}\binom{n}{\lfloor\frac{k-1}{2}\rfloor}$ variables in ϕ, one variable $z_{(S_1,S_2)}$ for each pair of sets $S_1, S_2 \subseteq [n]$ where either $|S_1| = \lceil\frac{k-1}{2}\rceil$ and $|S_2| = \lfloor\frac{k-1}{2}\rfloor$ or $|S_1| = \lfloor\frac{k-1}{2}\rfloor$ and $|S_2| = \lceil\frac{k-1}{2}\rceil$.

For each $u \in [p]$, $C \in \mathcal{H}^{(L)}_u$ and $C' \in \mathcal{H}^{(R)}_u$, we arbitrarily partition C into sets $S_1 \cup S_2$ and C' into sets $S'_1 \cup S'_2$, where $|S_1| = |S'_1| = \lceil\frac{k-1}{2}\rceil$ and $|S_2| = |S'_2| = \lfloor\frac{k-1}{2}\rfloor$. We then add the constraint $z_{(S_1,S'_2)} z_{(S_2,S'_1)} = b_{u,C} b_{u,C'}$ to ϕ.

It is intuitive to think of clauses from $\mathcal{H}^{(L)}_u$ and $\mathcal{H}^{(R)}_u$ as having different colors, and

each variable $z_{(S_1,S_2')}$ contains roughly $k/2$ of each color. See Fig. 8.2 for an example of a 2-XOR ϕ constructed from a bipartite k-XOR ψ.

Observation 8.4.2 (Size of ϕ)**.** The number of variables in ϕ is at most n^{k-1} (for both even and odd k). Since each $|\mathcal{H}_u| = m/p$, $|\mathcal{H}_u^{(L)}| = |\mathcal{H}_u^{(R)}| = \frac{m}{2p}$, and the number of constraints in ϕ is exactly $p \cdot (\frac{m}{2p})^2 = \frac{m^2}{4p}$. In particular, when $m \geqslant n^{\frac{k-1}{2}}\sqrt{p} \cdot \beta$ for $\beta = \text{poly}(\log n)$ as assumed in Lemma 8.3.2, the average degree of ϕ is at least $\frac{1}{4}\beta^2$.

Remark 8.4.3 (Corrupted constraints in ϕ)**.** A constraint $z_{(S_1,S_2')}z_{(S_2,S_1')} = b_{u,C}b_{u,C'}$ in ϕ is *corrupted* if exactly one of $b_{u,C}$ and $b_{u,C'}$ is corrupted in ψ. Thus, if each constraint in ψ is corrupted with probability $\eta \in (0, 1/2)$, then each constraint in ϕ is corrupted with probability $2\eta(1-\eta) < 1/2$. Note, however, that the constraints in ϕ are not corrupted independently.

We need some more definitions about the constraint graph of ϕ.

Definition 8.4.4 (Constraint graph of ϕ)**.** Let $G(\phi) = (V, E)$ be the constraint graph of ϕ. Notice that each edge $e \in E$ uniquely identifies $u(e) \in [p]$ and $C_L(e) \in \mathcal{H}_{u(e)}^{(L)}$, $C_R(e) \in \mathcal{H}_{u(e)}^{(R)}$. For each $u \in [p]$, $C \in \mathcal{H}_u^{(L)}$, define $G_{u,C}^{(L)}(\phi)$ to be the subgraph of G that C participates in, i.e., with edge set $\{e \in E : u(e) = u,\ C_L(e) = C\}$. We similarly define $G_{u,C'}^{(R)}(\phi)$ for $C' \in \mathcal{H}_u^{(R)}$.

We next make the important observation that the degree of a vertex in $G_{u,C}^{(L)}(\phi)$ is upper bounded by the number of $C' \in \mathcal{H}_u^{(R)}$ sharing at least $\lfloor \frac{k-1}{2} \rfloor$ vertices. See Fig. 8.2 also for an illustration. Therefore, assuming that ψ is τ-spread, we have a maximum degree bound on $G_{u,C}^{(L)}(\phi)$ and $G_{u,C'}^{(R)}(\phi)$ for all $u \in [p]$, $C \in \mathcal{H}_u^{(L)}$ and $C' \in \mathcal{H}_u^{(R)}$.

Lemma 8.4.5 (Degree bounds for $G_{u,C}^{(L)}$, $G_{u,C'}^{(R)}$)**.** *Let ψ be an τ-spread p-bipartite k-XOR instance. Then, for any $u \in [p]$, $C \in \mathcal{H}_u^{(L)}$ and $C' \in \mathcal{H}_u^{(R)}$, the maximum degree of $G_{u,C}^{(L)}(\phi)$, $G_{u,C'}^{(R)}(\phi)$ is at most $1/\tau^2$.*

Proof. Consider any $C \in \mathcal{H}_u^{(L)}$ and two adjacent edges $\{z_{(S_1,S_2')}, z_{(S_2,S_1')}\}$ and $\{z_{(S_1,S_2'')}, z_{(S_2,S_1'')}\}$ in $G_{u,C}^{(L)}(\phi)$ formed by joining $C = S_1 \cup S_2$ with $C' = S_1' \cup S_2'$ and $C'' = S_1'' \cup S_2'' \in \mathcal{H}_u^{(R)}$. As the edges are adjacent, it must be the case that either $S_1' = S_1''$ or $S_2' = S_2''$, which means that $|C' \cap C''| \geqslant \lfloor \frac{k-1}{2} \rfloor$. Thus, the degree of a vertex $z_{(S_1,S_2')}$ in G is upper bounded by the maximum number of $C' \in \mathcal{H}_u^{(R)}$ that all share the same $\lfloor \frac{k-1}{2} \rfloor$ variables.

Suppose ψ is τ-spread, meaning that $\deg_u(Q) \leqslant \frac{1}{\tau^2}\max(1, n^{\frac{k}{2}-1-|Q|})$ for $Q \subseteq [n]$. Since $\frac{k}{2} - 1 - \lfloor \frac{k-1}{2} \rfloor \leqslant 0$, we have that $G_{u,c}^{(L)}(\phi)$ has maximum degree $\leqslant 1/\tau^2$. □

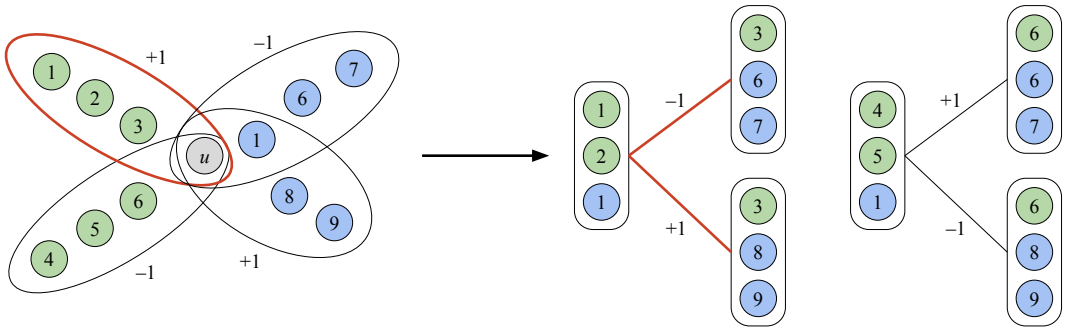


Figure 8.2: An example of the 2-XOR instance ϕ from a bipartite 4-XOR ψ (Definition 8.4.1). On the left, $\mathcal{H}_u^{(L)}$ consists of $C_1 = \{1,2,3\}$ and $C_2 = \{4,5,6\}$ (with green vertices), and $\mathcal{H}_u^{(R)}$ consists of $C_1' = \{1,6,7\}$ and $C_2' = \{1,8,9\}$ (with blue vertices). On the right, the constraint graph $G(\phi)$ has vertices z_{S_1,S_2} where either $|S_1| = 2$, $|S_2| = 1$ or $|S_1| = 1$, $|S_2| = 2$ (we can view S_1, S_2 as having green, blue vertices). Each edge corresponds to two clauses in ψ; for example, the edge $\left\{z_{\{1,2\},\{1\}}, z_{\{3\},\{6,7\}}\right\}$ comes from the clauses C_1 and C_1'.

Corruptions. In the figure, we label a clause -1 if it is corrupted and $+1$ otherwise. An edge in G is corrupted if exactly one of the two corresponding clauses in ψ is corrupted.

Degree of $G_{u,C}^{(L)}(\phi)$. For $C_1 \in \mathcal{H}_u^{(L)}$, the subgraph $G_{u,C_1}^{(L)}(\phi)$ corresponds to the edges colored red, i.e., all edges that C_1 participates in. The vertex $z_{\{1,2\},\{1\}}$ has degree 2 in $G_{u,C_1}^{(L)}(\phi)$ because $|C_1' \cap C_2'| = 1$.

8.4.2 Proof outline

With the setup in Section 8.4.1 in hand, our proof now proceeds in three conceptual steps.

Step 1: graph pruning and expander decomposition. Suppose the instance ϕ has average degree d. We first prune the instance using Lemma 2.3.1 such that the resulting constraint graph has minimum degree $\geqslant \varepsilon d$ while only removing ε fraction of the constraints, where $\varepsilon = o(1)$. We further apply expander decomposition (Fact 2.3.2) to the pruned instance to obtain subinstances $\phi_1, \ldots, \phi_T$ while discarding only a ε fraction of the constraints of ϕ such that the constraint graph of each ϕ_i has spectral gap $\widetilde{\Omega}(\varepsilon^2)$.

Step 2: relative spectral approximation and recovery of corrupted pairs. We show that for each expanding subinstance ϕ_i, the basic SDP for the 2-XOR instance ϕ_i is equal to $x^*(x^*)^\top$, where x^* is the planted assignment for ϕ. That is, the SDP solution is *rank* 1 and agrees with the *planted assignment* for ϕ. We show this by arguing that, for each ϕ_i, the Laplacian of the corrupted constraints in ϕ_i is a *spectral sparsifier* of the Laplacian of the constraint graph of ϕ_i (see Lemma 8.1.4). Here, we crucially use that each such constraint graph has large minimum degree and spectral gap.

From this, it is trivial to identify the corrupted edges in each ϕ_i, as they are the ones violated by the SDP solution. We are not quite done yet, however, because each constraint in ϕ corresponds to a *pair* of constraints in the original instance ψ.

Step 3: recovery of corrupted constraints from corrupted pairs. The previous step shows that for all but a ε fraction of tuples (u, C, C') where $u \in [p]$, $C \in \mathcal{H}_u^{(L)}$, and $C' \in \mathcal{H}_u^{(R)}$, we can recover the product $\xi_u(C)\xi_u(C')$, where $\xi_u(C) = -1$ if (u, C) is noisy in ψ, and is $+1$ otherwise. Because ε is small, it must be the case that for most $u \in [p]$, we know the product $\xi_u(C)\xi_u(C')$ (from Step 2) for *most* pairs (C, C') with $C \in \mathcal{H}_u^{(L)}$ and $C' \in \mathcal{H}_u^{(R)}$.

Suppose we knew $\xi_u(C)\xi_u(C')$ for all $(C, C') \in \mathcal{H}_u^{(L)} \times \mathcal{H}_u^{(R)}$. Then, it is trivial to decode $\xi_u(C)$ *up to a global sign*. Formally, we could obtain $z \in \{\pm 1\}^{\mathcal{H}_u}$ where $z_C = \alpha\xi_u(C)$ for some $\alpha \in \{\pm 1\}$. From this, it is easy to obtain $\xi_u(C)$, as the fraction of $C \in \mathcal{H}_u$ for which $\xi_u(C) = -1$ should be roughly $\eta < \frac{1}{2}$; so, if z has $< \frac{1}{2}$-fraction of -1's, then $z = \xi_u(C)$, and otherwise $-z = \xi_u(C)$. This, however, requires $|\mathcal{H}_u| \geqslant \Omega\left(\frac{\log n}{(1-2\eta)^2}\right)$ for a high-probability result.

Additionally, we do not quite know $\xi_u(C)\xi_u(C')$ for all $(C, C') \in \mathcal{H}_u^{(L)} \times \mathcal{H}_u^{(R)}$: we only know this for all but a ε_u-fraction of the pairs. By forming a graph G_u where we have an edge (C, C') if (C, C') is a pair where we know $\xi_u(C)\xi_u(C')$, we can thus obtain such a z for all C in the largest connected component of G_u. Because G_u is obtained by taking a *complete biclique* and deleting only a ε_u-fraction of all edges, the largest connected component has size $(1 - \varepsilon_u)|\mathcal{H}_u|$, and so we can recover $\xi_u(C)$ for all but a ε_u-fraction of constraints in $\mathcal{H}_u$. We do this for each partition u, which finishes the proof.

8.4.3 Graph pruning and expander decomposition

This step is a simple combination of graph pruning and expander decomposition.

Lemma 8.4.6. *Fix $\varepsilon \in (0, 1)$. There is a polynomial-time algorithm such that, given a 2-XOR instance ϕ whose constraint graph has m edges and average degree d, outputs subinstances $\phi_1, \ldots, \phi_T$ on disjoint variables with the following guarantees: $\phi_1, \ldots, \phi_T$ contain at least $1 - \varepsilon$ fraction of the constraints in ϕ, and for each $i \in [T]$, the constraint graph G_i of ϕ_i, after adding some self-loops, has minimum degree at least $\frac{1}{3}\varepsilon d$ and $\lambda_2(\widetilde{L}_{G_i}) \geqslant \Omega(\varepsilon^2 / \log^2 m)$.*

The self-loops in Lemma 8.4.6 are only for the analysis of $\widetilde{L}_{G_i}$ and do not correspond to actual constraints in ϕ_i. Observe that adding self-loops to a graph G does not change the *unnormalized* Laplacian L_G, but as D_G (the degree matrix) increases, the spectral gap of the *normalized* Laplacian, i.e. $\lambda_2(\widetilde{L}_G) = \lambda_2(D_G^{-1/2} L_G D_G^{-1/2})$, may decrease. The expander decomposition algorithm (Fact 2.3.2) guarantees that each piece, even after adding self-loops to preserve degrees, has large spectral gap. This does not change the

subinstances $\phi_1, \dots, \phi_T$, but in the next section, it is crucial that we use this stronger guarantee to ensure a lower bound on the minimum degree.

Proof of Lemma 8.4.6. We first apply the graph pruning algorithm (Lemma 2.3.1) such that the resulting instance has minimum degree $\geqslant \frac{\varepsilon}{3}d$ and at least $(1-\frac{2}{3}\varepsilon)m$ constraints. Then, we apply expander decomposition (Fact 2.3.2) that partitions the vertices of the pruned graph G' into $V_1, \dots, V_T$ such that the number of edges across partitions is at most $\frac{\varepsilon}{3}m$, and for each $i \in [T]$, the normalized Laplacian satisfies $\lambda_2(\widetilde{L}_{G'\{V_i\}}) \geqslant \Omega(\varepsilon^2/\log^2 m)$. Here we recall that $G'\{V_i\}$ is the induced subgraph of G' with self-loops such that the vertices in $G'\{V_i\}$ have the same degrees as in G'.

In total, we have removed at most εm edges. This completes the proof. □

8.4.4 Rank-1 SDP solution from expansion and relative spectral approximation

We next show that for each subinstance ϕ_i obtained from Lemma 8.4.6, its constraint graph G and the subgraph of corrupted edges H satisfy $L_H \prec \frac{1}{2}L_G$. Recall from Lemmas 8.1.4 and 8.1.5 that this implies the basic SDP for the 2-XOR ϕ_i is rank 1 and agrees with the planted assignment of ϕ.

The next lemma is analogous to Lemma 8.1.5 but differs in an important way: a constraint in ϕ is corrupted if and only if exactly one of the two corresponding constraints in ψ is corrupted; thus, the corruptions in ϕ are *correlated*. This is why each constraint in ϕ is obtained from one clause in $\mathcal{H}_u^{(L)}$ and one clause in $\mathcal{H}_u^{(R)}$ (recall Definition 8.4.1), so that in the proof below we have independent randomness to perform a "2-step sparsification" proof. It is also worth noting that the following lemma requires not just a lower bound on the minimum degree and spectral gap of G but also that the original bipartite k-XOR instance ψ is *well-spread*, which allows us to apply Lemma 8.4.5.

Same as Lemma 8.1.5, the following lemma is a purely graph-theoretic statement.

Lemma 8.4.7 (Relative spectral approximation with correlated subsamples)**.** *Suppose $G = (V, E)$ is an n-vertex graph with minimum degree $d_{\min}$ (self-loops and parallel edges allowed) and spectral gap $\lambda_2(\widetilde{L}_G) = \lambda > 0$. Let $m_1, m_2 \in \mathbb{N}$, $\eta \in [0, 1/2)$, and let $\xi_1^{(1)}, \dots, \xi_{m_1}^{(1)}, \xi_1^{(2)}, \dots, \xi_{m_2}^{(2)}$ be i.i.d. random variables that take value -1 with probability η and $+1$ otherwise. Suppose there is an injective map that maps each edge $e \mapsto (c_1(e), c_2(e)) \in [m_1] \times [m_2]$, and for each $i \in [m_1]$ (resp. $j \in [m_2]$) define $G_i^{(1)}$ (resp. $G_j^{(2)}$) be the subgraph of G with edge set $\{e \in E : c_1(e) = i\}$ (resp. $\{e \in E : c_2(e) = j\}$). Moreover, suppose $G_i^{(1)}$ and $G_j^{(2)}$ have maximum degree $\leqslant \Delta$ for all $i \in [m_1]$, $j \in [m_2]$.*

Let H be the subgraph of G with edge set $\{e \in E : \xi_{c_1(e)}^{(1)} \xi_{c_2(e)}^{(2)} = -1\}$. There is a universal

constant $B > 0$ such that if $d_{\min}\lambda \geqslant B\Delta \log n$, then with probability $1 - O(n^{-2})$,

$$L_H \preceq \max\left((1+\delta)\cdot 2\eta(1-\eta),\ \frac{1}{3}\right)\cdot L_G$$

for $\delta = \sqrt{\frac{B\Delta \log n}{d_{\min}\lambda}}$.

Let $\gamma := 1 - 2\eta > 0$ since $\eta < \frac{1}{2}$. Notice that $2\eta(1-\eta) = \frac{1}{2}(1-\gamma^2)$, which approaches $\frac{1}{2}$ as $\eta \to \frac{1}{2}$. Thus, if $\delta \leqslant \gamma^2$, then $(1+\delta)\cdot 2\eta(1-\eta) \leqslant (1+\gamma^2)\cdot\frac{1}{2}(1-\gamma^2) < \frac{1}{2}$, and $L_H \prec \frac{1}{2}L_G$ suffices to conclude via Lemma 8.1.4 that the SDP relaxation on the expanding subinstance is rank 1 and recovers the planted assignment, which also gives us the set of corrupted constraints.

Proof of Lemma 8.4.7. First, note that by the definition of Laplacian and the spectral gap of L_G, $\mathrm{span}(\vec{1})$ is exactly the null space of L_G and is contained in the null space of L_H. Therefore, recalling that $L_G = D_G^{1/2}\widetilde{L}_G D_G^{1/2}$, it suffices to prove that

$$\left\|(\widetilde{L}_G^{\dagger})^{1/2} D_G^{-1/2} L_H D_G^{-1/2} (\widetilde{L}_G^{\dagger})^{1/2}\right\|_2 \leqslant \max\left((1+\delta)\cdot 2\eta(1-\eta),\ \frac{1}{3}\right). \tag{8.1}$$

Here $\widetilde{L}_G^{\dagger}$ is the pseudo-inverse of $\widetilde{L}_G$, and $\|\widetilde{L}_G^{\dagger}\|_2 \leqslant 1/\lambda$. For simplicity, for any graph G', we will write $\widehat{L}_{G'} := (\widetilde{L}_G^{\dagger})^{1/2} D_G^{-1/2} L_{G'} D_G^{-1/2} (\widetilde{L}_G^{\dagger})^{1/2}$. Thus,

$$\widehat{L}_H = \sum_{e\in E} \mathbf{1}\left(\xi^{(1)}_{c_1(e)}\xi^{(2)}_{c_2(e)} = -1\right)\cdot \widehat{L}_e\,, \text{ and } \mathbb{E}[\widehat{L}_H] = 2\eta(1-\eta)\sum_{e\in E}\widehat{L}_e\,.$$

Note that $\sum_{e\in E}\widehat{L}_e = \widehat{L}_G$, a projection matrix, thus $\left\|\sum_{e\in E}\widehat{L}_e\right\|_2 = 1$.

For each $i \in [m_1]$, we further define $G^{(1)}_{i,+}$ and $G^{(1)}_{i,-}$ to be (random) edge-disjoint subgraphs of $G^{(1)}_i$ where $G^{(1)}_{i,+}$ has edge set $\left\{e \in E : c_1(e) = i, \xi^{(2)}_{c_2(e)} = +1\right\}$ and $G^{(1)}_{i,-}$ has edge set $\left\{e \in E : c_1(e) = i, \xi^{(2)}_{c_2(e)} = -1\right\}$. Note that $G^{(1)}_{i,+}$, $G^{(1)}_{i,-}$ are independent of $\xi^{(1)} = (\xi^{(1)}_1, \ldots, \xi^{(1)}_{m_1})$. By the maximum degree bound on $G^{(1)}_i$, we have that $\left\|L_{G^{(1)}_{i,+}}\right\|_2$ and $\left\|L_{G^{(1)}_{i,-}}\right\|_2 \leqslant \left\|L_{G^{(1)}_i}\right\|_2 \leqslant 2\Delta$. Thus,

$$\left\|\widehat{L}_{G^{(1)}_{i,+}}\right\|_2,\ \left\|\widehat{L}_{G^{(1)}_{i,-}}\right\|_2 \leqslant \left\|\widehat{L}_{G^{(1)}_i}\right\|_2 \leqslant 2\Delta\cdot\left\|\widetilde{L}_G^{\dagger}\right\|_2\cdot\left\|D_G^{-1}\right\|_2 \leqslant \frac{2\Delta}{d_{\min}\lambda}\,. \tag{8.2}$$

Similarly, for $j \in [m_2]$, $G^{(2)}_{j,+}$ and $G^{(2)}_{j,-}$ are (random) edge-disjoint subgraphs of $G^{(2)}_j$ independent of $\xi^{(2)} = (\xi^{(2)}_1, \ldots, \xi^{(2)}_{m_2})$ such that $\left\|\widehat{L}_{G^{(2)}_{j,+}}\right\|_2$ and $\left\|\widehat{L}_{G^{(2)}_{j,-}}\right\|_2 \leqslant \frac{2\Delta}{d_{\min}\lambda}$.

Now, we first fix $\xi^{(2)} \in \{\pm 1\}^{m_2}$. Observe that we can write $\widehat{L}_H$ as

$$\widehat{L}_H = \sum_{i\in[m_1]} \mathbf{1}(\xi^{(1)}_i = +1)\cdot\widehat{L}_{G^{(1)}_{i,-}} + \mathbf{1}(\xi^{(1)}_i = -1)\cdot\widehat{L}_{G^{(1)}_{i,+}}\,, \tag{8.3}$$

and

$$\begin{aligned}\mathbb{E}[\widehat{L}_H|\xi^{(2)}] &= (1-\eta)\sum_{i\in[m_1]}\widehat{L}_{G^{(1)}_{i,-}} + \eta\sum_{i\in[m_1]}\widehat{L}_{G^{(1)}_{i,+}} \\ &= \sum_{e\in E}\left((1-\eta)\cdot\mathbf{1}(\xi^{(2)}_{c_2(e)}=-1)+\eta\cdot\mathbf{1}(\xi^{(2)}_{c_2(e)}=+1)\right)\cdot\widehat{L}_e \\ &:= \sum_{e\in E} w_{c_2(e)}\cdot\widehat{L}_e\,.\end{aligned} \tag{8.4}$$

Here $w_{c_2(e)}\in\{\eta,1-\eta\}$, thus $\left\|\mathbb{E}[\widehat{L}_H|\xi^{(2)}]\right\|_2 \geqslant \eta\left\|\sum_{e\in E}\widehat{L}_e\right\|_2 = \eta$.

We now split the analysis into two cases. Let $\eta_0 := 1/12$.

Case 1: $\eta \geqslant \eta_0$.

In light of Eq. (8.3), we define $X_i := \mathbf{1}(\xi^{(1)}_i=+1)\cdot\widehat{L}_{G^{(1)}_{i,-}} + \mathbf{1}(\xi^{(1)}_i=-1)\cdot\widehat{L}_{G^{(1)}_{i,+}}$ such that $\widehat{L}_H = \sum_{i\in[m_1]}X_i$. Moreover, we have that $X_i \succeq 0$ and $\|X\|_2 \leqslant \frac{2\Delta}{d_{\min}\lambda}$ almost surely from Eq. (8.2). Thus, applying matrix Chernoff (Fact 2.4.2), we get

$$\begin{aligned}\Pr_{\xi^{(1)}}\left[\left\|\widehat{L}_H\right\|_2 \geqslant (1+\delta)\left\|\mathbb{E}[\widehat{L}_H|\xi^{(2)}]\right\|_2\right] &\leqslant n\cdot\exp\left(-\frac{1}{3}\delta^2\left\|\mathbb{E}[\widehat{L}_H|\xi^{(2)}]\right\|_2\cdot\frac{d_{\min}\lambda}{2\Delta}\right) \\ &\leqslant n\cdot\exp\left(-\frac{\delta^2\eta d_{\min}\lambda}{6\Delta}\right),\end{aligned} \tag{8.5}$$

which is at most $O(n^{-2})$ as long as $\delta^2 \geqslant \frac{B_1\Delta\log n}{d_{\min}\lambda}$ for a large enough constant B_1.

Next, we similarly prove concentration for $\left\|\mathbb{E}[\widehat{L}_H|\xi^{(2)}]\right\|_2$ over $\xi^{(2)}$. Recalling Eq. (8.4),

$$\mathbb{E}[\widehat{L}_H|\xi^{(2)}] = \sum_{e\in E} w_{c_2(e)}\cdot\widehat{L}_e = \sum_{j\in[m_2]} w_j \sum_{e\in G^{(2)}_j}\widehat{L}_e = \sum_{j\in[m_2]} w_j\cdot\widehat{L}_{G^{(2)}_j}\,.$$

$\mathbb{E}[w_j] = 2\eta(1-\eta)$, and $\left\|\mathbb{E}_{\xi^{(2)}}\mathbb{E}[\widehat{L}_H|\xi^{(2)}]\right\|_2 = 2\eta(1-\eta)\left\|\sum_{e\in E}\widehat{L}_e\right\|_2 = 2\eta(1-\eta)$. Since $\left\|w_j\widehat{L}_{G^{(2)}_j}\right\|_2 \leqslant \frac{2(1-\eta)\Delta}{d_{\min}\lambda}$, we can apply matrix Chernoff again:

$$\Pr_{\xi^{(2)}}\left[\left\|\mathbb{E}[\widehat{L}_H|\xi^{(2)}]\right\|_2 \geqslant (1+\delta')\cdot 2\eta(1-\eta)\right] \leqslant n\cdot\exp\left(-\frac{1}{3}\delta'^2\cdot 2\eta(1-\eta)\cdot\frac{d_{\min}\lambda}{2(1-\eta)\Delta}\right) \tag{8.6}$$

which is at most $O(n^{-2})$ as long as $\delta'^2 \geqslant \frac{B_2\Delta\log n}{d_{\min}\lambda}$ for a large enough constant B_2. Combining both tail bounds, by the union bound, we have that with probability at least $1-O(n^{-2})$, $\left\|\widehat{L}_H\right\|_2 \leqslant (1+\delta)\cdot 2\eta(1-\eta)$ as long as $\delta^2 \geqslant \frac{B\Delta\log n}{d_{\min}\lambda}$ for a large enough B. This establishes Eq. (8.1), proving the lemma for this case.

Case 2: $\eta < \eta_0$. To handle this case, observe that the exact same analysis goes through for $\widetilde{H} = \{e\in E : \xi^{(1)}_{c_1(e)} = -1 \text{ or } \xi^{(2)}_{c_2(e)} = -1\} \supseteq H$. Indeed, similar to Eq. (8.3) and (8.4), we have $\widehat{L}_{\widetilde{H}} = \sum_{i\in[m_1]}\widetilde{X}_i$ where $\widetilde{X}_i = \mathbf{1}(\xi^{(1)}_i=+1)\cdot\widehat{L}_{G^{(1)}_{i,-}} + \mathbf{1}(\xi^{(1)}_i=-1)\cdot\widehat{L}_{G^{(1)}_i}$ (notice

the 2nd term is $G_i^{(1)}$ instead of $G_{i,+}^{(1)}$), and

$$\mathbb{E}[\widehat{L}_{\widetilde{H}}|\xi^{(2)}] = (1-\eta)\sum_{i\in[m_1]}\widehat{L}_{G_{i,-}^{(1)}} + \eta\sum_{i\in[m_1]}\widehat{L}_{G_i^{(1)}} = \sum_{e\in E}\widetilde{w}_{c_2(e)}\cdot\widehat{L}_e = \sum_{j\in[m_2]}\widetilde{w}_j\cdot\widehat{L}_{G_j^{(2)}},$$

where $\widetilde{w}_j = 1$ if $\xi_j^{(2)} = -1$ and η if $\xi_j^{(2)} = +1$, hence $\mathbb{E}[\widetilde{w}_j] = \eta + \eta(1-\eta) = \eta(2-\eta)$. Moreover, $\left\|\mathbb{E}_{\xi^{(2)}}\mathbb{E}[\widehat{L}_{\widetilde{H}}|\xi^{(2)}]\right\|_2 = \eta(2-\eta)\left\|\sum_{e\in E}\widehat{L}_e\right\|_2 = \eta(2-\eta)$.

First, set $\eta = \eta_0$, and let $\widetilde{H}_0$ be the random subgraph as defined above. Similar to Eq. (8.5) and (8.6), we apply matrix Chernoff (Fact 2.4.2) and get that with probability $1 - O(n^{-2})$, $\left\|\widehat{L}_{\widetilde{H}_0}\right\|_2 \leqslant (1+\delta)\cdot\eta_0(2-\eta_0)$ for $\delta = \sqrt{\frac{B\Delta\log n}{d_{\min}\lambda}} \leqslant 1$. In particular, this means that $L_{\widetilde{H}_0} \preceq 2\eta_0(2-\eta_0)L_G \preceq \frac{1}{3}L_G$ when $\eta_0 = 1/12$.

Now, fix any $\eta < \eta_0$. We can obtain a coupling between this case and the case when $\eta = \eta_0$ by randomly changing $\xi_i^{(1)}$ and $\xi_j^{(2)}$ from $+1$ to -1 (while not flipping the ones with -1). Notice that $\widetilde{H}$ is monotone increasing as we change any $+1$ to -1 (whereas H is not!), thus we must have $\widetilde{H} \subseteq \widetilde{H}_0$ in this coupling. Then, as $H \subseteq \widetilde{H}$, we have

$$L_H \preceq L_{\widetilde{H}} \preceq L_{\widetilde{H}_0} \preceq \frac{1}{3}L_G$$

with probability $1 - O(n^{-2})$. This finishes the proof of Lemma 8.4.7. □

8.4.5 Recovery of corrupted constraints from corrupted pairs

We have thus shown that, with probability $\geqslant 1 - 1/\operatorname{poly}(n)$, we can *exactly* recover the set of corrupted constraints within each expanding subinstance $\phi_1, \ldots, \phi_T$. Recall that after pruning and expander decomposition (Lemma 8.4.6), the expanding subinstances contain a $(1-\varepsilon)$-fraction of all edges in the instance ϕ, and the set of edges removed only depends on the constraint graph and not the right-hand sides of ϕ. As stated in Observation 8.4.2, the instance ϕ has exactly $m^2/4p$ edges, and they correspond exactly to the set $\{(u, C, C') : u \in [p], C \in \mathcal{H}_u^{(L)}, C' \in \mathcal{H}_u^{(R)}\}$, and moreover an edge e in ϕ is corrupted if and only if exactly one of the two constraints $(u, C), (u, C')$ is corrupted in the original instance ψ, where e corresponds to (u, C, C'). For each $u \in [p]$ and $C \in \mathcal{H}_u = \mathcal{H}_u^{(L)} \cup \mathcal{H}_u^{(R)}$, let $\xi_u(C) = -1$ if (u, C) is corrupted in ψ, and 1 otherwise. It thus follows that we have learned, for $1-\varepsilon$ fraction of all $\{(u, C, C') : u \in [p], C \in \mathcal{H}_u^{(L)}, C' \in \mathcal{H}_u^{(R)}\}$, the product $\xi_u(C)\cdot\xi_u(C')$.

It now remains to show how to recover $\xi_u(C)$ for most $u \in [p]$, $C \in \mathcal{H}_u$. For each $u \in [p]$, let $P_u \subseteq \{(C, C') : C \in \mathcal{H}_u^{(L)}, C' \in \mathcal{H}_u^{(R)}\}$ such that we have determined $\xi_u(C)\cdot\xi_u(C')$, and let $P = \cup_{u\in[p]}P_u$. We know that $|P| \geqslant (1-\varepsilon)\frac{m^2}{4p}$. Let ε_u be chosen so that $|P_u| = (1-\varepsilon_u)\frac{m^2}{4p^2}$, i.e., ε_u is the fraction of pairs in $\mathcal{H}_u^{(L)} \times \mathcal{H}_u^{(R)}$ that were deleted

in Lemma 8.4.6. Notice that we have

$$\begin{aligned}(1-\varepsilon)\frac{m^2}{4p} \leqslant |P| = \sum_{u\in[p]} |P_u| = \frac{m^2}{4p^2}\sum_{u\in[p]}(1-\varepsilon_u)\\ \implies \frac{1}{p}\sum_{u\in[p]}\varepsilon_u \leqslant \varepsilon\,.\end{aligned} \tag{8.7}$$

One can think of this problem as a collection of disjoint *satisfiable* (noiseless) 2-XOR instances on P_u, where each P_u is a biclique ($\frac{m}{2p}$ vertices on each side) with ε_u fraction of edges are removed.

Algorithm 8.4.8 (Recover corrupted constraints from corrupted pairs)**.**

Given: For each $u \in [p]$, a set $P_u \subseteq \mathcal{H}_u^{(L)} \times \mathcal{H}_u^{(R)}$ such that $|P_u| = (1-\varepsilon_u)\frac{m^2}{4p^2}$ for $\varepsilon_u \in [0,1]$, along with "right-hand sides" $\xi_u(C)\cdot\xi_u(C')$ for each $(C,C') \in P_u$.

Output: For each $u \in [p]$, disjoint subsets $\mathcal{A}_u^{(1)}, \mathcal{A}_u^{(2)} \subseteq \mathcal{H}_u$.

Operation:

1. **Initialize:** $\mathcal{A}_u^{(1)}, \mathcal{A}_u^{(2)} = \emptyset$ for each $u \in [p]$.
2. **For each** $u \in [p]$:
   (a) If $\varepsilon_u \geqslant 1/3$, set $\mathcal{A}_u^{(1)} = \mathcal{H}_u$ and $\mathcal{A}_u^{(2)} = \emptyset$.
   (b) Else if $\varepsilon_u < 1/3$, let G_u be the graph with vertex set $\mathcal{H}_u = \mathcal{H}_u^{(L)} \cup \mathcal{H}_u^{(R)}$ with edges given by P_u, and let S_u be the size of the largest connected component in G_u.
   (c) As S_u is connected in G_u, and we know $\xi_u(C)\xi_u(C')$ for each edge (C,C') in G_u, by solving a linear system of equations we obtain $z \in \{\pm 1\}^{\mathcal{H}_u}$ such that either $z_C = \xi_u(C)$ for all $C \in S_u$, or $z_C = -\xi_u(C)$ for all $C \in S_u$. That is, $z_C = \xi_u(C)$ up to a global sign.
   (d) Pick the global sign to minimize the number of $C \in S_u$ for which $z_C = -1$. Set $\mathcal{A}_u^{(1)} = \mathcal{H}_u \setminus S_u$ and $\mathcal{A}_u^{(2)} = \{C \in S_u : z_C = -1\}$.
3. Output $\{\mathcal{A}_u^{(1)}\}_{u\in[p]}$, $\{\mathcal{A}_u^{(2)}\}_{u\in[p]}$.

We now analyze Algorithm 8.4.8 via the following lemma.

Lemma 8.4.9. *Let $\eta \in [0,1/2)$, and let $|\mathcal{H}_u| = \frac{m}{p} \geqslant \frac{24k}{(1-2\eta)^2}\log n$ and $|P_u| = (1-\varepsilon_u)\frac{m^2}{4p^2}$ with $\varepsilon_u \in [0,1]$ for each $u \in [p]$, and $\frac{1}{p}\sum_{u\in[p]}\varepsilon_u \leqslant \varepsilon$. The outputs of Algorithm 8.4.8 satisfy the following: (1) $\sum_{u\in[p]}|\mathcal{A}_u^{(1)}| \leqslant 4\varepsilon m$, and (2) with probability $1-n^{-k}$ over the noise $\{\xi_u(C)\}_{u\in[p],C\in\mathcal{H}_u}$, for every $u \in [p]$ we have that $\mathcal{A}_u^{(2)} = \{C \in \mathcal{H}_u : \xi_u(C) = -1\} \setminus \mathcal{A}_u^{(1)}$.*

Proof. Suppose that $\varepsilon_u < 1/3$. Observe that G_u is a graph obtained by taking a biclique with left vertices $\mathcal{H}_u^{(L)}$ and right vertices $\mathcal{H}_u^{(R)}$, i.e., with $m/2p$ left vertices and $m/2p$

right vertices. The following lemma shows that the largest connected component S_u in G_u has size at least $\frac{m}{p}(1-\varepsilon_u)$.

Claim 8.4.10. *Let $K_{n,n}$ be the complete bipartite graph with n left vertices L and n right vertices R. Let G be a graph obtained by deleting εn^2 edges from $K_{n,n}$. Then, the largest connected component in G has size $\geqslant 2n(1-\varepsilon)$.*

We postpone the proof of Claim 8.4.10 to the end of the section, and continue with the proof of Lemma 8.4.9.

We now argue that we can efficiently obtain the vector z in Step (2c) of Algorithm 8.4.8. Indeed, this is done as follows. First, pick one $C_0 \in S_u$ arbitrarily, and set $z_{C_0} = 1$. Then, we propagate in a breadth-first search manner: for any edge (C, C') in S_u where z_C is determined, set $z_{C'} = z_C \cdot \xi_u(C)\xi_u(C')$. We repeat this process until we have labeled all of S_u. Notice that as S_u is a connected component, fixing z_{C_0} for any $C_0 \in S_u$ uniquely determines the assignment of all S_u, thus we have obtained $z_C = \xi_u(C)$ up to a global sign.

Now, we observe that S_u does not depend on the noise in ψ. Indeed, this is because the pruning and expander decomposition (and thus the graph G_u) depends solely on the constraint graph G of the instance ϕ, and not on the right-hand sides of the constraints. The following lemma thus shows that with high probability over the noise, the number of $C \in S_u$ where $\xi_u(C) = -1$ is strictly less than $1/2|S_u|$. Hence, in Step (2d), by picking the assignment $\pm z$ that minimizes the number of $C \in S_u$ with $\xi_u(C) = -1$, we see that $\mathcal{A}_u^{(2)} = \{C \in S_u : z_C = -1\} = \{C \in S_u : \xi_u(C) = -1\}$.

Claim 8.4.11. *Let $\eta \in (0, 1/2)$ be the corruption probability, and assume that $p \leqslant n^k$ and $\frac{m}{p} \geqslant \frac{24k}{(1-2\eta)^2} \log n$. With probability $1 - n^{-k}$ over the noise in ψ, it holds that for each $u \in [p]$ with $\varepsilon_u < 1/3$, $|\{C \in S_u : \xi_u(C) = -1\}| < \frac{1}{2}|S_u|$.*

We postpone the proof of Claim 8.4.11, and finish the proof of Lemma 8.4.9. We next bound $\sum_{u\in[p]} |\mathcal{A}_u^{(1)}|$. By Eq. (8.7) we have that $\frac{1}{p}\sum_u \varepsilon_u \leqslant \varepsilon$. Thus,

$$\sum_{u:\varepsilon_u \geqslant 1/3} |\mathcal{H}_u| \leqslant \frac{m}{p} \sum_{u:\varepsilon_u \geqslant 1/3} 3\varepsilon_u \leqslant 3\varepsilon m\,.$$

Moreover, by Claim 8.4.10 we have $|S_u| \geqslant (1-\varepsilon_u)|\mathcal{H}_u| = (1-\varepsilon_u)\frac{m}{p}$. Thus,

$$\sum_{u:\varepsilon_u < 1/3} |\mathcal{H}_u \setminus S_u| \leqslant \sum_{u:\varepsilon_u < 1/3} \varepsilon_u \cdot \frac{m}{p} \leqslant \varepsilon m\,.$$

Therefore, combining the two,

$$\sum_{u\in[p]} |\mathcal{A}_u^{(1)}| = \sum_{u:\varepsilon_u < 1/3} |\mathcal{H}_u \setminus S_u| + \sum_{u:\varepsilon_u \geqslant 1/3} |\mathcal{H}_u| \leqslant 4\varepsilon m\,,$$

which finishes the proof of Lemma 8.4.9. □

In the following, we prove Claims 8.4.10 and 8.4.11.

Proof of Claim 8.4.10. Let $S_1, \dots, S_t$ be the connected components of G. Let $\ell_i = |S_i \cap L|$ and $r_i = |S_i \cap R|$. The number of edges in G is at most $\sum_{i=1}^t \ell_i r_i$.

Now, suppose that the largest connected component of G has size at most M. Then, we have that $\ell_i + r_i \leqslant M$ for all $i \in [t]$. Notice that the number of edges deleted from $K_{n,n}$ to produce G must be at least $n^2 - \sum_{i=1}^t \ell_i r_i$, and this is at most εn^2. Hence, by maximizing the quantity $\sum_{i=1}^t \ell_i r_i$ subject to $\ell_i + r_i \leqslant M$ for all $i \in [t]$ and $\sum_{i=1}^t \ell_i + r_i = 2n$, we can obtain a lower bound on the number of edges deleted from $K_{n,n}$ in order for the largest connected component of G to have size at most M. We have that

$$\sum_{i=1}^t \ell_i r_i \leqslant \sum_{i=1}^t \left(\frac{\ell_i + r_i}{2}\right)^2 \leqslant \frac{M}{2} \cdot \sum_{i=1}^t \frac{\ell_i + r_i}{2} = \frac{nM}{2},$$

where the first inequality is by the AM-GM inequality. Thus,

$$\varepsilon n^2 \geqslant n^2 - \frac{nM}{2} \implies M \geqslant 2n(1-\varepsilon),$$

which finishes the proof. □

Proof of Claim 8.4.11. Let u be such that $\varepsilon_u < 1/3$, and let S_u be the largest connected component in G_u. Observe that S_u is determined solely by the constraint graph of ϕ, and in particular does not depend on the noise in ϕ (and hence on the noise in ψ). As $p \leqslant n^k$ by assumption, it thus suffices to show that for each $u \in [p]$, with probability $1 - n^{-2k}$ it holds that $|\{C \in S_u : \xi_u(C) = -1\}| < \frac{1}{2}|S_u|$. Notice that $|\{C \in S_u : \xi_u(C) = -1\}|$ is simply the sum of $|S_u|$ Bernoulli(η) random variables. By Hoeffding's inequality, with probability $\geqslant 1 - \exp(-2\delta^2|S_u|)$ it holds that $|\{C \in S_u : \xi_u(C) = -1\}| \leqslant (\eta + \delta)|S_u|$. We choose $\delta = \frac{1}{2}(\frac{1}{2} - \eta)$ such that $\eta + \delta < \frac{1}{2}$ for $\eta \in (0, \frac{1}{2})$. Then, by noting that $2\delta^2|S_u| \geqslant 2\delta^2(1-\varepsilon_u)|\mathcal{H}_u| \geqslant \frac{1}{2}(\frac{1}{2} - \eta)^2 \cdot \frac{2}{3} \cdot \frac{m}{p} \geqslant 2k \log n$ since $\frac{m}{p} \geqslant \frac{24k}{(1-2\eta)^2} \log n$, Claim 8.4.11 follows. □

8.4.6 Finishing the proof of Lemma 8.3.2

Proof of Lemma 8.3.2. We are given an τ-spread p-bipartite k-XOR instance ψ with constraint graph $\mathcal{H} = \{\mathcal{H}_u\}_{u \in [p]}$, where we recall from Definition 8.3.1 that (1) $m = |\mathcal{H}|$ and each $|\mathcal{H}_u| = \frac{m}{p} \geqslant 2\lfloor \frac{1}{2\tau^2} \rfloor$ and $\frac{m}{p}$ is even, and (2) for any $Q \subseteq [n]$, $\deg_u(Q) \leqslant \frac{1}{\tau^2} \max(1, n^{\frac{k}{2}-1-|Q|})$. For convenience, let $m \geqslant n^{\frac{k-1}{2}} \sqrt{p} \cdot \beta$ where $\beta := C \cdot \frac{(k \log n)^{3/2}}{\tau \gamma^2 \varepsilon^{3/2}}$ and $\gamma := 1 - 2\eta \in (0, 1]$ since $\eta \in [0, \frac{1}{2})$.

First, we construct the 2-XOR instance ϕ defined in Definition 8.4.1. As stated in Observation 8.4.2, the average degree is at least $d := \frac{1}{4}\beta^2$, and furthermore, by Lemma 8.4.5,

the maximum degree of $G^{(L)}_{u,C}(\phi)$ and $G^{(R)}_{u,C'}(\phi)$ for any $u \in [p]$, $C \in \mathcal{H}^{(L)}_u$ and $C' \in \mathcal{H}^{(R)}_u$ is bounded by $\Delta := 1/\tau^2$. The algorithm then follows the steps outlined in Section 8.4.2.

Step 1. We apply graph pruning and expander decomposition (Lemma 8.4.6) with parameter $\varepsilon' := \frac{1}{4}\varepsilon$, which decomposes ϕ into $\phi_1, \dots, \phi_T$ such that they contain $1-\varepsilon'$ fraction of the constraints in ϕ, and their constraint graphs (after adding some self-loops due to expander decomposition) have minimum degree $d_{\min} \geqslant \frac{1}{3}\varepsilon' d = \frac{1}{48}\varepsilon\beta^2$ and spectral gap $\lambda \geqslant \Omega(\varepsilon'^2/\log^2 m) = \Omega(\varepsilon^2/(k^2\log^2 n))$.

Step 2. We solve the SDP relaxation for each subinstance ϕ_i. Let G be the constraint graph of ϕ_i (with at most $N \leqslant n^{k-1}$ vertices) and H be the corrupted edges of G. We apply the relative spectral approximation result (Lemma 8.4.7) with $\xi^{(1)}_1, \dots, \xi^{(1)}_{m/2p}$ (resp. $\xi^{(2)}_1, \dots, \xi^{(2)}_{m/2p}$) being $\{\pm 1\}$ random variables indicating whether each $C \in \mathcal{H}^{(L)}_u$ (resp. $C' \in \mathcal{H}^{(R)}_u$) is corrupted. Moreover, the subgraphs $G^{(1)}_i$ and $G^{(2)}_j$ in Lemma 8.4.7 (which are simply subgraphs of $G^{(L)}_{u,C}(\phi)$ and $G^{(R)}_{u,C'}(\phi)$) have maximum degree $\leqslant \Delta = 1/\tau^2$. Thus, we have that with probability $1 - O(N^{-2})$,

$$L_H \preceq \max\left((1+\delta)\cdot 2\eta(1-\eta), \frac{1}{3}\right)\cdot L_G$$

where $\delta = \sqrt{\frac{B\Delta\log N}{d_{\min}\lambda}} \leqslant O\left(\sqrt{\frac{k^3\log^3 n}{\tau^2\varepsilon^3\beta^2}}\right)$. Plugging in β (for large enough C), we get that $\delta \leqslant \gamma^2 = 1 - 4\eta(1-\eta)$. Therefore, we have $(1+\delta)\cdot 2\eta(1-\eta) \leqslant (1+\gamma^2)\cdot\frac{1}{2}(1-\gamma^2) < \frac{1}{2}$, hence $L_H \prec \frac{1}{2}L_G$. By union bound over all $T \leqslant N$ subinstances, this holds for all subinstances ϕ_i with probability $1 - \frac{1}{\mathrm{poly}(n)}$ over the randomness of the noise.

Then, by Lemma 8.1.4, the SDP relaxation has a unique optimum which is the planted assignment. Thus, we can identify the set of corrupted edges in each ϕ_i.

Step 3. So far we have identified, for $\geqslant 1-\varepsilon'$ fraction of all $\{(u, C, C') : u \in [p], C \in \mathcal{H}^{(L)}_u, C' \in \mathcal{H}^{(R)}_u\}$, the product $\xi_u(C)\cdot\xi_u(C')$, where $\xi_u(C) = -1$ if (u, C) is corrupted in ψ, and $+1$ otherwise. Let $P_u \subseteq \{(C, C') : C \in \mathcal{H}^{(L)}_u, C' \in \mathcal{H}^{(R)}_u\}$ be such pairs for each $u \in [p]$, and let $P = \cup_{u\in[p]} P_u$. Note that $|P| \geqslant (1-\varepsilon')\frac{m^2}{4p}$ and P depends only on $\mathcal{H}$ and not on the noise.

We then run Algorithm 8.4.8. By the assumption that $\tau \leqslant \frac{c\gamma}{\sqrt{k\log n}}$ for a small enough c, we have $|\mathcal{H}_u| = \frac{m}{p} \geqslant 2\lfloor\frac{1}{2\tau^2}\rfloor \geqslant \frac{24k}{(1-2\eta)^2}$, which is the condition we need in Lemma 8.4.9. Thus, with probability $1 - n^{-k}$, Algorithm 8.4.8 outputs (1) $\mathcal{A}_1 \subseteq \mathcal{H}$ which only depends on $\mathcal{H}$ and such that $|\mathcal{A}_1| \leqslant 4\varepsilon' m = \varepsilon m$, and (2) $\mathcal{A}_2 \subseteq \mathcal{H}$, the set of corrupted constraints in $\mathcal{H}\setminus\mathcal{A}_1$. This completes the proof of Lemma 8.3.2. □

8.5 Notions of relative approximation

We have encountered several notions of relative graph approximations. Let G be an n-vertex graph, and let H be a random subgraph of G by selecting each edge with a fixed probability $\eta \in (0,1)$. We are interested in the sufficient conditions on G for each of the following to hold with probability $1 - o(1)$ (for some $\delta = o(1)$):

(1) **Relative cut approximation**: $x^\top L_H x \leqslant (1+\delta)\eta \cdot x^\top L_G x$ for all $x \in \{\pm 1\}^n$.

(2) **Relative SDP approximation**: $\langle X, L_H \rangle \leqslant (1+\delta)\eta \cdot \langle X, L_G \rangle$ for all symmetric matrices $X \succeq 0$ with $\mathrm{diag}(X) = \mathbb{I}$.

(3) **Relative spectral approximation:** $L_H \preceq (1+\delta)\eta \cdot L_G$.

Here, we only state one-sided inequalities, as solving noisy XOR requires only an upper bound on L_H. Note also that the above is in increasing order: relative spectral approximation implies relative SDP approximation, which in turn implies relative cut approximation.

Recall from Lemma 8.1.3 that a lower bound on the min-cut of G suffices for cut approximation to hold, while Lemma 8.1.5 shows that lower bounds on the minimum degree and spectral gap of G suffice for spectral approximation to hold. It is natural to wonder whether a min-cut lower bound is sufficient for SDP approximation as well, since it allows us to efficiently recover the planted assignment in a noisy planted 2-XOR via solving an SDP relaxation (see Lemma 8.1.4). Unfortunately, there is a counterexample.

Separation of cut and SDP approximation. The example is the same graph that separates cut and spectral approximation described in [ST11]. Let n be even and $k = k(n)$. Define $G = (V, E)$ be a graph on $N = nk$ vertices where $V = \{0, 1, \ldots, n-1\} \times \{1, \ldots, k\}$ and (u, i), $(v, j) \in V$ are connected if $v = u \pm 1 \mod n$. Moreover, there is one additional edge e^* between $(0, 1)$ and $(n/2, 1)$. In other words, G consists of n clusters of vertices of size k, where the clusters form a ring with a complete bipartite graph between adjacent clusters, along with a special edge e^* in the middle.

Clearly, the minimum cut of G is $2k$, which means that cut approximation holds. Essentially, the special edge e^* does not play a role here.

However, we will show that e^* breaks SDP approximation. Define vector $x_0 \in \mathbb{R}^V$ such that the (u, i) entry is

$$x_0(u, i) = \min(u, n - u)\,,$$

and vectors $x_1, \ldots, x_{n-1}$ to be cyclic shifts of x_0: for $w \in \{0, 1, \ldots, n-1\}$,

$$x_w(u, i) = x_0(u - w \pmod n,\ i)\,.$$

We note that x_0 is the vector shown in [ST11] that breaks spectral approximation. We now show that $X = \sum_{w=0}^{n-1} x_w x_w^\top$ (scaled so that X has all 1s on the diagonal) breaks SDP approximation.

First, it is easy to see that the diagonal entries of X are all equal due to symmetry. Thus, for some scaling c, $cX \succeq 0$ and $\mathrm{diag}(cX) = \mathbb{I}$.

Observe that for $w \leqslant \frac{n}{2} - 1$, $x_w(0,1) = w$ and $x_w(\frac{n}{2},1) = \frac{n}{2} - w$. For $w \geqslant \frac{n}{2}$, $x_w(0,1) = n - w$ and $x_w(\frac{n}{2},1) = w - \frac{n}{2}$. Thus, as $x_w^\top L_{e^*} x_w = \big(x_w(0,1) - x_w(\frac{n}{2},1)\big)^2$,

$$\langle X, L_{e^*} \rangle = \sum_{w=0}^{n-1} x_w^\top L_{e^*} x_w = \sum_{w=0}^{\frac{n}{2}-1} \left(\frac{n}{2} - 2w\right)^2 + \sum_{w=\frac{n}{2}}^{n-1} \left(\frac{3n}{2} - 2w\right)^2 = \Theta(n^3)\,.$$

On the other hand, $x_w^\top L_{G\setminus e^*} x_w = nk^2$ for any w, thus $\langle X, L_{G\setminus e^*} \rangle = n^2k^2$. This is $o(n^3)$, i.e. dominated by $\langle X, L_{e^*} \rangle$, when $k = o(\sqrt{n})$. Since e^* is selected in H with probability η, we have that with probability η,

$$\langle X, L_H \rangle \geqslant \langle X, L_{e^*} \rangle \geqslant (1 - o(1)) \cdot \langle X, L_G \rangle\,,$$

which violates the desired SDP approximation.

Chapter 9

Rounding Large Independent Sets on Expanders

In this chapter, we prove Theorems 6.3.2 and 6.3.3, which we restate below.

Theorem (Restatement of Theorem 6.3.2)**.** *There is a polynomial-time algorithm that, given an n-vertex regular 10^{-4}-almost 3-colorable graph with normalized 2nd eigenvalue $\lambda_2 \leqslant 10^{-4}$, finds an independent set of size $\geqslant 10^{-4}n$.*

Theorem (Restatement of Theorem 6.3.3)**.** *For every positive $\varepsilon \leqslant 0.001$, there is a polynomial-time algorithm that, given an n-vertex regular graph that contains an independent set of size $(\frac{1}{2} - \varepsilon)n$ and has normalized 2nd eigenvalue $\lambda_2 \leqslant 1 - 40\varepsilon$, outputs an independent set of size at least $10^{-3}n$.*

We also prove the following hardness result for almost 4-colorable expanders.

Proposition (Restatement of Proposition 6.3.1)**.** *Assuming the Unique Games Conjecture, for any constants $\varepsilon, \gamma > 0$, it is NP-hard to find an independent set of size γn in an n-vertex regular graph that is ε-almost 4-colorable and has normalized 2nd eigenvalue $\lambda_2 \leqslant o_n(1)$.*

Organization. In Section 9.1, we give a technical overview of our algorithmic results. More specifically, we give full proofs of our combinatorial *clustering* results about independent sets and colorings in expanders (Lemmas 9.1.2 and 9.1.6) in Sections 9.1.2 and 9.1.3, as well as an overview of how they lead to our rounding algorithms.

In Section 9.2, we prove the algorithmic result for graphs containing $(\frac{1}{2} - \varepsilon)n$-sized independent sets (Theorem 6.3.3). Then, in Section 9.3, we prove the result for almost 3-colorable graphs (Theorem 6.3.2).

On the hardness side, in Section 9.4, we prove Proposition 6.3.1 and also similar hardness results for *exactly* k-colorable expanders (as opposed to almost k-colorable), assuming a variant of the 2-to-1 conjecture [Kho02, DMR06].

Finally, in Section 9.5, we give a proof of a folklore result: given a graph containing an independent set of size $(\frac{1}{2} - \varepsilon)n$, one can find an independent set of size at least $(\varepsilon n)^{1-O(\varepsilon)}$. The algorithm is a variant of the rounding algorithm by Karger, Motwani and Sudan [KMS98].

Notation. In this chapter, we will use $\lambda_2(G)$ to denote the normalized 2nd eigenvalue $\lambda_2(\widetilde{A}_G)$, where $\widetilde{A}_G$ is the normalized adjacency matrix of the graph G.

9.1 Technical overview

We provide a brief overview of our rounding framework and analysis in this section. In Section 9.1.1, we briefly discuss the clustering property and how it leads to our rounding algorithm for one-sided spectral expanders. Then, we describe the proof of the clustering property of independent sets in Section 9.1.2 and the clustering property of 3-colorings in Section 9.1.3. The rounding for 3-colorable graphs follows a similar rounding framework, and we refer the reader to Section 9.3 for details.

Polynomial Formulation and SoS Relaxation. Our algorithm rounds a constant-degree sum-of-squares relaxations (see Section 2.5 for background) of the following system of polynomial inequalities that encode independent sets of size $\geqslant (1/2 - \varepsilon)n$ in the input graph on $G(V, E)$.

$$\begin{aligned} \frac{1}{n}\sum_{u\in V} x_u &\geqslant \frac{1}{2} - \varepsilon, \\ x_u^2 &= x_u\,, \qquad \forall u \in V\,, \\ x_u x_v &= 0\,, \qquad \forall \{u,v\} \in E. \end{aligned} \tag{9.1}$$

The relaxation outputs a *pseudo-distribution* over solutions to (9.1). For a reader unfamiliar with the sum-of-squares method for algorithm design, it is helpful to think of μ as constant-degree moments (i.e., expectations under μ of any constant-degree polynomial of x) of a probability distribution over $x \in \{0,1\}^n$ satisfying (9.1).

We will repeatedly use the following simple fact.

Fact 9.1.1. *For a graph $G = (V, E)$, let μ be a pseudo-distribution of degree at least 2 that satisfies the independent set constraints, i.e., $x_u^2 = x_u$ for all $u \in V$ and $x_u x_v = 0$ for all $\{u, v\} \in E$. Then, the set of vertices $\{u \in V : \widetilde{\mathbb{E}}_\mu[x_u] > \frac{1}{2}\}$ forms an independent set in G.*

Proof. For all $\{u, v\} \in E$, from the independent set constraints we can derive that $(x_u + x_v)^2 = x_u^2 + 2x_u x_v + x_v^2 = x_u + x_v$, i.e., $(x_u + x_v)$ satisfies the Booleanity constraint, thus $x_u + x_v \leqslant 1$. Thus, we have $\widetilde{\mathbb{E}}_\mu[x_u + x_v] \leqslant 1$, which means that u, v cannot both be in the set $\{u \in V : \widetilde{\mathbb{E}}_\mu[x_u] > \frac{1}{2}\}$. □

9.1.1 Rounding large independent sets on one-sided spectral expanders

Let G be any regular one-sided spectral expander with $\lambda_2(G) \leqslant 1 - O(\varepsilon)$ containing an independent set of size $(1/2 - \varepsilon)n$. Our approach can be summarized as follows:

(1) **An extremal clustering property of independent sets:** We show (in Lemma 9.1.2) that there are only two *essentially distinct* $(1/2 - \varepsilon)n$-sized independent sets in G. Specifically, given any three independent sets $x^{(1)}, x^{(2)}, x^{(3)}$, at least two of them have a non-trivially large intersection[1] i.e., $\mathbb{E}_u[x_u^{(i)} x_u^{(j)}] > 1/2 - \nu$ for some $\nu \approx 0$ and $i \neq j \in [3]$.

(2) **Recasting large intersection as a polynomial inequality:** Given any three $(1/2 - \varepsilon)n$-sized independent sets, $\boldsymbol{x} := (x^{(1)}, x^{(2)}, x^{(3)})$, we define $\Phi(\boldsymbol{x}) := \mathbb{E}_u[x_u^{(1)} x_u^{(2)}]^2 + \mathbb{E}_u[x_u^{(2)} x_u^{(3)}]^2 + \mathbb{E}_u[x_u^{(1)} x_u^{(3)}]^2$ which is at least $(1/2 - \nu)^2 \geqslant 1/4 - \nu$ as a consequence of (1). Here, $\mathbb{E}_u$ is the average with respect to a uniformly random $u \in [n]$ and thus $\Phi(\boldsymbol{x})$ measures the (squared) average pairwise intersections between $x^{(1)}, x^{(2)}, x^{(3)}$.

(3) **A low-degree sum-of-squares proof of largeness of $\Phi(x)$:** We show how the above property can be "SoS-ized". That is, $\Phi(\boldsymbol{x})$ is large in expectation over $x^{(1)}, x^{(2)}, x^{(3)}$ drawn independently from any pseudo-distribution μ satisfying the independent set constraints in Eq. (9.1), i.e., $\Phi(\mu) := \widetilde{\mathbb{E}}_{\boldsymbol{x} \sim \mu^{\otimes 3}}[\Phi(\boldsymbol{x})] \geqslant 1/4 - \nu$.

(4) **Rounding:** We give a simple rounding algorithm for μ with analysis relying on (3) to obtain a large independent set in G.

Our rounding analysis actually works as long as the intersection in (1) is *non-trivially larger than expected*, i.e., intersection $\geqslant (1/4 + \nu)n$, where $n/4$ is the expected intersection between random sets of size $\approx n/2$. However, we would need a different function $\Phi(\boldsymbol{x})$. For the sake of simplicity, we stick to the case where the intersection is $\geqslant 1/2 - \nu$.

Our final rounding algorithm relies on the idea of *rounding from multiple samples* from a pseudo-distribution first introduced in [BBKSS21]. In their application for rounding Unique Games on certified small-set expanders, they considered a certain "shift-partition potential" (which measured the correlation between two solutions for the input UG instance). Our analysis will rely instead on the above "average agreement function" $\Phi(\mu)$.

We will prove the clustering property stated in (1) in Section 9.1.2. Here, let us see how to round when $\Phi(\mu)$ is large.

Rounding when $\Phi(\mu)$ is large. In order to understand the intuition behind our rounding, notice that for 3 *random* subsets of $[n]$ of size $(1/2 - \varepsilon)n$, the pairwise agreement function Φ would be $\approx 3 \cdot (1/4)^2 = 3/16$. Thus, if $\Phi(\mu) \geqslant 1/4 - \nu > 3/16$ for some

[1] Throughout this paper, we write $\mathbb{E}_u$ to denote the expectation over a uniformly random vertex $u \in [n]$ of the input graph. Notice that $\mathbb{E}_u[x_u x'_u]$ then equals $\langle x, x' \rangle / n$.

small ν, then three draws from μ must be non-trivially correlated. We interpret this property as saying that the (pseudo)-distribution μ is "supported" over only two "distinct" independent sets. Concretely,

$$\widetilde{\mathbb{E}}_{\boldsymbol{x}\sim\mu^{\otimes 3}}[\Phi(\boldsymbol{x})] = 3\cdot\widetilde{\mathbb{E}}_{x^{(1)},x^{(2)}\sim\mu}\Big[\mathbb{E}_u[x_u^{(1)}x_u^{(2)}]^2\Big] \geqslant 1/4-\nu\,,$$

implying that $\widetilde{\mathbb{E}}_{x^{(1)},x^{(2)}}[\mathbb{E}_u[x_u^{(1)}x_u^{(2)}]^2] \geqslant 1/12 - O(\nu) > 1/16+\eta$ for a constant $\eta>0$ if ν is a small enough constant. Using the independence of $x^{(1)}$ and $x^{(2)}$ this resolves to:

$$\widetilde{\mathbb{E}}_{x^{(1)},x^{(2)}\sim\mu}\Big[\mathbb{E}_u[x_u^{(1)}x_u^{(2)}]^2\Big] = \widetilde{\mathbb{E}}\Big[\mathbb{E}_{u,v}[x_u^{(1)}x_v^{(1)}x_u^{(2)}x_v^{(2)}]\Big] = \mathbb{E}_{u,v}\Big[\widetilde{\mathbb{E}}[x_ux_v]^2\Big] > 1/16+\eta\,.$$

The classical idea of *rounding by conditioning* SoS solutions now suggests that we may be able to condition μ to obtain a μ' that is essentially supported on a *unique* assignment. Concretely, we argue that by applying a certain repeated conditioning procedure (that reduces "global correlation" [BRS11, RT12]; see Lemma 2.5.7) we obtain a modified pseudo-distribution μ' that satisfies all the original constraints and, in addition, satisfies that for most pairs of vertices $u,v\in[n]$, we have $\widetilde{\mathbb{E}}_{\mu'}[x_ux_v] \approx \widetilde{\mathbb{E}}_{\mu'}[x_u]\widetilde{\mathbb{E}}_{\mu'}[x_v]$ (where the approximation hides additive constant errors). Thus,

$$\mathbb{E}_{u,v}\Big[\widetilde{\mathbb{E}}_{\mu'}[x_ux_v]^2\Big] \approx \mathbb{E}_{u,v}\Big[\widetilde{\mathbb{E}}_{\mu'}[x_u]^2\widetilde{\mathbb{E}}_{\mu'}[x_v]^2\Big] = \mathbb{E}_u\Big[\widetilde{\mathbb{E}}_{\mu'}[x_u]^2\Big]^2 > 1/16+\eta\,.$$

We interpret this as saying that two independent "samples" from the (pseudo)-distribution μ' have a larger intersection than the intersection that random sets (of the same size) typically have: $\widetilde{\mathbb{E}}_{x^{(1)},x^{(2)}\sim\mu'}[\mathbb{E}_u[x_u^{(1)}x_u^{(2)}]] = \mathbb{E}_u[\widetilde{\mathbb{E}}_{\mu'}[x_u]^2] > 1/4+\Omega(\eta)$. An averaging argument now yields that $\widetilde{\mathbb{E}}_{\mu}[x_u] > \frac{1}{2}$ for at least an $\Omega(\eta)$ fraction of the vertices. This subset forms an independent set (see Fact 9.1.1) of size $\Omega(\eta n)$.

9.1.2 Clustering of independent sets in one-sided expanders

Let us now return to the combinatorial guts of our approach. We present in full here, a proof of the following extremal combinatorics statement that eventually can be imported into the low-degree sum-of-squares proof system.

Lemma 9.1.2. *Let G be a regular graph containing an independent set of size $(\frac{1}{2}-\varepsilon)n$ and has $\lambda_2(G)\leqslant 1-C\varepsilon$ for any small enough ε and some large enough constant $C>0$. Then, for any 3 independent sets of size at least $(\frac{1}{2}-\varepsilon)n$, two of them have an intersection of size $\geqslant(\frac{1}{2}-O(\frac{1}{C})-\varepsilon)n$.*

Intersection between 2 independent sets. Let us analyze the intersection between 2 independent sets I_1, I_2 (indicated by $x,y\in\{0,1\}^n$) in G. By assigning every vertex u of G the label (x_u,y_u), we obtain a partition of vertices of G into subsets with labels

in $\{00, 01, 10, 11\}$. Consider now a graph on the label set of 4 vertices and add an edge between two such labels, say ℓ_1, ℓ_2 (including self-loops) if there are vertices u with label ℓ_1, v with label ℓ_2 such that $\{u, v\} \in E$ (see Figure 9.1).

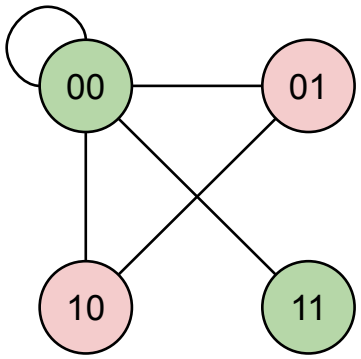


Figure 9.1: The gadget for 2 independent sets.

No edges can exist between $01, 10$ and 11 because x, y indicate independent sets. There can, however, be edges between vertices in the set 00, hence the self-loop. The graph in Figure 9.1 is the tensor product $H \otimes H$ where H is graph on $\{0, 1\}$ and edges $\{0, 0\}$ and $\{0, 1\}$.

Let $\mathsf{wt}(ij)$ denote the fraction of vertices u in G such that $x_u = i$ and $y_u = j$. Then, $|I_1 \cap I_2| = \mathsf{wt}(11)$. Let us now observe:

Claim 9.1.3. $\mathsf{wt}(00) \leqslant \mathsf{wt}(11) + 2\varepsilon$.

Proof. Since $|I_1|, |I_2| \geqslant (\frac{1}{2} - \varepsilon)n$, we have: $\mathsf{wt}(11) + \mathsf{wt}(10)$ and $\mathsf{wt}(11) + \mathsf{wt}(01) \geqslant \frac{1}{2} - \varepsilon$. Thus, $\mathsf{wt}(01) + \mathsf{wt}(10) + 2\mathsf{wt}(11) \geqslant 1 - 2\varepsilon$. We use $\mathsf{wt}(00) + \mathsf{wt}(01) + \mathsf{wt}(10) + \mathsf{wt}(11) = 1$ to finish. □

Let's now see how expansion of G enters the picture:

Claim 9.1.4. *Fix $\varepsilon > 0$ small enough. If $\lambda_2 \leqslant 1 - C\varepsilon$ for some large enough constant $C > 0$, then either $\mathsf{wt}(00) + \mathsf{wt}(11) \leqslant O(\frac{1}{C})$ or $\mathsf{wt}(11) \geqslant \frac{1}{2} - O(\frac{1}{C}) - \varepsilon$.*

Proof. For any subset $S \subseteq V$, we have $e(S, \overline{S}) \geqslant (1 - \lambda_2) \cdot (|S|/n)(1 - |S|/n)$. Here, $e(S, \overline{S})$ denotes the *weight* of edges between S and $\overline{S}$, which is $|E_G(S, \overline{S})|/nd$ for a d-regular graph. Applying this to the set of vertices with labels in $\{00, 11\}$, we have: $e(00, 01) + e(00, 10) \geqslant (1 - \lambda_2) \cdot \mathsf{wt}(\{00, 11\})(1 - \mathsf{wt}(\{00, 11\}))$.

On the other hand, since G is regular, $\mathsf{wt}(11) = e(00, 11)$ as there are no edges between $01, 10$ and 11. Similarly, we have $\mathsf{wt}(00) = \sum_{\alpha \in \{0,1\}^2} e(00, \alpha)$. Subtracting the two, we get $\mathsf{wt}(00) - \mathsf{wt}(11) = e(00, 00) + e(00, 01) + e(00, 10) \geqslant e(00, 01) + e(00, 10)$.

Therefore, we have

$$\begin{aligned}(1-\lambda_2)\cdot \mathsf{wt}(\{00,11\}))(1-\mathsf{wt}(\{00,11\})) &\leqslant e(00,01)+e(00,10)\\ &\leqslant \mathsf{wt}(00)-\mathsf{wt}(11)\leqslant 2\varepsilon\,.\end{aligned}$$

Thus, if $\lambda_2 \leqslant 1 - C\varepsilon$ for some large enough constant C, then either $\mathsf{wt}(\{00,11\}) \leqslant \eta$ or $\mathsf{wt}(\{00,11\}) \geqslant 1-\eta$ for $\eta = O(1/C)$. In the latter case, since $\mathsf{wt}(00) \leqslant \mathsf{wt}(11) + 2\varepsilon$, we have $\mathsf{wt}(11) \geqslant \frac{1}{2} - O(\frac{1}{C}) - \varepsilon$. □

Proof of Lemma 9.1.2. Let's now consider 3 independent sets. We can now naturally partition the vertices of G into 8 subsets labeled by elements of $\{0,1\}^3$. In the following, we will use "$*$" to denote both possible values. For example, $00*$ means $\{000, 001\}$.

From Claim 9.1.4, we know that $\mathsf{wt}(00*) + \mathsf{wt}(11*)$ (and analogously $\mathsf{wt}(0*0) + \mathsf{wt}(1*1)$ and $\mathsf{wt}(*00) + \mathsf{wt}(*11)$) is either $\leqslant O(\frac{1}{C}) < \frac{1}{3}$ or $\geqslant 1 - O(\frac{1}{C})$ for a large enough constant C. We now argue that the first possibility cannot simultaneously hold for all three pairs, and thus at least one pair of independent sets must have an intersection of at least $\frac{1}{2} - O(\frac{1}{C}) - \varepsilon$, completing the proof. Indeed, $\{00*, 11*\} \cup \{0*0, 1*1\} \cup \{*00, *11\}$ covers all strings $\{0,1\}^3$, since each $\alpha \in \{0,1\}^3$ must have either two 0s or two 1s. And thus, $\mathsf{wt}(\{00*, 11*\}) + \mathsf{wt}(\{0*0, 1*1\}) + \mathsf{wt}(\{*00, *11\}) \geqslant 1$, thus at least one of the three terms exceeds $1/3$. □

9.1.3 Clustering in 3-colorable one-sided spectral expanders

We now discuss an analogous extremal clustering property of 3-colorings in one-sided spectral expanders. This property is stated in terms of pairwise "agreement" between different 3-colorings — a natural generalization of intersection that "mods" out the symmetry between colors.

Definition 9.1.5. The relative *agreement* between two valid 3-colorings x and y according to a permutation π is defined by:

$$\mathrm{agree}_\pi(x,y) := \mathbb{E}_{u\in V}[\pi(x_u) = y_u]\,,$$

and the agreement between x and y is defined as the maximum over all permutations:

$$\mathrm{agree}(x,y) := \max_{\pi\in S_3} \mathrm{agree}_\pi(x,y)\,.$$

The agreement between two relabelings of the same coloring is 1 (the maximum possible).

We will prove the following extremal clustering property of 3-colorings in a spectral expander that informally says that in any collection of three non-trivial 3-colorings, two must have a better-than-random agreement.

Lemma 9.1.6. *Let $G = (V, E)$ be a regular 3-colorable graph with $\lambda_2(G) \leqslant \frac{\varepsilon}{1+\varepsilon}$ for some small enough ε. Then, given any 3 valid 3-colorings of G such that no color class has size $> (\frac{1}{2} + \varepsilon)n$, there exist two with an agreement $\geqslant \frac{1}{2} + \varepsilon$.*

Agreement between 2 valid 3-colorings. We now analyze the agreement between 2 valid 3-colorings $x, y \in [3]^n$ of G. Similar to Section 9.1.2, the colorings induce a partition of the vertices into 9 subsets indexed by $\{1,2,3\}^2$, where set ij contains vertices that are assigned i and j by x and y respectively (see Figure 9.2). The 9-vertex graph in Figure 9.2 is exactly $H = K_3 \otimes K_3$ where K_3 is a triangle.

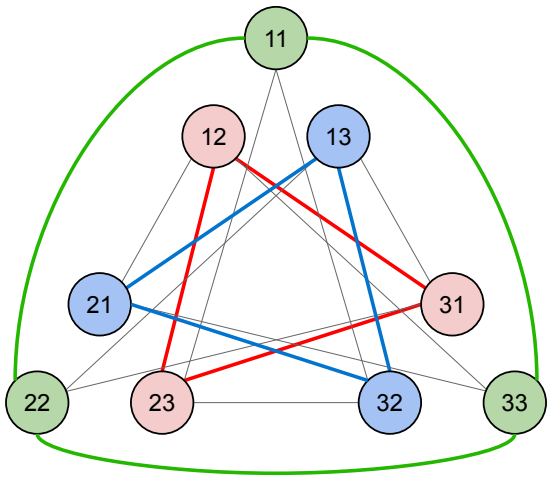


Figure 9.2: The triangle gadget for 2 valid 3-colorings. There are 2 ways to partition the 9 vertices into 3 disjoint triangles. The highlighted triangles show the partition $\{S_\pi\}_{\pi \in \mathbb{S}_3^+}$.

Define the set

$$S_\pi := \{(\sigma, \pi(\sigma)) : \sigma \in \{1,2,3\}\} .$$

Then, for any $\pi \in \mathbb{S}_3$, $\text{agree}_\pi(x, y) = \text{wt}(S_\pi)$.

Claim 9.1.7. *If $\lambda_2 \leqslant 1 - \frac{1}{1+\varepsilon}$, then,*

$$\sum_{\pi \in \mathbb{S}_3} \text{wt}(S_\pi)^2 \geqslant 2 - \frac{1}{1-\lambda_2} \geqslant 1 - \varepsilon . \tag{9.2}$$

Proof. Observe that for any π, S_π forms a triangle in H. In fact, there are exactly two ways to partition the 9 vertex graph above into 3 disjoint triangles: (1) $\{11, 22, 33\}$, $\{12, 23, 31\}$, $\{13, 21, 32\}$ (highlighted in Figure 9.2), and (2) $\{11, 23, 32\}$, $\{12, 21, 33\}$, $\{13, 22, 31\}$, where each of the 6 triangles appearing in the list above corresponds to a permutation $\pi \in \mathbb{S}_3$.

Now, $e(S_\pi, \overline{S}_\pi) \geqslant (1 - \lambda_2) \cdot \text{wt}(S_\pi)(1 - \text{wt}(S_\pi))$ for each π. Summing up the inequalities over $\pi \in \mathbb{S}_3$ gives $(1 - \lambda_2) \sum_\pi \text{wt}(S_\pi)(1 - \text{wt}(S_\pi)) = (1 - \lambda_2)(2 - \sum_\pi \text{wt}(S_\pi)^2)$ on

the right-hand side and $\sum_\pi e(S_\pi, \overline{S}_\pi) = 1$ on the left-hand side. Thus, rearranging gives us

$$\sum_{\pi \in \mathbb{S}_3} \mathsf{wt}(S_\pi)^2 \geqslant 2 - \frac{1}{1-\lambda_2} \geqslant 1 - \varepsilon\,. \qquad \square$$

Small agreement + expansion implies almost bipartite. We show the following claim:

Claim 9.1.8. *Suppose $\lambda_2 \leqslant \frac{\varepsilon}{1+\varepsilon}$ and $\mathrm{agree}(x,y) \leqslant \frac{1}{2} + \varepsilon$ for small enough ε, then one of $\{w(S_\pi)\}_{\pi \in \mathbb{S}_3^+}$ and one of $\{w(S_\pi)\}_{\pi \in \mathbb{S}_3^-}$ is at most $O(\varepsilon)$.*

As a result, G is almost bipartite, i.e., removing an $O(\varepsilon)$ fraction of vertices results in a bipartite graph.

Recall that $\mathrm{agree}(x,y) \leqslant \frac{1}{2} + \varepsilon$ means that $\mathsf{wt}(S_\pi) \leqslant \frac{1}{2} + \varepsilon$ for all $\pi \in \mathbb{S}_3$. To prove Claim 9.1.8, we formulate it as a 6-variable lemma (see Lemma 9.3.10): let $z_1, z_2, \ldots, z_6$ be such that $0 \leqslant z_i \leqslant \frac{1}{2} + \varepsilon$ for each i, $z_1 + z_2 + z_3 = z_4 + z_5 + z_6 = 1$, and $\|z\|_2^2 \geqslant 1 - \varepsilon$, then one of z_1, z_2, z_3 and one of z_4, z_5, z_6 must be $\leqslant O(\varepsilon)$.

With this lemma, the first statement in Claim 9.1.8 immediately follows from $\mathsf{wt}(S_\pi) \leqslant \frac{1}{2} + \varepsilon$, $\sum_{\pi \in \mathbb{S}_3^+} \mathsf{wt}(S_\pi) = \sum_{\pi \in \mathbb{S}_3^-} \mathsf{wt}(S_\pi) = 1$, and Eq. (9.2).

For the second statement, let $\pi^+ \in \mathbb{S}_3^+$ and $\pi^- \in \mathbb{S}_3^-$ be the permutations such that $\mathsf{wt}(S_{\pi^+})$, $\mathsf{wt}(S_{\pi^-}) \leqslant O(\varepsilon)$. Note that since π^+ and π^- have different signs, S_{π^+} and S_{π^-} intersect in exactly one string $\alpha \in [3]^2$. In fact, α uniquely determines π^+, π^- since there are exactly two permutations with different signs that map α_1 to α_2. Assume without loss of generality (due to symmetry) that $\alpha = 11$, so that $S_{\pi^+} = \{11, 22, 33\}$ and $S_{\pi^-} = \{11, 23, 32\}$. Then, we have $\mathsf{wt}(\{11, 22, 33\})$, $\mathsf{wt}(\{11, 23, 32\}) \leqslant O(\varepsilon)$. This means that $\mathsf{wt}(\{12, 13, 21, 23\}) \geqslant 1 - O(\varepsilon)$. Observe that $\{12, 13, 21, 23\}$ forms a bipartite structure between $\{12, 13\}$ and $\{21, 23\}$, as shown in Figure 9.3. In particular, the first coloring labels the entire left side with the same color, while the second labels the right side with the same color.

Agreement between 3 valid 3-colorings. Naturally, we consider the graph as being partitioned into 27 subsets indexed by strings $[3]^3$. Again, we will use "$*$" to denote "free" coordinate, so for example $11*$ means $\{111, 112, 113\}$, i.e., the set 11 if we ignore the third coloring.

Suppose for contradiction that the agreement between each pair of 3-colorings is at most $\frac{1}{2} + \varepsilon$. Then, by Claim 9.1.8, we have that each pair (i, j) of colorings gives a bipartite structure, denoted $T^{(ij)}$, such that $\mathsf{wt}(T^{(ij)}) \geqslant 1 - O(\varepsilon)$. This is best explained by example. Suppose $T^{(12)} = \{12*, 13*, 21*, 23*\}$, $T^{(13)} = \{1*2, 1*3, 2*1, 3*1\}$ and $T^{(23)} = \{*11, *13, *22, *32\}$. Then, we can see that $T := T^{(12)} \cap T^{(13)} \cap T^{(23)} = \{122, 132, 211, 311\}$. This is a bipartite structure between $\{122, 132\}$ and $\{211, 311\}$, where the first coloring labels the entire left side with the same color, while the second and third label the right side with the same color.

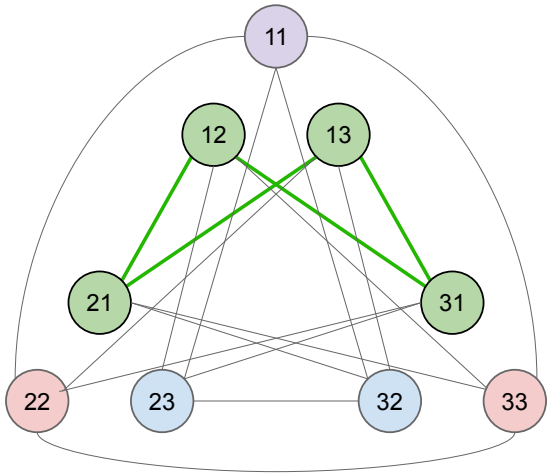


Figure 9.3: $S_{\pi^+} = \{11, 22, 33\}$ and $S_{\pi^-} = \{11, 23, 32\}$, and $\mathsf{wt}(S_{\pi^+}), \mathsf{wt}(S_{\pi^-}) \leqslant O(\varepsilon)$, which means that $\mathsf{wt}(\{12, 13, 21, 23\}) \geqslant 1 - O(\varepsilon)$. Here $\{12, 13, 21, 23\}$ forms a bipartite structure.

Moreover, we have $\mathsf{wt}(T) \geqslant 1 - O(\varepsilon)$. We now use this to derive a contradiction. Suppose no colors have size larger than $(\frac{1}{2} + \varepsilon)n$, so $\mathsf{wt}(\{122, 132\}), \mathsf{wt}(\{211, 311\}) \leqslant \frac{1}{2} + \varepsilon$. This implies that $\mathsf{wt}(\{122, 132\}), \mathsf{wt}(\{211, 311\}) \geqslant \frac{1}{2} - O(\varepsilon)$. Next, observe that $\{122, 211, 311\} \subseteq \{*11, *22\} \subseteq S_\pi$ between the second and third colorings for some π. Similarly, $\{132, 211, 311\} \subseteq \{*11, *32\} \subseteq S_{\pi'}$ for some π'. Thus, one of them has weight at least $\mathsf{wt}(\{211, 311\}) + \frac{1}{2}\mathsf{wt}(\{122, 132\}) \geqslant \frac{3}{4} - O(\varepsilon)$, contradicting that each pairwise agreement is $\leqslant \frac{1}{2} + \varepsilon$.

One can verify that the above holds in general; T will contain at most 4 strings in $[3]^3$ and form the bipartite structure explained above. This proves Lemma 9.1.6.

9.2 Independent sets on spectral expanders

In this section, we prove Theorem 6.3.3. Our algorithm starts by considering a constant degree SoS relaxation of the integer program for Independent Set (9.1) and obtaining a pseudo-distribution μ'. We then apply a simple rounding algorithm to obtain an independent set in G as shown below.

Algorithm 9.2.1 (Find independent set in an expander)**.**
Input: A graph $G = (V, E)$.
Output: An independent set of G.
Operation:

1. Solve a degree $D = O(1)$ SoS relaxation of the integer program (9.1) to

obtain a pseudo-distribution μ'.

2. Choose a uniformly random set of $t = O(1)$ vertices $i_1, \dots, i_t \sim [n]$ and draw $(\sigma_{i_1}, \dots, \sigma_{i_t}) \sim \mu'$. Let μ be the pseudo-distribution obtained by conditioning μ' on $(x_{i_1} = \sigma_{i_1}, \dots, x_{i_t} = \sigma_{i_t})$.
3. Output the set $\{u \in V : \widetilde{\mathbb{E}}_{\mu}[x_u] > \frac{1}{2}\}$.

9.2.1 Multiple assignments from μ: definitions and facts

Fix $t \in \mathbb{N}$. Throughout this section, we will work with t assignments $x^{(1)}, x^{(2)}, \dots, x^{(t)}$ that the reader should think of as independent samples from the pseudo-distribution μ, i.e. each $x^{(i)}$ is an n-dimensional vector which is the indicator of a $(1/2 - \varepsilon)n$-sized independent set in G and therefore it satisfies the constraints of the integer program (9.1). Given $x^{(1)}, x^{(2)}, \dots, x^{(t)}$ we use boldface $\boldsymbol{x}$ to denote $(x^{(1)}, \dots, x^{(t)})$, i.e., the collection of variables $x_u^{(i)}$ for $u \in [n]$ and $i \in [t]$. Moreover, for $U \subseteq [t]$, we write $\boldsymbol{x}^U := (x^{(i)})_{i \in U}$.

Definition 9.2.2. We denote $\mathcal{A}_G^{\mathsf{bool}}(x) := \{x_u^2 - x_u = 0, \ \forall u \in V\}$, i.e., the Booleanity constraints. Moreover, we write $\mathcal{A}_G^{\mathsf{IS}}(x)$ to denote the *independent set constraints*:

$$\mathcal{A}_G^{\mathsf{IS}}(x) := \mathcal{A}_G^{\mathsf{bool}}(x) \cup \{x_u x_v = 0, \ \forall \{u, v\} \in E\}.$$

Moreover, with slight abuse of notation, for $t \in \mathbb{N}$ and vectors $x^{(1)}, x^{(2)}, \dots, x^{(t)}$,

$$\mathcal{A}_G^{\mathsf{bool}}(\boldsymbol{x}) := \bigcup_{i \in [t]} \mathcal{A}_G^{\mathsf{bool}}(x^{(i)}), \quad \mathcal{A}_G^{\mathsf{IS}}(\boldsymbol{x}) := \bigcup_{i \in [t]} \mathcal{A}_G^{\mathsf{IS}}(x^{(i)}).$$

We will drop the dependence on G when the graph is clear from context.

Given assignments $x^{(1)}, \dots, x^{(t)} \in \{0,1\}^n$ and $\alpha \in \{0,1\}^t$, for each vertex $u \in [n]$, we define below the event that u is assigned α_i by $x^{(i)}$, which is viewed as a degree-t multilinear polynomial of $\boldsymbol{x}$. Similarly, for $S \subseteq \{0,1\}^t$, we define the event that u receives one of the assignments in S.

Definition 9.2.3. Let $t \in \mathbb{N}$, and let $\boldsymbol{x} = (x^{(1)}, x^{(2)}, \dots, x^{(t)})$. For each $u \in [n]$, $\alpha \in \{0,1\}^t$ and $S \subseteq \{0,1\}^t$, we define the following events,

$$\mathbf{1}(u \leftarrow \alpha) := \mathbf{1}(x_u^{(1)} = \alpha_1, \dots, x_u^{(t)} = \alpha_t) = \prod_{i \in [t]} \left(x_u^{(i)}\right)^{\alpha_i} \left(1 - x_u^{(i)}\right)^{1 - \alpha_i},$$

$$\mathbf{1}(u \leftarrow S) := \sum_{\alpha \in S} \mathbf{1}(u \leftarrow \alpha).$$

For convenience, we omit the dependence on $\boldsymbol{x}$. We will also consider the quantity $\mathsf{wt}(\alpha)$ which is the fraction of vertices that get assigned α:

$$\mathsf{wt}(\alpha) := \mathbb{E}_{u \in [n]}[\mathbf{1}(u \leftarrow \alpha)].$$

Similarly, $\mathsf{wt}(S) := \mathbb{E}_{u\in[n]}[\mathbf{1}(u \leftarrow S)]$ for $S \subseteq \{0,1\}^t$.

Moreover, we will use the symbol "$*$" to denote "free variables" — for $\beta \in \{0,1,*\}^t$, $\mathbf{1}(u \leftarrow \beta) := \mathbf{1}(u \leftarrow S_\beta)$ and $\mathsf{wt}(\beta) := \mathsf{wt}(S_\beta)$ where $S_\beta = \{\alpha \in \{0,1\}^t : \alpha_i = \beta_i \text{ if } \beta_i \neq *\}$. For example, $\mathsf{wt}(00*) = \mathsf{wt}(000) + \mathsf{wt}(001)$.

We note some simple facts (written in SoS form) that will be useful later.

Fact 9.2.4. *The following can be easily verified: for $\alpha, \beta \in \{0,1\}^t$,*

(1) $\mathcal{A}^{\mathsf{bool}}(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{2t}\right. \{\mathbf{1}(u \leftarrow \alpha)^2 = \mathbf{1}(u \leftarrow \alpha)\}$*, i.e.,* $\mathbf{1}(u \leftarrow \alpha)$ *satisfies the Booleanity constraint.*

(2) $\mathcal{A}^{\mathsf{bool}}(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{2t}\right. \{\mathbf{1}(u \leftarrow \alpha)\cdot\mathbf{1}(u \leftarrow \beta) = 0\}$ *for $\alpha \neq \beta$. This also implies that $\mathbf{1}(u \leftarrow S)$ satisfies the Booleanity constraint for any $S \subseteq \{0,1\}^t$.*

(3) $\left|\frac{\boldsymbol{x}}{t}\right. \left\{\sum_{\alpha\in\{0,1\}^t} \mathbf{1}(u \leftarrow \alpha) = 1\right\}$.

We next prove the following lemma, which is an "SoS proof" that if $x^{(1)}, \ldots, x^{(t)}$ are indicators of independent sets and $\{u,v\} \in E$, then u and v cannot be both assigned 1 by any $x^{(i)}$. As a consequence, any vertex that is assigned all 1s can only be connected to vertices that are assigned all 0s, meaning that if v gets $\vec{1}$ then u must get $\vec{0}$.

Lemma 9.2.5. *Let $t \in \mathbb{N}$ and $\boldsymbol{x} = (x^{(1)}, x^{(2)}, \ldots, x^{(t)})$ be variables. For any graph $G = (V,E)$ and any $\alpha, \beta \in \{0,1\}^t$ such that $\mathrm{supp}(\alpha) \cap \mathrm{supp}(\beta) \neq \varnothing$, then for all $\{u,v\} \in E$ we have*

$$\mathcal{A}^{\mathsf{IS}}_G(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{2t}\right. \{\mathbf{1}(u \leftarrow \alpha)\mathbf{1}(v \leftarrow \beta) = 0\}\ .$$

In particular, for all $\{u,v\} \in E$,

$$\mathcal{A}^{\mathsf{IS}}_G(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{2t}\right. \left\{\left(1 - \mathbf{1}(u \leftarrow \vec{0})\right)\cdot\mathbf{1}(v \leftarrow \vec{1}) = 0\right\}$$
$$\left|\frac{\boldsymbol{x}}{2t}\right. \left\{\mathbf{1}(u \leftarrow \vec{0}) \geqslant \mathbf{1}(v \leftarrow \vec{1})\right\}\ .$$

Proof. Let $i \in [t]$ be the index such that $\alpha_i = \beta_i = 1$. Then, by Definition 9.2.3, $\mathbf{1}(u \leftarrow \alpha)\cdot\mathbf{1}(v \leftarrow \beta) = x^{(i)}_u x^{(i)}_v \cdot f(\boldsymbol{x})$ for some polynomial f (not depending on $x^{(i)}$). The first statement follows since $x^{(i)}_u x^{(i)}_v = 0$ is in the independent set constraints.

For the second statement, $(1 - \mathbf{1}(u \leftarrow \vec{0}))\mathbf{1}(v \leftarrow \vec{1}) = 0$ follows from the polynomial equality $\sum_{\alpha\in\{0,1\}^t} \mathbf{1}(u \leftarrow \alpha) = 1$ and that $\vec{1}$ intersects with all $\alpha \neq \vec{0}$. Moreover, $\mathbf{1}(u \leftarrow \alpha)$ satisfies the Booleanity constraints (Fact 9.2.4). Denoting $a := \mathbf{1}(u \leftarrow \vec{0})$ and $b := \mathbf{1}(u \leftarrow \vec{1})$ for convenience, from $(1-a)b = 0$ and $a^2 = a$, $b^2 = b$ we have

$$a - b = (a-b)^2 - 2(1-a)b + (a - a^2) + (b - b^2) \geqslant 0\,,$$

which completes the proof. □

9.2.2 Spectral gap implies a unique solution

Recall the definitions from Definition 9.2.3. We need some definitions for edge sets in the graph.

Definition 9.2.6. Let G be a graph, let $t \in \mathbb{N}$, and let $\boldsymbol{x} = (x^{(1)}, \dots, x^{(t)})$. For $\alpha, \beta \in \{0,1\}^t$, define

$$e(\alpha, \beta) := \frac{1}{2|E(G)|} \sum_{\{u,v\} \in E(G)} \mathbf{1}(u \leftarrow \alpha)\mathbf{1}(v \leftarrow \beta) + \mathbf{1}(u \leftarrow \beta)\mathbf{1}(v \leftarrow \alpha)\,.$$

Here, we omit the dependence on $\boldsymbol{x}$ and G for simplicity.

Similarly, for $S, T \subseteq \{0,1\}^t$, we denote $e(S,T) := \sum_{\alpha \in S} \sum_{\beta \in T} e(\alpha, \beta)$.

Given assignments $x^{(1)}, \dots, x^{(t)} \in \{0,1\}^n$ and $\alpha, \beta \in \{0,1\}^t$, one should view $e(\alpha, \beta)$ as the (normalized) number of edges between vertices that are assigned α and vertices assigned β. We note a few properties which can be easily verified:

Fact 9.2.7. *The following can be easily verified:*

(1) Symmetry: $e(\alpha, \beta) = e(\beta, \alpha)$ by definition.

(2) Sum of edge weights (double counted) equals 1: $\left|\frac{\boldsymbol{x}}{2t}\right. \{\sum_{\alpha,\beta \in \{0,1\}^t} e(\alpha, \beta) = 1\}$, which follows from $\left|\frac{\boldsymbol{x}}{t}\right. \{\sum_{\alpha \in \{0,1\}^t} \mathbf{1}(u \leftarrow \alpha) = 1\}$.

(3) For a regular graph, the weight of a subset equals the weight of incident edges: for any $\alpha \in \{0,1\}^t$, $\left|\frac{\boldsymbol{x}}{2t}\right. \{\sum_{\beta \in \{0,1\}^t} e(\alpha, \beta) = \mathsf{wt}(\alpha)\}$.

(4) $\mathcal{A}_G^{\mathsf{IS}}(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{2t}\right. \{e(\alpha, \beta) = 0\}$ for any α, β such that $\mathrm{supp}(\alpha) \cap \mathrm{supp}(\beta) \neq \varnothing$ due to Lemma 9.2.5.

We next show the following lemma relating the Laplacian to the cut in the graph.

Lemma 9.2.8. *Let G be a graph and L_G be its Laplacian matrix. Let $t \in \mathbb{N}$, $S \subseteq \{0,1\}^t$, and let $y_u := \mathbf{1}(u \leftarrow S)$ for each vertex u, we have*

$$\mathcal{A}^{\mathsf{bool}}(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{2t}\right. \left\{ \frac{1}{2|E(G)|} \cdot y^\top L_G y = e(S, \overline{S}) \right\}.$$

Proof. Since y_u satisfies the Booleanity constraint and $1 - \mathbf{1}(u \leftarrow S) = \mathbf{1}(u \leftarrow \overline{S})$, for any u, v,

$$\begin{aligned} \mathcal{A}^{\mathsf{bool}}(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{2t}\right. \Big\{ (y_u - y_v)^2 &= \mathbf{1}(u \leftarrow S) + \mathbf{1}(v \leftarrow S) - 2 \cdot \mathbf{1}(u \leftarrow S)\mathbf{1}(v \leftarrow S) \\ &= \mathbf{1}(u \leftarrow S)\mathbf{1}(v \leftarrow \overline{S}) + \mathbf{1}(v \leftarrow S)\mathbf{1}(u \leftarrow \overline{S}) \Big\}. \end{aligned}$$

The lemma then follows by noting that $y^\top L_G y = \sum_{\{u,v\} \in E(G)} (y_u - y_v)^2$. □

For rounding independent sets on spectral expanders, we will only consider $t = 2$ and 3. For $t = 2$, we get a simple bound that $\mathsf{wt}(00) - \mathsf{wt}(11) \leqslant 2\varepsilon$ given that the graph has an independent set of size $(\frac{1}{2} - \varepsilon)n$, i.e., $\mathbb{E}_u[x_u^{(1)}]$ and $\mathbb{E}_u[x_u^{(2)}] \geqslant \frac{1}{2} - \varepsilon$.

Lemma 9.2.9. *Let $\boldsymbol{x} = (x^{(1)}, x^{(2)})$.*

$$\left|\frac{\boldsymbol{x}}{2}\right. \left\{ \mathsf{wt}(00) - \mathsf{wt}(11) = 2\varepsilon - \sum_{t\in[2]} \left(\mathbb{E}_u[x_u^{(t)}] - \left(\frac{1}{2} - \varepsilon\right)\right)\right\}.$$

Proof. First note that $\mathbb{E}_u[x_u^{(1)}] = \mathsf{wt}(10) + \mathsf{wt}(11)$ and $\mathbb{E}_u[x_u^{(2)}] = \mathsf{wt}(01) + \mathsf{wt}(11)$. Summing up $\mathbb{E}_u[x_u^{(1)}] - (\frac{1}{2} - \varepsilon)$ and $\mathbb{E}_u[x_u^{(2)}] - (\frac{1}{2} - \varepsilon)$ gives $\mathsf{wt}(01) + \mathsf{wt}(10) + 2\mathsf{wt}(11) - (1 - 2\varepsilon)$. Then, noting that $\mathsf{wt}(00) + \mathsf{wt}(01) + \mathsf{wt}(10) + \mathsf{wt}(11) = 1$ completes the proof. □

We next lower bound $\mathsf{wt}(00) - \mathsf{wt}(11)$ by the expansion of the graph.

Lemma 9.2.10. *Let G be a d-regular n-vertex graph with $\lambda_2 := \lambda_2(G) > 0$. Let $\boldsymbol{x} = (x^{(1)}, x^{(2)})$. Then,*

$$\mathcal{A}^{\mathsf{IS}}_G(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{4}\right. \{\mathsf{wt}(00) - \mathsf{wt}(11) \geqslant (1 - \lambda_2) \cdot \mathsf{wt}(\{00, 11\})(1 - \mathsf{wt}(\{00, 11\}))\}.$$

Proof. Let $S = \{01, 10\}$, and define $y_u := \mathbf{1}(u \leftarrow S)$. By Lemma 9.2.8,

$$\mathcal{A}^{\mathsf{IS}}_G(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{4}\right. \left\{ \frac{1}{nd} \cdot y^\top L_G y = e(S, \overline{S}) = e(00, 01) + e(00, 10) \leqslant \mathsf{wt}(00) - \mathsf{wt}(11)\right\} \tag{9.3}$$

where $e(S, \overline{S}) = e(00, 01) + e(00, 10)$ because $\mathcal{A}^{\mathsf{IS}}(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{2t}\right. \{e(01, 11) = e(10, 11) = 0\}$ (Fact 9.2.7), and the last inequality follows from $\mathsf{wt}(00) = \sum_{\alpha\in\{0,1\}^2} e(00, \alpha)$ and $\mathsf{wt}(11) = e(00, 11)$ (again because $e(01, 11) = e(10, 11) = 0$).

On the other hand, the trivial eigenvector of L_G is $\vec{1}$ with eigenvalue 0 while $\lambda_2(\frac{1}{d}L_G) = 1 - \lambda_2$, so we have

$$\left|\frac{y}{2}\right. \left\{ \frac{1}{nd} y^\top L_G y \geqslant \frac{1}{n} \cdot (1 - \lambda_2) \left(\|y\|_2^2 - \frac{1}{n}\langle \vec{1}, y\rangle^2\right)\right\}.$$

By the Booleanity constraints, $\mathcal{A}^{\mathsf{bool}}(y) \left|\frac{y}{2}\right. \frac{1}{n}(\|y\|_2^2 - \frac{1}{n}\langle \vec{1}, y\rangle^2) = \mathbb{E}_u[y_u] - \mathbb{E}_u[y_u]^2 = \mathsf{wt}(S)(1 - \mathsf{wt}(S)) = \mathsf{wt}(\overline{S})(1 - \mathsf{wt}(\overline{S}))$, where $\overline{S} = \{00, 11\}$. This combined with Eq. (9.3) finishes the proof. □

Combining Lemmas 9.2.9 and 9.2.10, we have that $\mathbb{E}_u[x_u^{(t)}] \geqslant \frac{1}{2} - \varepsilon$ (i.e., the independent set indicated by $x^{(t)}$ has size at least $(\frac{1}{2} - \varepsilon)n$) together with the expansion of the graph imply that

$$(1 - \lambda_2) \cdot \mathsf{wt}(\{00, 11\})(1 - \mathsf{wt}(\{00, 11\})) \leqslant \mathsf{wt}(00) - \mathsf{wt}(11) \leqslant 2\varepsilon.$$

When $\lambda_2 \leqslant 1 - C\varepsilon$ for some large enough constant C, then the above implies either $\mathsf{wt}(\{00,11\}) \leqslant \gamma$ or $\mathsf{wt}(\{00,11\}) \geqslant 1-\gamma$ for some small constant $\gamma < \frac{1}{3}$. In the latter case, since $\mathsf{wt}(11) \geqslant \mathsf{wt}(00) - 2\varepsilon$, we have $\mathsf{wt}(11) \geqslant \frac{1}{2} - \frac{\gamma}{2} - \varepsilon$.

Now, we now consider 3 assignments, where each pair of assignments satisfy the above, i.e., $\mathsf{wt}(\{00*,11*\})(1-\mathsf{wt}(\{00*,11*\})) \leqslant \frac{2\varepsilon}{1-\lambda_2} \leqslant \frac{2}{C}$ for all 3 "$*$" locations. Then, we claim that one of them, say $\mathsf{wt}(\{00*,11*\})$, must be $\geqslant 1-\gamma$. To see this, notice that the 3 pairs $\mathsf{wt}(\{00*,11*\})$ must sum up to at least 1 because each $\alpha \in \{0,1\}^3$ is covered, i.e., has either two 0s or two 1s. Thus, all 3 being $\leqslant \gamma$ leads to a contradiction.

We now formalize this reasoning as an SoS proof.

Lemma 9.2.11. *Let G be a d-regular n-vertex graph with $\lambda_2 := \lambda_2(G) > 0$, and let $\varepsilon > 0$. Let $\boldsymbol{x} = (x^{(1)}, x^{(2)}, x^{(3)})$. Let $\mathcal{A}$ be the constraints $\mathcal{A}^{\mathsf{IS}}_G(\boldsymbol{x}) \cup \left\{\mathbb{E}_u[x_u^{(t)}] \geqslant \frac{1}{2} - \varepsilon,\ \forall t \in [3]\right\}$. Then,*

$$\mathcal{A} \left|\frac{\boldsymbol{x}}{6}\right. \left\{ (\mathsf{wt}(11*)+\varepsilon)^2 + (\mathsf{wt}(1*1)+\varepsilon)^2 + (\mathsf{wt}(*11)+\varepsilon)^2 \geqslant \frac{1}{4}\left(1 - \frac{6\varepsilon}{1-\lambda_2}\right)\right\}.$$

Proof. By Lemma 9.2.9, we have $\mathcal{A}$ implies that $\mathsf{wt}(00*) \leqslant \mathsf{wt}(11*) + 2\varepsilon$. Moreover, by Lemma 9.2.10, we have

$$\begin{aligned} 2\varepsilon \geqslant \mathsf{wt}(00*) - \mathsf{wt}(11*) &\geqslant (1-\lambda_2)\cdot\Big((\mathsf{wt}(00*)+\mathsf{wt}(11*)) - (\mathsf{wt}(00*)+\mathsf{wt}(11*))^2\Big) \\ &\geqslant (1-\lambda_2)\cdot\Big((\mathsf{wt}(00*)+\mathsf{wt}(11*)) - 4(\mathsf{wt}(11*)+\varepsilon)^2\Big). \end{aligned}$$

Next, we sum up the inequalities for all 3 "$*$" locations. Observe that $\{00*,11*\} \cup \{0*0,1*1\} \cup \{*00,*11\} = \{0,1\}^3$, as any $\alpha \in \{0,1\}^3$ must have either 2 zeros or 2 ones. This means that the sum of $\mathsf{wt}(00*) + \mathsf{wt}(11*)$ must be $\geqslant 1$. Thus,

$$\mathcal{A} \left|\frac{\boldsymbol{x}}{6}\right. \left\{(1-\lambda_2)\cdot\Big(1 - 4\Big((\mathsf{wt}(11*)+\varepsilon)^2 + (\mathsf{wt}(1*1)+\varepsilon)^2 + (\mathsf{wt}(*11)+\varepsilon)^2\Big)\Big) \leqslant 6\varepsilon\right\},$$

and rearranging the above completes the proof. □

9.2.3 Analysis of Algorithm 9.2.1

We now prove that Algorithm 9.2.1 successfully outputs an independent set of size $\Omega(n)$.

Lemma 9.2.12. *Let $\eta, \delta \in (0,1)$ such that $\delta \leqslant \eta^2/18$, and let μ be a pseudo-distribution over $\{0,1\}^n$ such that $\mathbb{E}_{u,v\in[n]} I_\mu(X_u; X_v) \leqslant \delta$. Suppose $\widetilde{\mathbb{E}}_{\mu^{\otimes 2}}[\mathsf{wt}(11)^2] \geqslant \frac{1}{16} + \eta$, then the set of vertices u such that $\widetilde{\Pr}_\mu[x_u = 1] > \frac{1}{2}$ forms an independent set of size $\eta n/4$.*

Proof. Recall from Definition 9.2.3 that $\mathsf{wt}(11) = \mathbb{E}_{u\sim[n]}[x_u^{(1)} x_u^{(2)}]$, thus

$$\begin{aligned} \widetilde{\mathbb{E}}_{\mu^{\otimes 2}}[\mathsf{wt}(11)^2] &= \widetilde{\mathbb{E}}_{\mu^{\otimes 2}} \mathbb{E}_{u,v\sim[n]}\left[x_u^{(1)} x_u^{(2)} x_v^{(1)} x_v^{(2)}\right] \\ &= \mathbb{E}_{u,v\sim[n]}\left[\widetilde{\mathbb{E}}_\mu[x_u x_v]^2\right] = \mathbb{E}_{u,v\sim[n]}\left[\widetilde{\Pr}_\mu[x_u = 1,\ x_v = 1]^2\right]. \end{aligned}$$

Now, given that μ has small average correlation, by Pinsker's inequality (Fact 2.5.8),

$$\mathbb{E}_{u,v\sim[n]}\left|\widetilde{\Pr}_{\mu}[x_u=1,\ x_v=1]-\widetilde{\Pr}_{\mu}[x_u=1]\widetilde{\Pr}_{\mu}[x_v=1]\right|\leqslant\sqrt{\delta/2}\,.$$

Then, using the fact that $p^2=q^2+2q(p-q)+(p-q)^2\leqslant q^2+3|p-q|$ for all $p,q\in[0,1]$, we have

$$\begin{aligned}\mathbb{E}_{u,v\sim[n]}\left[\widetilde{\Pr}_{\mu}[x_u=1,\ x_v=1]^2\right]&\leqslant\mathbb{E}_{u,v\sim[n]}\left[\widetilde{\Pr}_{\mu}[x_u=1]^2\widetilde{\Pr}_{\mu}[x_v=1]^2\right]+3\sqrt{\delta/2}\\&\leqslant\mathbb{E}_{u\sim[n]}\left[\widetilde{\Pr}_{\mu}[x_u=1]^2\right]^2+3\sqrt{\delta/2}\,.\end{aligned}$$

Thus, since $\widetilde{\mathbb{E}}_{\mu^{\otimes 2}}[\mathsf{wt}(11)^2]\geqslant\frac{1}{16}+\eta$ and $\delta\leqslant\eta^2/18$, we have $\mathbb{E}_{u\sim[n]}\left[\widetilde{\Pr}_{\mu}[x_u=1]^2\right]^2\geqslant\frac{1}{16}+\frac{\eta}{2}$, which means that $\mathbb{E}_{u\sim[n]}\left[\widetilde{\Pr}_{\mu}[x_u=1]^2\right]\geqslant\sqrt{\frac{1}{16}+\frac{\eta}{2}}\geqslant\frac{1}{4}+\frac{\eta}{2}$. It follows that at least $\eta/4$ fraction of vertices have $\widetilde{\Pr}_{\mu}[x_u=1]>\frac{1}{2}$. By Fact 9.1.1, these vertices form an independent set. □

Proof of Theorem 6.3.3. By the assumption that G contains an independent set of size $(\frac{1}{2}-\varepsilon)n$, the pseudo-distribution μ satisfies the constraint $\mathbb{E}_u[x_u]\geqslant\frac{1}{2}-\varepsilon$. Let $\boldsymbol{x}=(x^{(1)},x^{(2)},x^{(3)})\sim\mu^{\otimes 3}$, then Lemma 9.2.11 states that

$$\widetilde{\mathbb{E}}_{\mu^{\otimes 3}}\left[(\mathsf{wt}(11*)+\varepsilon)^2+(\mathsf{wt}(1*1)+\varepsilon)^2+(\mathsf{wt}(*11)+\varepsilon)^2\right]\geqslant\frac{1}{4}\left(1-\frac{6\varepsilon}{1-\lambda_2}\right).$$

By symmetry, the 3 terms on the left-hand side are equal, and

$$\widetilde{\mathbb{E}}_{\mu^{\otimes 3}}[(\mathsf{wt}(11*)+\varepsilon)^2]=\widetilde{\mathbb{E}}_{\mu^{\otimes 2}}[\mathsf{wt}(11)^2+2\varepsilon\cdot\mathsf{wt}(11)+\varepsilon^2]\leqslant\widetilde{\mathbb{E}}_{\mu^{\otimes 2}}[\mathsf{wt}(11)^2]+2\varepsilon+\varepsilon^2\,.$$

Thus, if $\varepsilon\leqslant 0.001$ and $\lambda_2\leqslant 1-C\varepsilon$ with $C=40$, then we have $\widetilde{\mathbb{E}}_{\mu^{\otimes 2}}[\mathsf{wt}(11)^2]\geqslant\frac{1}{12}(1-\frac{6}{C})-(2\varepsilon+\varepsilon^2)\geqslant\frac{1}{15}>\frac{1}{16}$.

By Lemma 2.5.7, after we condition μ' on the values of $O(1/\delta)$ variables as done in Step (2) of Algorithm 9.2.1 to get μ, we have $\mathbb{E}_{u,v\in[n]}[I_{\mu}(X_u;X_v)]\leqslant\delta$, where δ is a small enough constant. Then, by Lemma 9.2.12, at least $\frac{1}{4}(\frac{1}{15}-\frac{1}{16})\geqslant\frac{1}{1000}$ fraction of the vertices have $\widetilde{\Pr}_{\mu}[x_u=1]>\frac{1}{2}$. By Fact 9.1.1, this must be an independent set, thus completing the proof. □

9.3 Independent sets on almost 3-colorable spectral expanders

Recall that an ε-almost 3-colorable graph is a graph which is 3-colorable if one removes ε fraction of the vertices.

Theorem 9.3.1 (Formal version of Theorem 6.3.2)**.** *For any $\varepsilon \in [0, 10^{-4}]$, let G be an n-vertex regular ε-almost 3-colorable graph with $\lambda_2(G) \leqslant 10^{-4}$. Then, there is an algorithm that runs in $\mathrm{poly}(n)$ time and outputs an independent set of size at least $10^{-4}n$.*

If the graph has a color class of size at least $(\frac{1}{2} + \Omega(1))n$, then we are already done by the well-known 2-approximation algorithm for minimum vertex cover:

Fact 9.3.2. *If an n-vertex graph G has an independent set of size at least $(\frac{1}{2} + \gamma)n$, then there exists a polynomial-time algorithm that outputs an independent set of size at least $2\gamma n$.*

Below, we state our algorithm.

Algorithm 9.3.3 (Find independent set in a 3-colorable expander)**.**
Input: A graph $G = (V, E)$.
Output: An independent set of G.
Operation: Fix $\gamma = 10^{-3}$ and $\varepsilon = 10^{-4}$.

1. Run the polynomial-time algorithm from Fact 9.3.2 and exit if that outputs an independent set of size at least γn.
2. Solve the degree-d SoS algorithm to obtain a pseudo-distribution μ' that satisfies the almost 3-coloring constraints and the constraints $\mathbb{E}_u[\mathbf{1}(x_u = \sigma)] \leqslant \frac{1}{2} + \gamma$ for all $\sigma \in [3]$ and $\mathbb{E}_u[\mathbf{1}(x_u = \perp)] \leqslant \varepsilon$.
3. Choose a uniformly random set of $t = O(1)$ vertices $i_1, \ldots, i_t \sim [n]$ and draw $(\sigma_{i_1}, \ldots, \sigma_{i_t}) \sim \mu'$. Let μ be the pseudo-distribution obtained by conditioning μ' on $(x_{i_1} = \sigma_{i_1}, \ldots, x_{i_t} = \sigma_{i_t})$.
4. For each $\sigma \in [3]$, let $I_\sigma = \{u \in V : \widetilde{\mathbb{E}}_\mu[\mathbf{1}(x_u = \sigma)] > \frac{1}{2}\}$. Output the largest one.

9.3.1 Almost 3-coloring formulation and agreement

We define an almost 3-coloring of a graph to be an assignment of vertices to $\{1, 2, 3, \perp\}$ where $\{1, 2, 3\}$ are the color classes and the fraction of vertices assigned to $\perp$ is small.

Definition 9.3.4 (Almost 3-coloring constraints)**.** Denote $\Sigma := [3] \cup \{\perp\}$. Given a graph $G = (V, E)$ and parameter $\varepsilon \geqslant 0$, let $\overline{x} = \{\overline{x}_{u,\sigma}\}_{u \in V, \sigma \in \Sigma}$ be indeterminants. We define the almost 3-coloring constraints as follows:

$$\mathcal{A}_G^{\mathsf{Col}}(\overline{x}) := \mathcal{A}^{\mathsf{bool}}(\overline{x}) \cup \left\{ \sum_{\sigma \in \Sigma} \overline{x}_{u,\sigma} = 1,\ \forall u \in V \right\} \cup \{\overline{x}_{u,\sigma}\overline{x}_{v,\sigma} = 0,\ \forall \{u, v\} \in E,\ \sigma \in [3]\}\ .$$

Moreover, with slight abuse of notation, for $t \in \mathbb{N}$ and assignments $\overline{x}^{(1)}, \overline{x}^{(2)}, \ldots, \overline{x}^{(t)}$,

$$\mathcal{A}_G^{\mathsf{Col}}(\overline{\boldsymbol{x}}) := \bigcup_{i \in [t]} \mathcal{A}_G^{\mathsf{Col}}(\overline{x}^{(i)})\,.$$

We will drop the dependence on G when it is clear from context.

Notation. We remark that there is a one-to-one correspondence between almost 3-coloring assignments $x \in \{1,2,3,\perp\}^n$ and $\overline{x} \in \{0,1\}^{n\times 4}$. Even though formally the SoS program is over variables $\overline{x}$, from here on we will use the notation $x \in \{1,2,3,\perp\}^n$ as it is equivalent and more intuitive. For example, we will write $\mathbf{1}(x_u = \sigma)$ to mean $\overline{x}_{u,\sigma}$, and similarly $\widetilde{\Pr}_\mu[x_u = \sigma] = \widetilde{\mathbb{E}}_\mu[\overline{x}_{u,\sigma}]$.

The following definition is almost identical to Definition 9.2.3.

Definition 9.3.5. Let $t \in \mathbb{N}$, and let $\boldsymbol{x} = (x^{(1)}, x^{(2)}, \ldots, x^{(t)})$. For each $\alpha \in \Sigma^t$, we define the following multilinear polynomials,

$$\mathbf{1}(u \leftarrow \alpha) := \prod_{i\in[t]} \mathbf{1}(x_u^{(i)} = \alpha_i)\,, \quad \text{for each } u \in [n]\,,$$
$$\mathsf{wt}(\alpha) := \mathbb{E}_{u\in[n]}[\mathbf{1}(u \leftarrow \alpha)]\,.$$

For convenience, we omit the dependence on $\boldsymbol{x}$.

For $S \subseteq \Sigma^t$, we denote $\mathbf{1}(u \leftarrow S) := \sum_{\alpha\in S} \mathbf{1}(u \leftarrow \alpha)$ and $\mathsf{wt}(S) := \sum_{\alpha\in S} \mathsf{wt}(\alpha)$. Moreover, we will denote $S_\perp := \{\alpha \in \Sigma^t : \exists i \in [t],\ \alpha_i = \perp\}$.

As explained in Section 9.1.3, due to the symmetry of the color classes, we need to define the relative *agreement* between two valid almost 3-colorings according to some permutation $\pi \in \mathbb{S}_3$. For example, consider a coloring $x \in \Sigma^n$ and suppose $y \in \Sigma^n$ is obtained by permuting the 3 color classes of x. The agreement between x and y should be close to 1. Thus, we define the agreement between x and y as

$$\max_{\pi\in\mathbb{S}_3} \mathbb{E}_{u\in V}[\pi(x_u) = y_u \neq \perp]\,.$$

Here for simplicity we assume $\pi(\perp) = \perp$. Formally,

Definition 9.3.6 (Agreement between 2 valid 3-colorings)**.** Let $\pi \in \mathbb{S}_3$. Define

$$S_\pi := \{(\sigma, \pi(\sigma)) : \sigma \in [3]\}\,.$$

For almost 3-colorings $x, y \in \Sigma^n$, we define the agreement between x and y according to permutation π to be

$$\mathrm{agree}_\pi(x,y) := \mathsf{wt}(S_\pi) = \mathbb{E}_{u\in[n]}\left[\sum_{\sigma\in[3]} \mathbf{1}(x_u = \sigma,\ y_u = \pi(\sigma))\right].$$

Furthermore, for any $\ell \in \mathbb{N}$, we write

$$\mathrm{agree}^{(\ell)}(x,y) = \sum_{\pi\in\mathbb{S}_3} \mathrm{agree}_\pi(x,y)^\ell\,.$$

Here $\text{agree}^{(\ell)}(x,y)$ should be viewed as a polynomial approximation of $\max_\pi \text{agree}_\pi(x,y)^\ell$. We note some simple facts (written in SoS form) that will be useful later.

Fact 9.3.7. *For any $t \in \mathbb{N}$, the following can be easily verified:*

(1) $\mathcal{A}^{\mathsf{bool}}(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{2t}\right. \{\mathbf{1}(u \leftarrow \alpha)^2 = \mathbf{1}(u \leftarrow \alpha)\}$, *i.e.,* $\mathbf{1}(u \leftarrow \alpha)$ *satisfies the Booleanity constraint.*

(2) $\mathcal{A}^{\mathsf{Col}}(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{2t}\right. \{\mathbf{1}(u \leftarrow \alpha) \cdot \mathbf{1}(u \leftarrow \beta) = 0\}$ *for $\alpha \neq \beta$. This also implies that $\mathbf{1}(u \leftarrow S)$ satisfies the Booleanity constraint for any $S \subseteq \Sigma^t$.*

(3) $\mathcal{A}^{\mathsf{Col}}(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{t}\right. \{\sum_{\alpha\in\Sigma^t} \mathbf{1}(u \leftarrow \alpha) = 1\}$, *thus* $\mathcal{A}^{\mathsf{Col}}(\boldsymbol{x}) \left|\frac{\boldsymbol{x}}{t}\right. \{\sum_{\alpha\in\Sigma^t} \mathsf{wt}(\alpha) = 1\}$.

Each S_π corresponds to a triangle in Figure 9.2, and we see that there are two ways to partition the graph into 3 disjoint triangles. The next lemma can essentially be proved by looking at Figure 9.2 (there $S_\perp$ is not shown), and it is crucial for our analysis.

Lemma 9.3.8. *Let G be a regular graph. Let $\mathbb{S}_3^+$ be the set of 3 permutations with sign (a.k.a. parity) $+1$ and $\mathbb{S}_3^-$ be the ones with sign -1. Then,*

$$\mathcal{A}_G^{\mathsf{Col}}(x,y) \left|\frac{x,y}{2}\right. \left\{ \sum_{\pi\in\mathbb{S}_3^+} \mathsf{wt}(S_\pi) = \sum_{\pi\in\mathbb{S}_3^-} \mathsf{wt}(S_\pi) = 1 - \mathsf{wt}(S_\perp) \right\}.$$

Moreover,

$$\mathcal{A}_G^{\mathsf{Col}}(x,y) \left|\frac{x,y}{2}\right. \left\{ \sum_{\pi\in\mathbb{S}_3} e(S_\pi, \overline{S}_\pi) \leqslant 1 \right\}.$$

Proof. The first statement follows by noting that for each $i,j \in [3]$, there are exactly two permutations with opposite signs that map i to j. Thus, $\{S_\pi : \pi \in \mathbb{S}_3^+\} \cup \{S_\perp\}$ and $\{S_\pi : \pi \in \mathbb{S}_3^-\} \cup \{S_\perp\}$ are partitions of the whole graph. One can also prove this directly from Figure 9.2.

For the second statement, note that each edge $(i_1,j_1),(i_2,j_2) \in [3]^2$ in the gadget uniquely identifies the permutation π such that $\pi(i_1) = j_1$ and $\pi(i_2) = j_2$. This means that each edge not incident to S_π is contained in exactly one S_π, and we have $\sum_\pi e(S_\pi, S_\pi) = e(\overline{S}_\perp, \overline{S}_\perp) \geqslant 1 - 2\mathsf{wt}(S_\perp)$. On the other hand, from the first statement we have $\sum_\pi \mathsf{wt}(S_\pi) = 2 - 2\mathsf{wt}(S_\perp)$ Thus, $\sum_\pi e(S_\pi, \overline{S}_\pi) = \sum_\pi (\mathsf{wt}(S_\pi) - e(S_\pi, S_\pi)) \leqslant (2 - 2\mathsf{wt}(S_\perp)) - (1 - 2\mathsf{wt}(S_\perp)) = 1$. □

9.3.2 Large spectral gap implies large agreement

Lemma 9.3.9. *Let G be a d-regular n-vertex graph with $\lambda_2 := \lambda_2(G) > 0$. Then,*

$$\mathcal{A}_G^{\mathsf{Col}}(x,y) \left|\frac{x,y}{4}\right. \left\{ \sum_{\pi\in\mathbb{S}_3} \mathsf{wt}(S_\pi)^2 \geqslant 2 - \frac{1}{1-\lambda_2} - 2\mathsf{wt}(S_\perp) \right\}.$$

Proof. Fix a permutation $\pi \in \mathbb{S}_3$, and let $y_u = \mathbf{1}(u \leftarrow S_\pi)$. By Lemma 9.2.8, we have that

$$\mathcal{A}_G^{\mathsf{Col}}(x,y) \left|\frac{x,y}{4}\right. \left\{ e(S_\pi, \overline{S}_\pi) = \frac{1}{nd} y^\top L_G y \geqslant \frac{1}{n} \cdot (1-\lambda_2) \left(\|y\|_2^2 - \frac{1}{n} \langle \vec{1}, y \rangle^2 \right) \right\}$$

Since y_u satisfies the booleanity constraints, we have $\frac{1}{n}(\|y\|_2^2 - \frac{1}{n}\langle \vec{1}, y\rangle^2) = \mathbb{E}_u[y_u] - \mathbb{E}_u[y_u]^2 = \mathsf{wt}(S_\pi)(1-\mathsf{wt}(S_\pi))$. Thus,

$$\mathcal{A}_G^{\mathsf{Col}}(x,y) \left|\frac{x,y}{4}\right. \left\{ e(S_\pi, \overline{S}_\pi) \geqslant (1-\lambda_2) \cdot \mathsf{wt}(S_\pi)(1-\mathsf{wt}(S_\pi)) \right\} .$$

Next, we sum over $\pi \in \mathbb{S}_3$. By Lemma 9.3.8, on the left-hand side we have $\sum_\pi e(S_\pi, \overline{S}_\pi) \leqslant 1$, and on the right-hand side we have $(1-\lambda_2)\sum_\pi \mathsf{wt}(S_\pi)(1-\mathsf{wt}(S_\pi)) = (1-\lambda_2)(2 - 2\mathsf{wt}(S_\perp) - \sum_\pi \mathsf{wt}(S_\pi)^2)$. Rearranging this completes the proof. □

In Theorem 9.3.1, we assume that the graph has spectral gap $1-\lambda_2 \geqslant 1-\gamma$ and the almost 3-coloring assignments satisfy $\mathsf{wt}(S_\perp) \leqslant \mathsf{wt}(\{\perp *\}) + \mathsf{wt}(\{*\perp\}) \leqslant 2\varepsilon$ for some small enough constants ε, γ. Thus, by Lemmas 9.3.8 and 9.3.9, the 6 variables $\{\mathsf{wt}(S_\pi)\}_{\pi \in \mathbb{S}_3}$ satisfy that $\sum_{\pi \in \mathbb{S}_3^+} \mathsf{wt}(S_\pi) = \sum_{\pi \in \mathbb{S}_3^-} \mathsf{wt}(S_\pi) \in [1-2\varepsilon, 1]$ and $\sum_\pi \mathsf{wt}(S_\pi)^2 \geqslant 1 - O(\gamma + \varepsilon)$. On the other hand, recall from Definition 9.3.6 that $\mathsf{wt}(S_\pi) = \mathsf{agree}_\pi(x,y)$.

We would like to prove Claim 9.1.8: assuming $\mathsf{agree}_\pi(x,y) \leqslant \frac{1}{2} + \gamma$ for all π, then one of $\{\mathsf{wt}(S_\pi)\}_{\pi \in \mathbb{S}_3^+}$ and one of $\{\mathsf{wt}(S_\pi)\}_{\pi \in \mathbb{S}_3^-}$ must be small. This is captured in the following lemma:

Lemma 9.3.10. *Fix $\gamma \in [0, 0.01]$. Let $z_1, z_2, \ldots, z_6$ be such that $0 \leqslant z_i \leqslant \frac{1}{2} + \gamma$ and $z_1 + z_2 + z_3 = z_4 + z_5 + z_6 \leqslant 1$. Suppose $\|z\|_2^2 \geqslant 1 - \gamma$. Then, one of z_1, z_2, z_3 and one of z_4, z_5, z_6 must be $\leqslant 8\gamma$.*

Proof. For any $i \in [6]$, we have $\|z\|_2^2 \leqslant z_i^2 + (\frac{1}{2} + \gamma)\sum_{j \neq i} z_j$ since $z_j \leqslant \frac{1}{2} + \gamma$ for all j. Then since $\|z\|_1 \leqslant 2$, for all $i \in [6]$ we have

$$\|z\|_2^2 \leqslant z_i^2 + \left(\frac{1}{2} + \gamma\right)(2 - z_i) = 1 + 2\gamma - z_i\left(\frac{1}{2} + \gamma - z_i\right) .$$

Since $\|z\|_2^2 \geqslant 1 - \gamma$, it follows that

$$z_i\left(\frac{1}{2} + \gamma - z_i\right) \leqslant 3\gamma , \quad \forall i \in [6] .$$

Then, by solving a quadratic inequality, one can verify that when $\gamma \leqslant 0.01$, the above implies that either $z_i \leqslant 8\gamma$ or $z_i \geqslant \frac{1}{2} - 8\gamma$. Therefore, since $z_1 + z_2 + z_3 \leqslant 1$, z_1, z_2, z_3 cannot all be the latter, i.e., one of them must be $\leqslant 8\gamma$. Similarly for z_4, z_5, z_6. □

We next consider 3 almost 3-coloring assignments. Recall that $\Sigma = [3] \cup \{\perp\}$.

Lemma 9.3.11. *Let $0 \leqslant \varepsilon, \gamma \leqslant 0.001$. Let $\{w(\alpha)\}_{\alpha \in \Sigma^3}$ be variables such that $0 \leqslant w(\alpha) \leqslant 1$ and $\sum_\alpha w(\alpha) = 1$. For any $S \subseteq \Sigma^3$, denote $w(S) = \sum_{\alpha \in S} w(\alpha)$, and let*

$$\begin{aligned} S_\pi^{(12)} &= \{(\sigma, \pi(\sigma), *) : \sigma \in [3]\}\ , \\ S_\pi^{(13)} &= \{(\sigma, *, \pi(\sigma)) : \sigma \in [3]\}\ , \\ S_\pi^{(23)} &= \{(*, \sigma, \pi(\sigma)) : \sigma \in [3]\}\ , \end{aligned}$$

Suppose $w(\sigma * *), w(*\sigma*), w(* * \sigma) \leqslant \frac{1}{2} + \gamma$ and $w(\perp * *), w(*\perp*), w(* * \perp) \leqslant \varepsilon$. Moreover, suppose $\sum_{\pi \in \mathfrak{S}_3} w(S_\pi^{(ij)})^2 \geqslant 1 - \gamma$ for all pairs $i < j \in [3]$, then there must be some π and $i < j$ such that $w(S_\pi^{(ij)}) \geqslant \frac{1}{2} + \gamma$.

Proof. Suppose by contradiction that all $w(S_\pi^{(ij)}) \leqslant \frac{1}{2} + \gamma$. Let $\mathfrak{S}_3^+$ be the set of 3 permutations with sign (a.k.a. parity) $+1$ and $\mathfrak{S}_3^-$ be the ones with sign -1. For each pair $i < j$ (say, (12) for now), by Lemma 9.3.8 we have $\sum_{\pi \in \mathfrak{S}_3^+} w(S_\pi^{(12)}) = \sum_{\pi \in \mathfrak{S}_3^-} w(S_\pi^{(12)}) \leqslant 1$.

Therefore, the 6 variables $\{w(S_\pi^{(12)})\}_{\pi \in \mathfrak{S}_3^+} \cup \{w(S_\pi^{(12)})\}_{\pi \in \mathfrak{S}_3^-}$ satisfy the conditions in Lemma 9.3.10, and thus there are some $\pi^+ \in \mathfrak{S}_3^+$ and $\pi^- \in \mathfrak{S}_3^-$ such that $w(S_{\pi^+}^{(12)}), w(S_{\pi^-}^{(12)}) \leqslant 8\gamma$. Furthermore, note that since π^+ and π^- have different signs, $S_{\pi^+}^{(12)}$ and $S_{\pi^-}^{(12)}$ intersect in exactly $(\beta_1, \beta_2, *)$ for some $\beta_1, \beta_2 \in [3]$. In fact, β_1, β_2 uniquely determine π^+ and π^-, as there are exactly two permutations with different signs that map β_1 to β_2.

Assume without loss of generality (due to symmetry) that $\beta_1 = \beta_2 = 1$, thus we have $S_{\pi^+}^{(12)} = \{11*, 22*, 33*\}$ and $S_{\pi^-}^{(12)} = \{11*, 23*, 32*\}$. Let $T^{(12)} := [3]^3 \setminus (S_{\pi^+}^{(12)} \cup S_{\pi^-}^{(12)}) = \{12*, 13*, 21*, 31*\}$, which equals $\{1 * *, *1*\} \setminus \{11*\}$ (here we do not include $\perp$). Notice the structure of $T^{(12)}$ — ignoring the third assignment, $T^{(12)}$ forms a 2×2 bipartite graph (between $\{12*, 13*\}$ and $\{21*, 31*\}$ in this case; see Figure 9.3) where one assignment labels the entire left-hand side as one color while the other assignment labels the entire right-hand side as one color.

Now, for all 3 pairs $(12), (13), (23)$, consider $T := T^{(12)} \cap T^{(23)} \cap T^{(13)} \subseteq [3]^3$. First, we have $w(T) \geqslant 1 - 48\gamma - \mathrm{wt}(S_\perp) \geqslant 1 - 48\gamma - 3\varepsilon$, since $\mathrm{wt}(S_\perp) \leqslant 3\varepsilon$ by assumption. Next, we claim that for all choices of π^+ and π^- for each pair, T can contain at most 4 strings in $[3]^3$ and must form a 2×2 bipartite structure such that each assignment colors one side with one color.

Let $T^{(12)} = \{a_1 * *, *a_2*\} \setminus \{a_1a_2*\}$, $T^{(13)} = \{b_1 * *, * * b_2\} \setminus \{b_1 * b_2\}$, and $T^{(23)} = \{*c_1*, * * c_2\} \setminus \{*c_1c_2\}$ for some $a_1, a_2, b_1, b_2, c_1, c_2 \in [3]$. We split into several cases:

- $a_1 = b_1$: in this case, $T^{(12)} \cap T^{(13)} = (\{a_1 * *\} \setminus \{a_1a_2*, a_1 * b_2\}) \cup (\{*a_2b_2\} \setminus \{a_1a_2b_2\})$.
  1. $c_1 \neq a_2, c_2 \neq b_2$: then, $T = (\{a_1c_1*\} \setminus \{a_1c_1b_2, a_1c_1c_2\}) \cup (\{a_1 * b_2\} \setminus \{a_1a_2b_2, a_1c_1b_2\})$, i.e., 2 strings in $[3]^3$. For example, $T = \{123, 131\}$.

2. $c_1 = a_2$, $c_2 \neq b_2$: then, $T = (\{a_1 * c_2\} \setminus \{a_1a_2c_2\}) \cup (\{*a_2b_2\} \setminus \{a_1a_2b_2\})$, i.e., 4 strings in $[3]^3$. For example, $T = \{122, 132, 211, 311\}$.
3. $c_1 = a_2$, $c_2 = b_2$: then, $T = \varnothing$.

- $a_1 \neq b_1$: in this case, $T^{(12)} \cap T^{(13)} = (\{a_1 * b_2\} \setminus \{a_1a_2b_2\}) \cup (\{b_1a_2*\} \setminus \{b_1a_2b_2\})$, which is already the same case as the second case above.

For the case when $T = \varnothing$ or T contains 2 strings, we have $w(T) \leqslant w(\sigma **)$ for some $\sigma \in [3]$, which means $1 - 48\gamma - 3\varepsilon \leqslant \frac{1}{2} + \gamma$. This is a contradiction.

For the case when T contains 4 strings, let $T = \{\alpha^1, \alpha^2, \beta^1, \beta^2\}$ such that $\{\alpha^1, \alpha^2\}$ and $\{\beta^1, \beta^2\}$ form the bipartite structure. Assume without loss of generality that the first assignment labels the left with the same color: $\alpha^1_1 = \alpha^2_1 \neq \beta^1_1, \beta^2_1$, and the second and third label the right with the same color: $\beta^1_2 = \beta^2_2 \neq \alpha^1_2, \alpha^2_2$ and $\beta^1_3 = \beta^2_3 \neq \alpha^1_3, \alpha^2_3$. Observe that $w(\alpha^1) + w(\alpha^2) \leqslant w(\alpha^1_1 **) \leqslant \frac{1}{2} + \gamma$ and $w(\beta^1) + w(\beta^2) \leqslant w(*\beta^1_2*) \leqslant \frac{1}{2} + \gamma$ by the assumptions. Since $w(T) \geqslant 1 - 48\gamma - 3\varepsilon$, it follows that $w(\alpha^1) + w(\alpha^2)$ and $w(\beta^1) + w(\beta^2) \geqslant \frac{1}{2} - 49\gamma - 3\varepsilon$.

On the other hand, $w(\alpha^1) + w(\beta^1) + w(\beta^2) \leqslant w(*\alpha^1_2\alpha^1_3) + w(*\beta^1_2\beta^1_3) \leqslant w(S^{(23)}_\pi)$ and $w(\alpha^2) + w(\beta^1) + w(\beta^2) \leqslant w(*\alpha^2_2\alpha^2_3) + w(*\beta^1_2\beta^1_3) \leqslant w(S^{(23)}_{\pi'})$ for some permutations $\pi, \pi' \in \mathbb{S}_3$. However, this means that one of $w(S^{(23)}_\pi)$, $w(S^{(23)}_{\pi'})$ is at least $\frac{3}{2}(\frac{1}{2} - 49\gamma - 3\varepsilon) > \frac{1}{2} + \gamma$ when $\varepsilon, \gamma \leqslant 0.001$, which is a contradiction. □

We next formalize Lemma 9.3.11 as an SoS proof.

Lemma 9.3.12 (SoS version of Lemma 9.3.11)**.** *Fix constants $\varepsilon, \gamma \in (0, 0.001]$ and $\ell \in \mathbb{N}$. Let $S^{(ij)}_\pi \subseteq [3]^3$ be as defined in Lemma 9.3.11, and let $\{w(\alpha)\}_{\alpha \in \Sigma^3}$ be indeterminants. Let $\mathcal{A}$ be the set of constraints including*

(1) $0 \leqslant w(\alpha) \leqslant 1$,
(2) $\sum_{\alpha \in \Sigma^3} w(\alpha) = 1$,
(3) $w(\sigma **), w(*\sigma*), w(**\sigma) \leqslant \frac{1}{2} + \gamma$ *for all* $\sigma \in [3]$,
(4) $w(\perp **), w(*\perp*), w(**\perp) \leqslant \varepsilon$,
(5) $\sum_{\pi \in \mathbb{S}_3} w(S^{(ij)}_\pi)^2 \geqslant 1 - \gamma$ *for all pairs* $i < j \in [3]$.

Then, there exists an integer $d = d(\varepsilon, \gamma, \ell)$ such that

$$\mathcal{A} \vdash^{\{w(\alpha)\}}_{d} \left\{ \sum_{i<j\in[3]} \sum_{\pi \in \mathbb{S}_3} w\left(S^{(ij)}_\pi\right)^\ell \geqslant \left(\frac{1+\gamma}{2}\right)^\ell \right\}.$$

Proof. Lemma 9.3.11 shows that assuming constraints $\mathcal{A}$, there must be some $w(S^{(ij)}_\pi) \geqslant \frac{1}{2} + \gamma$. This immediately implies that $\sum_{i<j\in[3]} \sum_{\pi \in \mathbb{S}_3} w(S^{(ij)}_\pi)^\ell \geqslant (\frac{1}{2} + \gamma)^\ell$.

Define $f(w) := \sum_{i<j\in[3]} \sum_{\pi \in \mathbb{S}_3} w(S^{(ij)}_\pi)^\ell - (\frac{1+\gamma}{2})^\ell$, a degree-$\ell$ polynomial in 64 variables with bounded coefficients. Note that $\mathcal{A}$ defines a subset $A \subseteq \mathbb{R}^n$ which is compact,

and $\min_{w \in A} f(w) \geqslant \theta$ for some constant $\theta = \theta(\gamma, \ell) > 0$. Thus, by the Positivstellensatz (Fact 2.5.5), $f(w) \geqslant 0$ has an SoS proof of degree d depending on $\varepsilon, \gamma, \ell$. □

9.3.3 Rounding with large agreement

We prove the following key lemma that large agreement and small correlation imply rounding. Using this, we finish the proof of Theorem 9.3.1 at the end of this section.

Lemma 9.3.13 (Rounding with large agreement)**.** *Fix $\gamma \in (0,1)$. There exist $\ell \in \mathbb{N}$ and $\delta \in (0,1)$ such that given a degree-ℓ pseudo-distribution μ satisfying the almost 3-coloring constraints such that*

$$\widetilde{\mathbb{E}}_{(x,y)\sim\mu^{\otimes 2}}\left[\mathrm{agree}^{(\ell)}(x,y)\right] \geqslant \left(\frac{1}{2}+\gamma\right)^{\ell},$$

and suppose μ is almost ℓ-wise independent on average:

$$\mathbb{E}_{u_1,\dots,u_\ell\in[n]} \mathrm{KL}(\mu(X_{u_1},\dots,X_{u_\ell}) \| \mu(X_{u_1}) \times \cdots \times \mu(X_{u_\ell})) \leqslant \delta,$$

then one of the sets $I_\sigma = \{u \in V : \widetilde{\Pr}_\mu[x_u = \sigma] > \frac{1}{2}\}$ for $\sigma \in [3]$ has size at least $\Omega(\gamma n)$.

The proof of Lemma 9.3.13 relies on the following definition.

Definition 9.3.14 (Collision probability)**.** Given a pseudo-distribution μ over Σ^n, we define the collision probability of a vertex $u \in [n]$ to be

$$\mathsf{CP}_\mu(x_u) := \widetilde{\mathbb{E}}_{x,x'\sim\mu}[\mathbf{1}(x_u = x'_u \neq \perp)] = \sum_{\sigma\in[3]} \widetilde{\Pr}_\mu[x_u = \sigma]^2.$$

Further, the (average) collision probability $\mathsf{CP}(\mu) = \mathbb{E}_{u\in[n]}\mathsf{CP}(x_u)$.

We next show a simple lemma which states that large collision probability implies a large fraction of vertices with $\widetilde{\Pr}_\mu[x_u = \sigma] > \frac{1}{2}$ for some color $\sigma \in [3]$ (and they form an independent set due to Fact 9.1.1).

Lemma 9.3.15. *Suppose a pseudo-distribution μ over Σ^n has collision probability $\mathsf{CP}(\mu) \geqslant \frac{1}{2}+\gamma$ for some $\gamma \in (0, 1/2]$, then there is a $\sigma \in [3]$ such that at least $\gamma/3$ fraction of $u \in [n]$ have $\widetilde{\Pr}_\mu[x_u = \sigma] \geqslant \frac{1}{2}+\frac{\gamma}{2}$.*

Proof. Observe that $\mathsf{CP}_\mu(x_u) \leqslant \max_{\sigma\in[3]} \widetilde{\Pr}_\mu[x_u = \sigma]$ because $\sum_{\sigma\in[3]} \widetilde{\Pr}_\mu[x_u = \sigma] \leqslant 1$. Thus, we have $\mathbb{E}_{u\in[n]} \max_{\sigma\in[3]} \widetilde{\Pr}_\mu[x_u = \sigma] \geqslant \frac{1}{2}+\gamma$. This implies that at least γ fraction of $u \in [n]$ has $\max_{\sigma\in[3]} \widetilde{\Pr}_\mu[x_u = \sigma] \geqslant \frac{1}{2}+\frac{\gamma}{2}$. Then, there must be a $\sigma \in [3]$ such that at least $\gamma/3$ fraction of $u \in [n]$ have $\widetilde{\Pr}_\mu[x_u = \sigma] \geqslant \frac{1}{2}+\frac{\gamma}{2}$. □

In light of Lemma 9.3.15, to prove Lemma 9.3.13, it suffices to show that the pseudo-distribution μ has large collision probability.

Proof of Lemma 9.3.13. We first prove an upper bound on $\widetilde{\mathbb{E}}[\mathrm{agree}^{(\ell)}(x,y)]$:

$$\widetilde{\mathbb{E}}_{x,y\sim\mu}\left[\mathrm{agree}^{(\ell)}(x,y)\right] \leqslant 6\left(\mathsf{CP}(\mu)^{\ell} + 2\sqrt{2\delta}\right). \tag{9.4}$$

For any permutation π, recalling Definition 9.3.6,

$$\begin{aligned}\mathrm{agree}_{\pi}(x,y)^{\ell} &= \Pr_{u_1,\dots,u_\ell\in[n]}\left[x_{u_i} = \pi(y_{u_i}) \neq \perp,\ \forall i\in[\ell]\right]\\ &= \mathbb{E}_{u_1,\dots,u_\ell\in[n]}\sum_{\sigma_1,\dots,\sigma_\ell\in[3]}\mathbf{1}\left(x_{u_i} = \pi(y_{u_i}) = \sigma_i,\ \forall i\in[\ell]\right).\end{aligned}$$

Thus, summing over $\pi\in\mathbb{S}_3$ and using the independence between x and y,

$$\begin{aligned}\widetilde{\mathbb{E}}_{\mu^{\otimes 2}}\left[\mathrm{agree}^{(\ell)}(x,y)\right] &= \mathbb{E}_{u_1,\dots,u_\ell\in[n]}\sum_{\pi\in\mathbb{S}_3}\sum_{\sigma_1,\dots,\sigma_\ell\in[3]}\widetilde{\Pr}_{\mu}[x_{u_i}=\sigma_i,\ \forall i]\cdot\widetilde{\Pr}_{\mu}[x_{u_i}=\pi^{-1}(\sigma_i),\ \forall i]\\ &\leqslant \mathbb{E}_{u_1,\dots,u_\ell\in[n]}\sum_{\pi\in\mathbb{S}_3}\sum_{\sigma_1,\dots,\sigma_\ell\in[3]}\frac{1}{2}\left(\widetilde{\Pr}_{\mu}[x_{u_i}=\sigma_i,\ \forall i]^2+\widetilde{\Pr}_{\mu}[x_{u_i}=\pi^{-1}(\sigma_i),\ \forall i]^2\right)\end{aligned}$$

then since the summation is over all permutations π and $\sigma_1,\dots,\sigma_\ell\in[3]$,

$$= |\mathbb{S}_3|\cdot\mathbb{E}_{u_1,\dots,u_\ell\in[n]}\sum_{\sigma_1,\dots,\sigma_\ell\in[3]}\widetilde{\Pr}_{\mu}[x_{u_i}=\sigma_i,\ \forall i]^2. \tag{9.5}$$

Now, suppose $\mathbb{E}_{u_1,\dots,u_\ell\in[n]}\mathrm{KL}(\mu(X_{u_1},\dots,X_{u_\ell})\|\mu(X_{u_1})\times\cdots\times\mu(X_{u_\ell}))\leqslant\delta$, then by Pinsker's inequality (Fact 2.5.8) and Jensen's inequality,

$$\mathbb{E}_{u_1,\dots,u_\ell\in[n]}\sum_{\sigma_1,\dots,\sigma_\ell\in[3]}\left|\widetilde{\Pr}_{\mu}[x_{u_i}=\sigma_i,\ \forall i]-\prod_{i=1}^{\ell}\widetilde{\Pr}_{\mu}[x_{u_i}=\sigma_i]\right|\leqslant\sqrt{2\delta}.$$

Then, using the fact that $p^2-q^2=(p-q)(p+q)\leqslant 2|p-q|$ for all $p,q\in[0,1]$, we can bound Eq. (9.5) by

$$\begin{aligned}\widetilde{\mathbb{E}}_{\mu^{\otimes 2}}\left[\mathrm{agree}^{(\ell)}(x,y)\right] &\leqslant 6\left(\mathbb{E}_{u_1,\dots,u_\ell\in[n]}\sum_{\sigma_1,\dots,\sigma_\ell\in[3]}\prod_{i=1}^{\ell}\widetilde{\Pr}_{\mu}[x_{u_i}=\sigma_i]^2+2\sqrt{2\delta}\right)\\ &= 6\left(\left(\mathbb{E}_{u\in[n]}\sum_{\sigma\in[3]}\widetilde{\Pr}_{\mu}[x_u=\sigma]^2\right)^{\ell}+2\sqrt{2\delta}\right)\\ &= 6\left(\mathsf{CP}(\mu)^{\ell}+2\sqrt{2\delta}\right).\end{aligned}$$

This completes the proof of Eq. (9.4).

Therefore, since $\widetilde{\mathbb{E}}_{\mu^{\otimes 2}}[\mathrm{agree}^{(\ell)}(x,y)]\geqslant(\frac{1}{2}+\gamma)^{\ell}$, we have

$$\mathsf{CP}(\mu)^{\ell}\geqslant\frac{1}{6}\left(\frac{1}{2}+\gamma\right)^{\ell}-2\sqrt{2\delta}.$$

For any $\gamma > 0$, there exists a large enough $\ell \in \mathbb{N}$ and small enough δ (here $\ell = O(1/\gamma)$ and $\delta = 2^{-O(\ell)}$ suffice) such that the above is at least $(\frac{1}{2} + \frac{\gamma}{2})^\ell$, which means that $\mathsf{CP}(\mu) \geqslant \frac{1}{2} + \frac{\gamma}{2}$.

Then, let $I_\sigma = \{u : \widetilde{\Pr}_\mu[x_u = \sigma] > \frac{1}{2}\}$ for $\sigma \in [3]$, which are independent sets. By Lemma 9.3.15, one of the sets has size at least $\Omega(\gamma n)$, thus completing the proof. □

We can now finish the analysis of Algorithm 9.3.3 and prove Theorem 9.3.1.

Proof of Theorem 9.3.1. Fix $\gamma = 10^{-3}$. If there is an independent set in G with size larger than $(\frac{1}{2} + \gamma)n$, then Fact 9.3.2 says that we can find an independent set of size at least $2\gamma n$, and the first step of Algorithm 9.3.3 would succeed. Therefore, let us assume that this is not the case, and in particular the second step of the algorithm outputs a valid pseudo-distribution μ' satisfying the constraints listed therein.

Fix $\ell = 10^4$, and let δ be some small enough constant as in Lemma 9.3.13. First, by Lemma 2.5.9, we can assume that the third step of Algorithm 9.3.3 reduces the total ℓ-wise correlation of μ' to output a pseudo-distribution μ with total ℓ-wise correlation $\leqslant \delta$.

By Lemma 9.3.9 we have

$$\mathcal{A}_G^{\mathsf{Col}}(x,y) \Big|^{x,y}_{4} \left\{ \sum_{\pi \in \mathbb{S}_3} \mathsf{wt}(S_\pi)^2 \geqslant 2 - \frac{1}{1-\lambda_2} - 2\varepsilon \geqslant 1 - \gamma \right\}$$

since $\widetilde{\mathbb{E}}[\mathsf{wt}(S_\perp)] \leqslant 2\varepsilon$ by the constraints on μ and $\lambda_2 \leqslant 10^{-4}$, $\varepsilon \leqslant 10^{-4}$. Then, consider 3 assignments $\boldsymbol{x} = (x^{(1)}, x^{(2)}, x^{(3)})$. By Lemma 9.3.12, it follows that the pseudo-distribution μ satisfies

$$\widetilde{\mathbb{E}}_{\mu^{\otimes 3}} \sum_{i<j \in [3]} \sum_{\pi \in \mathbb{S}_3} w\left(S_\pi^{(ij)}\right)^\ell \geqslant \left(\frac{1+\gamma}{2}\right)^\ell.$$

By symmetry between the 3 assignments, it follows that

$$\widetilde{\mathbb{E}}_{\mu^{\otimes 3}} \sum_{\pi \in \mathbb{S}_3} w(S_\pi)^\ell = \widetilde{\mathbb{E}}_{\mu^{\otimes 2}} \left[\mathsf{agree}^{(\ell)}(x,y)\right] \geqslant \frac{1}{3}\left(\frac{1+\gamma}{2}\right)^\ell \geqslant \left(\frac{1}{2} + \frac{\gamma}{4}\right)^\ell$$

since $\ell = 10^4$. Then, Lemma 9.3.13 shows that one of the sets $I_\sigma = \{u : \widetilde{\Pr}_\mu[x_u = \sigma] > \frac{1}{2}\}$ for $\sigma \in [3]$ has size at least $\Omega(\gamma n)$. The degree of the SoS algorithm required is $O(1/\delta) + d = O(1)$, where $d = d(\varepsilon, \gamma, \ell)$ is the constant from Lemma 9.3.12. □

9.4 Hardness of finding independent sets in k-colorable expanders

Bansal and Khot [BK09] proved the following hardness result of finding linear-sized independent sets in almost 2-colorable graphs, which is a strengthening of [KR08].

Proposition 9.4.1 ([BK09])**.** *Assuming the Unique Games Conjecture, for any constants $\varepsilon, \gamma > 0$, given an n-vertex graph G, it is NP-hard to decide between*

1. *G has 2 disjoint independent sets of size $(\frac{1}{2} - \varepsilon)n$,*
2. *G has no independent set of size larger than γn.*

Moreover, the above holds if we additionally assume that the graph has degrees $o(n)$.

Proposition 9.4.2 (Formal version of Proposition 6.3.1)**.** *Assuming the Unique Games Conjecture, for any constants $\varepsilon, \gamma > 0$, given a regular n-vertex graph G which is a one-sided spectral expander with $\lambda_2(G) \leqslant o_n(1)$, it is NP-hard to decide between*

1. *G is ε-almost 4-colorable,*
2. *G has no independent set of size larger than γn.*

Proof. We start the reduction from Proposition 9.4.1. Given a graph G, we add a regular bipartite graph H (potentially introducing multi-edges) such that H has degree $\Omega(n)$ and the second eigenvalue of its normalized adjacency matrix $\lambda_2(H) = o_n(1)$. If G is not regular, we can make the resulting graph G' regular by removing $o(1)$ fraction of edges, denoted H', from H.

If G is ε-almost 2-colorable, then G' is clearly ε-almost 4-colorable (since H is 2-colorable). On the other hand, adding edges cannot increase the size of the maximum independent set.

Next, we prove that G' has small normalized second eigenvalue. We can assume that G and H' have maximum degrees $d_G, d_{H'} = o(n)$ while H has degree $d_H = \Omega(n)$. Then, $\lambda_2(G') = \frac{1}{d_{G'}}\lambda_2(A_G + A_H - A_{H'}) = \frac{1}{d_{G'}} \cdot \max_{x \perp \vec{1}, \|x\|_2 = 1} x^\top (A_G + A_H - A_{H'})x \leqslant o_n(1)$. □

Hardness for k-colorable graphs. In this case (as opposed to almost k-colorable), we need a hardness conjecture with *perfect completeness*. The natural candidate is the 2-to-1 (or d-to-1) conjecture:

Conjecture 9.4.3 (2-to-1 conjecture with perfect completeness [Kho02])**.** *For every $\varepsilon > 0$, there exists some $R \in \mathbb{N}$ such that given a label cover instance ψ with alphabet size R such that all constraints are 2-to-2 constraints, it is NP-hard to decide between*

1. *ψ is satisfiable,*
2. *no assignment satisfies more than ε fraction of the constraints in ψ.*

Dinur, Mossel and Regev [DMR06] introduced the following variant of the 2-to-1 conjecture. We note that the "$\ltimes$" constraints (termed "alpha" or "fish-shaped" constraints) have also appeared in [DS05]. See [DS05, DMR06] for a precise statement.

Conjecture 9.4.4 ("$\ltimes$" variant of the 2-to-1 conjecture [DMR06]). *Conjecture 9.4.3 is true even assuming that all constraints in the label cover instance are "$\ltimes$" constraints.*

Dinur, Mossel and Regev [DMR06] proved the following,

Proposition 9.4.5. *Assuming Conjecture 9.4.4, for any constant $\gamma > 0$, given an n-vertex graph $G = (V, E)$, it is* NP*-hard to decide between*

1. *G is 3-colorable,*
2. *G has no independent set of size larger than γn.*

In particular, the gadget used to prove Proposition 9.4.5 is regular, hence we can additionally assume that the graph is regular. Moreover, we can assume that the degrees are $o(n)$.

With Proposition 9.4.5, we can prove the following:

Proposition 9.4.6. *Assuming Conjecture 9.4.4, for any constant $\gamma > 0$, given a regular n-vertex graph G which is a one-sided spectral expander with $\lambda_2(G) \leqslant o_n(1)$, it is NP-hard to decide between*

1. *G is 6-colorable,*
2. *G has no independent set of size larger than γn.*

Proof. Given a graph G, the reduction is to add a regular bipartite graph H (potentially introducing multi-edges) such that H has degree $\Omega(n)$ and the second eigenvalue of its normalized adjacency matrix $\lambda_2(H) = o_n(1)$. If G is k-colorable, then the resulting graph G' is clearly $2k$-colorable (since H is 2-colorable). On the other hand, adding edges cannot increase the size of the maximum independent set.

Next, we prove that G' has small normalized second eigenvalue. Since we can assume that G has degree $d_G = o(n)$ while H has degree $d_H = \Omega(n)$. Then, $\lambda_2(G') \leqslant \frac{1}{d_G + d_H}(d_G \lambda_2(G) + d_H \lambda_2(H)) \leqslant o_n(1)$. □

9.5 Rounding independent sets via Karger-Motwani-Sudan

In this section, we recall a folklore result (we were unable to find a reference, though this argument seems to be known to experts) that extends the rounding algorithm of Karger, Motwani and Sudan [KMS98] to prove the following:

Theorem 9.5.1. *For any $\varepsilon > 0$, there exists a polynomial-time algorithm such that given an n-vertex graph containing an independent set of size $(1/2 - \varepsilon)n$, it finds an independent set of size at least $(\varepsilon n)^{1 - O(\varepsilon)}$.*

We first prove the following crucial lemma.

Lemma 9.5.2. *Let $G = (V, E)$ be a graph and $\varepsilon > 0$. Suppose each vertex $i \in V$ is associated with a unit vector u_i such that for all $(i, j) \in E$, $\langle u_i, u_j \rangle \leqslant -1 + \varepsilon$. Then, there is a polynomial-time algorithm that finds an independent set in G of size at least $n^{1-O(\varepsilon)}$.*

Proof. Set $t := 4\sqrt{\varepsilon \log n}$. The algorithm is as follows,

(1) Sample a Gaussian vector $g \sim \mathcal{N}(0, \mathbb{I}_n)$.
(2) Let $S := \{i \in V : \langle g, u_i \rangle \geqslant t\}$.
(3) Output $T := \{i \in S : \forall j \in N(i),\ j \notin S\}$.

Here, $N(i)$ denotes the set of neighbors of i. By definition, T is an independent set. We next claim that in expectation over g, $|T| \geqslant n^{1-O(\varepsilon)}$, which finishes the proof.

First, note that $\Pr[i \in S] = \Pr_{h \sim \mathcal{N}(0,1)}[h \geqslant t] \geqslant \Omega(\frac{1}{t} e^{-t^2/2}) \geqslant n^{-O(\varepsilon)}$. Next, for each $i \in V$,

$$\begin{aligned}\Pr[i \in S \text{ and } \forall j \in N(i),\ j \notin S] &= \Pr[i \in S] \cdot (1 - \Pr[\exists j \in N(i),\ j \in S | i \in S]) \\ &\geqslant \Pr[i \in S] \cdot \left(1 - \sum_{j \in N(i)} \Pr[j \in S | i \in S]\right).\end{aligned}$$

where the second inequality follows by union bound.

We now analyze $\Pr[j \in S | i \in S]$. Since $\langle u_i, u_j \rangle \leqslant -1 + \varepsilon$, we can write $u_j = \alpha u_i + \beta w$, where $w \perp u_i$, $-1 \leqslant \alpha \leqslant -1 + \varepsilon$ and $\beta = \sqrt{1 - \alpha^2} \leqslant \sqrt{2\varepsilon}$. Then, $j \in S$ means that $\langle g, u_j \rangle = \alpha \langle g, u_i \rangle + \beta \langle g, w \rangle \geqslant t$, and combined with $\langle g, u_i \rangle \geqslant t$, we have $\langle g, w \rangle \geqslant (1 - \alpha) t / \beta \geqslant t / \beta$. Thus,

$$\Pr[j \in S | i \in S] \leqslant \Pr[\langle g, w \rangle \geqslant t/\beta] \leqslant e^{-t^2/2\beta^2} \leqslant 1/n^2,$$

since $\beta \leqslant \sqrt{2\varepsilon}$ and $t = 4\sqrt{\varepsilon \log n}$. As i has at most n neighbors, this implies that $\Pr[i \in S \text{ and } \forall j \in N(i),\ j \notin S] \geqslant \Pr[i \in S] \cdot (1 - o(1))$. In particular, we have $\mathbb{E}|T| \geqslant n \cdot \Pr[i \in S] \cdot (1 - o(1)) \geqslant n^{1-O(\varepsilon)}$, completing the proof. □

We now prove Theorem 9.5.1.

Proof of Theorem 9.5.1. Consider the following independent set formulation:

$$\begin{aligned} \max \quad & \sum_{i \in V} x_i \\ \text{s.t.} \quad & (1 + x_i)(1 + x_j) = 0 \quad \forall (i, j) \in E(G), \\ & x_i^2 = 1 \quad \forall i \in V(G). \end{aligned} \tag{9.6}$$

Note that any vector $x \in \{\pm 1\}^n$ where $\{i : x_i = 1\}$ is an independent set in G is a feasible solution to the above, since $(1 + x_i)(1 + x_j)$ is nonzero only if $x_i = x_j = 1$.

Since G has an independent set of size $(1/2-\varepsilon)n$, the above program has value at least $(1/2-\varepsilon)n-(1/2+\varepsilon)n=-2\varepsilon n$.

We can solve the SDP relaxation of (9.6) and obtain a pseudo-distribution μ, and we have that $\sum_{i\in V}\widetilde{\mathbb{E}}_\mu[x_i]\geqslant -2\varepsilon n$. Let $S\coloneqq\{i:\widetilde{\mathbb{E}}_\mu[x_i]\geqslant -4\varepsilon\}$. Then,

$$-2\varepsilon n\leqslant \sum_{i\in S}\widetilde{\mathbb{E}}_\mu[x_i]+\sum_{i\notin S}\widetilde{\mathbb{E}}_\mu[x_i]\leqslant |S|+(n-|S|)\cdot(-4\varepsilon)\implies |S|\geqslant \frac{2\varepsilon n}{1+4\varepsilon}\geqslant \varepsilon n\,.$$

For any $i\sim j\in S$, we have $\widetilde{\mathbb{E}}_\mu[x_ix_j]=-1-\widetilde{\mathbb{E}}_\mu[x_i+x_j]\leqslant -1+8\varepsilon$. Moreover, each vertex i is associated with a unit vector u_i such that $\langle u_i,u_j\rangle=\widetilde{\mathbb{E}}_\mu[x_ix_j]$. Thus, the subgraph $G[S]$ and the unit vectors satisfy the conditions in Lemma 9.5.2. Thus, there is a polynomial-time algorithm that finds an independent set in $G[S]$ of size at least $|S|^{1-O(\varepsilon)}\geqslant(\varepsilon n)^{1-O(\varepsilon)}$. □

Part III

Explicit Constructions of Vertex Expanders

Chapter 10

Introduction

In this part, we study the problem of constructing explicit *vertex expanders*. Vertex expansion refers to the property that every "small enough" set of vertices should have "many" distinct neighbors. In this thesis, we restrict our attention to bipartite graphs. Below, we first define bipartite vertex expanders formally.

Definition 10.0.1 (Two-sided vertex expander)**.** We say a (d_L, d_R)-biregular graph Z with left and right vertex sets L and R respectively is a γ-*two-sided vertex expander* if there is a constant $\delta > 0$ depending on γ, d_L, d_R such that:

1. Every subset $S \subseteq L$ with $|S| < \delta|L|$ has at least $\gamma \cdot d_L|S|$ neighbors in R.
2. Every subset $S \subseteq R$ with $|S| < \delta|R|$ has at least $\gamma \cdot d_R|S|$ neighbors in L.

In addition to vertex expansion, a well-studied notion of expansion is *unique-neighbor expansion*. A unique-neighbor of a set S is a vertex v with exactly one edge to S. We can similarly define bipartite unique-neighbor expanders.

Definition 10.0.2 (Two-sided unique-neighbor expander)**.** We say a (d_L, d_R)-biregular graph Z with left and right vertex sets L and R respectively is a γ-*two-sided unique-neighbor expander* if there is a constant $\delta > 0$ depending on γ, d_L, d_R such that:

1. Every subset $S \subseteq L$ with $|S| < \delta|L|$ has at least $\gamma \cdot d_L|S|$ unique-neighbors in R.
2. Every subset $S \subseteq R$ with $|S| < \delta|R|$ has at least $\gamma \cdot d_R|S|$ unique-neighbors in L.

Observe that $(1-\varepsilon)$-vertex expansion implies $(1-2\varepsilon)$-unique-neighbor expansion. We will call an infinite family of graphs *lossless* (vertex) expanders if $\gamma = 1 - \varepsilon$, where ε can be made arbitrarily small (for large enough degrees d_L, d_R). Thus, lossless expanders are graphs achieving the quantitatively strongest form of unique-neighbor expansion.

More generally, γ-vertex expansion implies $(2\gamma - 1)$-unique-neighbor expansion, and this is tight. In particular, $\frac{1}{2}$-vertex expansion, which is guaranteed by spectral expansion, does not guarantee unique-neighbor expansion [Kah95] (see Section 10.1 for more discussion).

The main result of this part is an explicit construction of constant-degree lossless expanders, resolving a longstanding open problem (see, e.g., [HLW06, Open problem 10.8]). Before stating our results, we first provide some background on prior work and applications of vertex expanders in Section 10.1. Then, in Section 10.2, we state our main result and discuss the sequence of works [HMMP24, HLMOZ25, HLMRZ25] that led to this construction.

10.1 History of vertex expanders

The quest for explicit lossless vertex expanders can be traced back to the seminal work of Sipser and Spielman [SS96] who identified vertex expansion as an important property for error correction. In particular, they showed that a *one-sided* lossless expander can be used to construct a good error-correcting code with a linear-time decoding algorithm. Around the same time, a parallel line of work on distributed routing in networks [Pip93, ALM96, BFSU98] identified vertex expansion as a crucial property of networks for designing routing protocols. At the time, it was well understood that a random graph is a lossless vertex expander with optimal parameters with high probability, but no explicit constructions were known.

The quest for explicit constructions I. The first work in the direction of obtaining explicit constructions was by Kahale [Kah95], who proved that any d-regular Ramanujan graph is a $(1/2 - o(1))$-vertex expander (see Theorem 3.2.1). Unfortunately, this barely fell short of being useful for applications, which needed small sets to have many *unique-neighbors*. In the same work, Kahale proved that $1/2$ was an inherent barrier to spectral techniques by constructing a near-Ramanujan graph along with a small subset S of vertices with only $d/2 \cdot |S|$ neighbors, and more strikingly, with *zero* unique-neighbors (see [MM21, KK22, KY24] for similar examples of such graphs).

The first explicit construction of unique-neighbor expanders was given by Alon and Capalbo [AC02]. Shortly after, in a breakthrough work, Capalbo, Reingold, Vadhan, and Wigderson [CRVW02] gave explicit constructions of one-sided lossless expanders.

Applications. We refer the reader to [CRVW02] for a detailed treatment of known applications of lossless expanders at the time in coding theory, distributed routing, fault tolerant networks, storage schemes, and proof complexity.

Ever since, the array of applications has expanded: [DSW06, BV09] proved that one can use codes arising from unique-neighbor expanders to construct *robustly testable codes*, and Viderman [Vid13] gave a linear-time decoding algorithm for codes constructed from $\frac{2}{3}$-vertex expanders. Vertex expanders have also seen applications in high-dimensional geometry: the works of [GLR10, Kar11, BGIKS08, GMM22] used unique-neighbor expanders to construct ℓ_p-*spread subspaces* and matrices satisfying the ℓ_p-*isometry property*. The work [HMP06] gave a construction of a family of deterministic and uniform circuits

for computing the (approximate) majority of n bits assuming the construction of fully lossless expanders. Motivated by randomness extractors, the works [TUZ07, GUV09] gave constructions of polynomially imbalanced one-sided lossless expanders.

More recently, in the wake of advances on constructing c^3-locally testable codes [DELLM22, PK22] and quantum LDPC codes [PK22], Lin and M. Hsieh gave alternate simpler constructions of both these objects: c^3-LTCs in [LH22a] based on one-sided lossless expanders, and quantum LDPC codes in [LH22b] based on two-sided lossless expanders with a free group action, which we now have (see Remark 10.2.3).

The quest for explicit constructions II. The work of Lin and M. Hsieh [LH22b] renewed interest in constructing vertex expanders, which led to a flurry of new work. Asherov and Dinur [AD24] gave a simple construction of one-sided unique-neighbor expanders, based on generalizing a construction in [AC02], which was simplified in a work of Kopparty, Ron-Zewi, and Saraf [KRS23]. Golowich [Gol24] and independently, Cohen, Roth and Ta-Shma [CRT23] proved that their construction instantiated with appropriate parameters in fact yields one-sided lossless expanders.

Using significantly different ideas, a recent work of [CGRZ24] studied the bipartite graphs of [KT22], which have polynomially large imbalance, and showed that they have two-sided lossless expansion — the first construction of two-sided lossless expanders in the unbalanced setting. The setting of polynomial imbalance is of interest in the literature on randomness extractors, , but are not known to give good quantum LDPC codes via [LH22b]. Here, we focus on bipartite graphs with constant degrees and constant imbalance.

10.2 Explicit lossless vertex expanders

In a sequence of works [HMMP24, HLMOZ25, HLMRZ25], culminating in [HLMRZ25], we give the first construction of constant-degree lossless expanders.

Theorem 10.2.1 (Constant-degree lossless expanders)**.** *For every $\varepsilon > 0$, there exists a sufficiently large integer $d_0 = d_0(\varepsilon)$ such that for every integer $d \geqslant d_0$, there is an explicit (deterministic polynomial-time constructible) infinite family of d-regular graphs G that are $(1-\varepsilon)$-vertex expanders.*

More generally, we construct *two-sided* lossless expanders with arbitrary constant *imbalance* $\beta = d_L/d_R$. Having arbitrary imbalance is important in many applications; for example, in coding theory, the imbalance determines the rate of the code.

Theorem 10.2.2. *For every $\varepsilon, \beta \in (0,1]$, there exist $k = k(\varepsilon), d_0 = d_0(\varepsilon,\beta) \in \mathbb{N}$ such that for any $d_L, d_R \geqslant d_0$ for which $\beta \leqslant d_L/d_R \leqslant \beta + \varepsilon$, there is an infinite family of (kd_L, kd_R)-biregular bipartite graphs $(Z_n)_{n\geqslant 1}$ for which Z_n is a two-sided $(1-\varepsilon)$-vertex expander on $\Theta(n)$*

vertices. Additionally, there is an algorithm that takes in a positive integer n as input, and in $\mathrm{poly}(n)$*-time outputs* Z_n*.*

We remark that Theorem 10.2.2 implies Theorem 10.2.1: In the special case where $d_L = d_R = d$, the construction can be made d-regular for *any* $d \geqslant d_0(\varepsilon)$. The trick is to begin with a $(\widetilde{d}, \widetilde{d})$-biregular graph G guaranteed by Theorem 10.2.2, where $\widetilde{d} \in \left[d, \left(1 + \frac{1}{k-1}\right)d\right]$. Since G is bipartite, it can be decomposed into $\widetilde{d}$ edge-disjoint perfect matchings. By taking the union of any d of these matchings, we obtain a d-regular subgraph. Such a d-regular subgraph can be seen to incur only a negligible loss in the vertex expansion.

Remark 10.2.3 (Free group action)**.** Our construction also admits a *free group action* by a group of size linear in the number of vertices in the graph, resolving a conjecture of Lin and M. Hsieh [LH22b, Conjecture 10]. By their work, our construction yields a new family of good quantum LDPC codes, which also admit a linear-time decoding algorithm. See Section 13.5 for details.

Tripartite line product. In all three papers [HMMP24, HLMOZ25, HLMRZ25], the main ingredient is the *tripartite line product*, which was first introduced in [HMMP24] and will be explained in detail in Chapter 11. It has 2 ingredients: a large base graph and a constant-size gadget graph. The gadget graph can be viewed as a random graph, which has lossless expansion (see Section 11.1). At a high level, the tripartite line product "lifts" the strong expansion properties of the gadget graph to a large graph, according to the base graph. This falls in the general framework of "local-to-global lifting" (see, e.g., Dinur's recent talk [Din24]).

Once the tripartite line product is defined, the main innovation lies in choosing the large base graphs, as the gadget graph is always assumed to be a random graph.

- In Chapter 12, we instantiate the base graphs using explicit (near-)Ramanujan bipartite graphs. This yields γ-unique-neighbor expanders, where $\gamma > 0$ is a small universal constant. The analysis is simple and only uses the bound on subgraph densities in near-Ramanujan bipartite graphs (Theorem 3.2.3), which was proved in Chapter 5.

  Moreover, using the generalized Moore bound (Theorem 3.1.3; proved in Chapter 4) and the high-girth properties of the explicit Ramanujan graphs by [OW20], we are able to show that small sets in this construction have lossless expansion. This chapter is based on [HMMP24].

- In [HLMOZ25], we use the face-vertex incidence graphs of Ramanujan simplicial complexes (i.e., high-dimensional expanders) by [LSV05a, LSV05b]. This gives $(3/5)$-unique-neighbor expanders, which notably surpass Kahale's spectral barrier of $1/2$ guaranteed by spectral expansion (recall Theorem 3.2.1). We do not

include this construction in the thesis, since many of the ideas are shared with those in the subsequent paper.

- In Chapter 13, we use the face-vertex incidence graphs of expanding *cubical complexes* based on LPS Ramanujan graphs [LPS88], which were constructed (in a different form) in [RSV19]. This achieves the final goal of constructing lossless expanders. This chapter is based on [HLMRZ25].

Chapter 11

Tripartite Line Product

Our construction of lossless expanders is based on the *tripartite line product* framework, which is a generalization of the *line product* introduced in [AC02].

Definition 11.0.1 (Tripartite line product)**.** Given the ingredients:

- Two bipartite *base graphs*, a (k, D_L)-biregular graph $G_L = (L, M, E_L)$, and a (k, D_R)-biregular graph $G_R = (R, M, E_R)$, along with injective functions $\mathrm{LNbr}_u : [D_L] \to L$ and $\mathrm{RNbr}_u : [D_R] \to R$ for every vertex $u \in M$ that index the left and right neighbors of u,
- A (d_L, d_R)-biregular *gadget graph* H where the left-hand side is $[D_L]$, and the right-hand side is $[D_R]$.

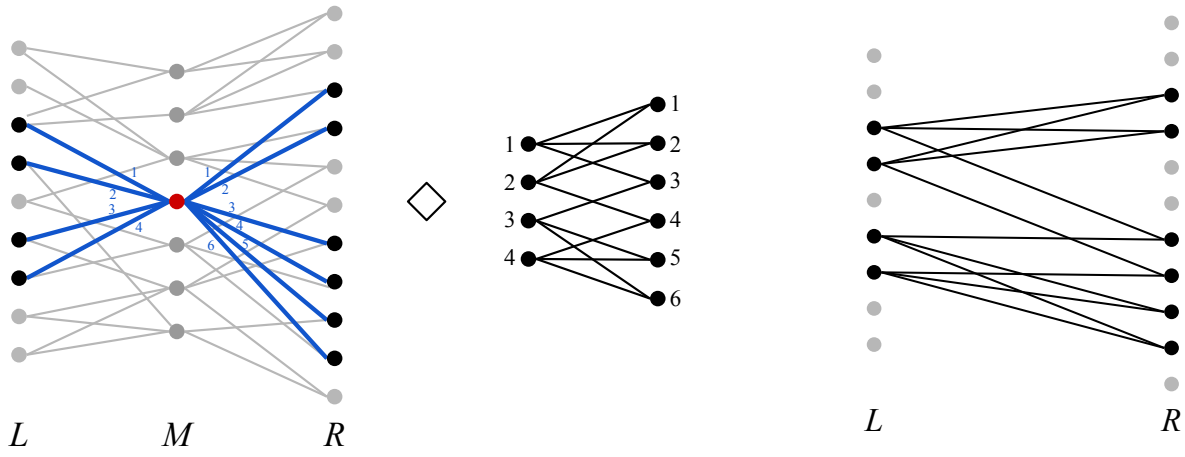
We define the *tripartite line product* of (G_L, G_R) and H as the (kd_L, kd_R)-biregular graph Z obtained by taking each middle vertex $u \in M$, and placing a copy of H between the left and right neighbors of u. Specifically, for every edge $(i, j) \in H$, we place an edge between $\mathrm{LNbr}_u(i)$ and $\mathrm{RNbr}_u(j)$.

See Figure 11.1 for an example of the tripartite line product.

Since the gadget graph H has constant size (since D_L, D_R are constants), we can find an H that satisfies strong expansion properties by brute force. In particular, we can find a gadget graph with the properties that a *random* graph satisfies (with high probability), such as lossless expansion. Therefore, it is convenient to think of H as a random graph.

On the other hand, the bipartite graphs G_L and G_R of the base graph are carefully chosen explicit, infinite families of bipartite expanders. As mentioned in Chapter 10, our goal is to choose base graphs that allow the tripartite line product to "lift" the expansion properties of the (random) gadget to much larger graphs.

Remark 11.0.2 (Generalizing the line product and routed product)**.** The *line product* and *routed product*, which feature in [AC02, AD24, Gol24], arise from instantiating the tripartite line product with appropriate base graphs. The line product can be obtained by choosing a $(2, D_L)$-biregular graph between L and M, and a $(D_R, 2)$-biregular graph



(a) The base graph G and gadget graph H.

(b) The product $G \diamond H$.

Figure 11.1: The tripartite line product between a base graph G and gadget graph H. In this figure, only the edges from the copy of H placed at the red vertex in M are drawn.

between M and R in the base graph. The routed product arises by choosing a (k, D_L)-biregular graph between L and M, and a $(D_R, 1)$-biregular graph between M and R in the base graph.

Parameters. In our constructions, the parameters k, D_L, D_R, d_L, d_R are all constants (large enough depending on the imbalance β and the lossless expansion parameter ε) compared to the size of the base graphs. However, it is convenient to treat k as fixed while d_L, d_R and $D := D_L + D_R$ grow (as we want constructions for infinitely many degrees), and we will use $o_D(1)$ to denote a quantity that can be made smaller than any constant by making D a large enough constant.

11.1 Gadget graph

The reader should think of the gadget graph as a random graph that satisfies strong expansion properties. Its properties were analyzed in [HMMP24, HLMOZ25], which we articulate in the following statement.

Lemma 11.1.1 ([HLMOZ25, Lemma 2.10]). *Let D_L, D_R, d_L, d_R, k, s be integers such that $D_L \cdot d_L = D_R \cdot d_R$, and $k \leqslant D^{0.1} \leqslant d_L, d_R \leqslant o_D(D)$ where $D := D_L + D_R$. Suppose for any distinct $a, b \in [k]$, there is an $r(a,b) \in \mathbb{N}$ and a partition $(Q_i^{a,b})_{i \in [r(a,b)]}$ of $[D_R]$ where each partition has size within $\left[\frac{D}{2s}, \frac{2D}{s}\right]$. Then, there exists a bipartite graph H on $[D_L] \cup [D_R]$ such that*

- ***(lossless expansion)** for any $A \subseteq [D_L]$ with $|A| \leqslant o_D(1) \cdot D_R/d_L$, we have $|N(A)| \geqslant$*

$(1 - o_D(1))d_L|A|$,

- ***(spread)*** *for any distinct $a, b \in [k]$, for any $A \subseteq [D_L]$ and any $W \subseteq [r(a,b)]$ with $|W| \geqslant \frac{s \log D}{d_L}$,*

$$\sum_{i \in W} |N(A) \cap Q_i| \leqslant 32|W| \cdot \max\left\{\frac{d_L|A|}{s}, \log D\right\}.$$

Additionally, H satisfies the above guarantees when the roles of "L" and "R" are swapped.

The spread condition above can be interpreted as follows: for any $A \subseteq [D_L]$ not too small, it has at most $d_L|A|$ neighbors, and these neighbors are well spread across the r partitions — any $|W|$ partitions contain at most an $O\left(\frac{|W|}{s}\right)$ fraction of them.

In Chapter 12, we will only use the expansion property — in fact, we will need unique-neighbor expansion of "slightly larger" subsets in a random gadget graph (see Lemma 12.2.1). In Chapter 13, we will exploit the high-dimensional structure in the base graphs, in which the spread property in Lemma 11.1.1 is crucial.

11.2 Outline of the analysis

Once the tripartite line product is defined, the main innovation lies in choosing the large base graphs, as the gadget graph is always assumed to be a random graph.

Chapter 12

Unique-Neighbor Expanders with Lossless Small-Set Expansion

In this chapter, we show that instantiating the tripartite line product with

- a base graph consisting of two explicit near-Ramanujan bipartite graphs, and
- a "random" gadget graph,

gives γ-unique-neighbor expanders, where $\gamma > 0$ is a universal constant.

Moreover, using the generalized Moore bound from Chapter 4 (Theorem 3.1.3), we prove that if the near-Ramanujan bipartite graphs have no *bicycle* (recall Definition 3.1.2) of size at most $g_n = \omega_n(1)$, then small subsets of size $\exp(g_n)$ expand losslessly. Such expanders were constructed in the works of [MOP20, OW20].

Theorem 12.0.1 (Special case of main theorem in [OW20])**.** *For every $c, d \geqslant 3$ and $\gamma > 0$, there is an explicit construction of an infinite family of (c,d)-biregular graphs $(G_n)_{n\geqslant 1}$ where $\lambda_2(G_n) \leqslant (\sqrt{c-1} + \sqrt{d-1}) \cdot (1+\gamma)$, and G_n has no bicycle on $o(\sqrt{\log |V(G_n)|})$ vertices.*

This gives the following result.

Theorem 12.0.2. *For every $\varepsilon, \beta \in (0,1]$, there are constants $\gamma > 0$ and $k, d_0 \in \mathbb{N}$ such that for any $d_L, d_R \geqslant d_0$ for which $\beta \leqslant d_L/d_R \leqslant \beta + \varepsilon$, there is an infinite family of (kd_L, kd_R)-biregular bipartite graphs $(Z_n)_{n\geqslant 1}$ such that:*

1. *Z_n is a two-sided γ-unique-neighbor expander,*
2. *every $S \subseteq L(Z_n)$ with $|S| \leqslant \exp(\Omega(\sqrt{\log |V(Z_n)|}))$ has $(1-\varepsilon) \cdot kd_L \cdot |S|$ unique-neighbors,*
3. *every $S \subseteq R(Z_n)$ with $|S| \leqslant \exp(\Omega(\sqrt{\log |V(Z_n)|}))$ has $(1-\varepsilon) \cdot kd_R \cdot |S|$ unique-neighbors.*

While Theorem 12.0.2 is ultimately subsumed by our construction of lossless expanders in Chapter 13, the goal of this chapter is to highlight the effectiveness and

flexibility of the tripartite line product. Notably, using well-known constructions of Ramanujan graphs [LPS88, Mor94, MOP20, OW20] — as opposed to the more involved base graphs via cubical complexes used in Chapter 13 — the product already yields non-trivial guarantees in unique-neighbor expansion.

Organization. In Section 12.1, we prove lossless expansion of small sets in graphs with no short cycles or bicycles (Lemma 12.1.1), generalizing a result of [Kah95]. Then, in Section 12.2, we prove Theorem 12.0.2.

12.1 Lossless expansion in high-girth graphs

We first restate the generalized Moore bound for convenience.

Theorem (Restatement of Theorem 3.1.3)**.** *Suppose G is a graph on n vertices, and let $\rho = \rho(B_G)$ be the spectral radius of its non-backtracking matrix B_G. Suppose $\rho > 1$, then G contains a cycle of size at most $2(\lfloor \log_\rho n \rfloor + 1)$ and a bicycle of size at most $3(\lfloor \log_\rho 2n \rfloor + 1)$.*

Finally, we will need the following statement about the expansion of small sets in graphs with no short cycles or bicycles, which generalizes [Kah95, Theorem 10].

Lemma 12.1.1 (Expansion of small sets)**.** *Let $G = (L \cup R, E)$ be a d-left-regular bipartite graph, and let $\varepsilon \in (0,1)$ such that $\varepsilon(d-1) > 1$. Suppose G has no cycle of length at most g, then for all $S \subseteq L$ with $|S| \leqslant \frac{1}{d+1}(\varepsilon(d-1))^{\frac{1}{4}g-\frac{1}{2}}$ we have $|N_G(S)| \geqslant (1-\varepsilon)d|S|$.*

Similarly, suppose G has no bicycle of length at most g, then for all $S \subseteq L$ with $|S| \leqslant \frac{1}{2(d+1)}(\varepsilon(d-1))^{\frac{1}{6}g-\frac{1}{2}}$ we have $|N_G(S)| \geqslant (1-\varepsilon)d|S|$.

Proof. Let $T := N_G(S) \subseteq R$. Suppose S does not expand losslessly, i.e., $|T| < (1-\varepsilon)d|S|$. Then, the subgraph $G[S \cup T]$ must have right average degree at least $\frac{d|S|}{(1-\varepsilon)d|S|} \geqslant \frac{1}{1-\varepsilon} \geqslant 1+\varepsilon$. Let $\rho > 0$ be the spectral radius of the non-backtracking matrix $B_{G[S\cup T]}$ so that $H_{G[S\cup T]}(1/\rho) \succeq 0$. Then, applying Lemma 3.2.4, we have

$$1+\varepsilon \leqslant 1 + \frac{\rho^2}{(d-1)} \implies \rho \geqslant \sqrt{\varepsilon(d-1)}\,.$$

Next, by Theorem 3.1.3, $G[S \cup T]$ must contain a cycle of size at most

$$2\log_\rho(|S|+|T|) + 2 \leqslant \frac{2\log((d+1)|S|)}{\log\sqrt{\varepsilon(d-1)}} + 2\,.$$

Suppose $|S| \leqslant \frac{1}{d+1}(\varepsilon(d-1))^{\frac{1}{4}(g-2)}$, then there exists a cycle of length at most g, which is a contradiction.

Similarly, by Theorem 3.1.3, $G[S \cup T]$ must contain a bicycle of size at most

$$3\log_\rho(2(|S|+|T|))+3 \leqslant \frac{3\log(2(d+1)|S|)}{\log\sqrt{\varepsilon(d-1)}}+3\,.$$

Suppose $|S| \leqslant \frac{1}{2(d+1)}(\varepsilon(d-1))^{\frac{1}{6}(g-3)}$, then there exists a bicycle of length at most g, which is a contradiction. □

We remark that in a follow-up work, Chen [Che25] gave a different proof using combinatorial and subsampling arguments, with slightly improved parameters compared to Lemma 12.1.1.

12.2 Proof of Theorem 12.0.2

From Lemma 11.1.1, we know that the random gadget H has lossless expansion for small sets. Here, we need that slightly larger sets have unique-neighbor expansion.

Lemma 12.2.1 (Lemma 4.3 of [HMMP24]). *Let $\beta \in (0,1/2]$, $\theta > 0$, and $\tau > 0$ be constants. For integers d_1, d_2, D_1, D_2 and $D := D_1 + D_2$ satisfying*

1. $\frac{d_1}{d_2} = \frac{D_2}{D_1}$,
2. $1 \geqslant \frac{d_1}{d_2} \geqslant \frac{\beta}{1-\beta}$,
3. $\theta\sqrt{D}/2 \leqslant d_1 + d_2 \leqslant \theta\sqrt{D}$,

there is a (d_1, d_2)-biregular graph H with D_1 vertices on the left and D_2 vertices on the right such that:

$$\min_{S\subseteq V(H):1\leqslant|S|\leqslant t} \frac{|\mathrm{UN}_H(S)|}{|S|} \geqslant (1-o_D(1))\cdot d_1 \cdot \exp(-\theta t/\sqrt{D})$$

for $1 \leqslant t \leqslant \tau\sqrt{D}$ where $o_D(1)$ hides constant factors depending only on β, θ and τ.

We now prove Theorem 12.0.2 using the tripartite line product, where the base graph consists of two near-Ramanujan bipartite graphs from Theorem 12.0.1. In the analysis, we will use the bound on subgraph densities in Ramanujan bipartite graphs, which we restate below.

Theorem (Restatement of Theorem 3.2.3). *Let $3 \leqslant c \leqslant d$ be integers, $\gamma \in [0,1]$, and $\varepsilon \in (0, 0.1)$. Let $G = (L \cup R, E)$ be a (c,d)-biregular graph such that $\lambda_2 \leqslant (\sqrt{c-1}+\sqrt{d-1})(1+\gamma/d)$. Then, there exists $\delta = \delta(\varepsilon, c, d) > 0$ such that for every $S_1 \subseteq L$ and $S_2 \subseteq R$ with $|S_1| + |S_2| \leqslant \delta|L \cup R|$, the left and right average degrees $d_1 = \frac{|E(S_1,S_2)|}{|S_1|}$ and $d_2 = \frac{|E(S_1,S_2)|}{|S_2|}$ in the induced subgraph $G[S_1 \cup S_2]$ must satisfy*

$$(d_1-1)(d_2-1) \leqslant \sqrt{(c-1)(d-1)}\cdot(1+O(\varepsilon+\sqrt{\gamma}))\,.$$

Proof of Theorem 12.0.2. The construction of Z_n is based on taking the tripartite line product of some base graph G_n and a bipartite gadget graph H from Lemma 11.1.1 with suitably chosen parameters.

- Let $K = K_1 = K_2 = \frac{1000}{\varepsilon^2}$.
- Let $d_0 \geqslant C_0 K$ where $C_0 = C_0(\varepsilon, \beta)$ is some large enough constant chosen later.
- Let $\widetilde{d_1} := d_1/K$ and $\widetilde{d_2} := d_2/K$, both at least C_0.
- Let $\theta := \frac{C}{\varepsilon}\sqrt{K/\beta}$ (depending only on ε, β) where C is a universal constant.
- Let $D_1 := \left\lceil \frac{\widetilde{d_1}+\widetilde{d_2}}{\theta^2} \right\rceil \cdot \widetilde{d_2}$ and $D_2 := \left\lceil \frac{\widetilde{d_1}+\widetilde{d_2}}{\theta^2} \right\rceil \cdot \widetilde{d_1}$, and define $D := D_1 + D_2$.

Note that $1 \geqslant \widetilde{d_1}/\widetilde{d_2} = d_1/d_2 \geqslant \frac{\beta}{1-\beta}$. One can verify that $KD_1 \leqslant 2\varepsilon^2 \widetilde{d_2}^2/C^2$ and $KD_2 \leqslant 2\varepsilon^2 \widetilde{d_1}^2/C^2$.

The above choice of parameters satisfy the requirements of Lemma 12.2.1. Thus, applying Lemma 12.2.1 with parameters θ and $\tau = 1$, there is a $(\widetilde{d_1}, \widetilde{d_2})$-biregular graph H with D_1 left vertices and D_2 right vertices such that

$$\min_{S \subseteq V(H): 1 \leqslant |S| \leqslant t} \frac{|\mathrm{UN}_H(S)|}{|S|} \geqslant (1 - o_D(1)) \cdot \widetilde{d_1} \cdot \exp(-\theta t/\sqrt{D}) \tag{12.1}$$

for $1 \leqslant t \leqslant \sqrt{D}$. We can set C_0 large enough (depending only on β, θ which only depend on ε, β) such that the $o_D(1)$ term is at most 0.1.

For the tripartite base graph $G_n = (L \cup M \cup R, E_1 \cup E_2)$, we construct $G_n^{(1)} = (L \cup M, E_1)$ and $G_n^{(2)} = (M \cup R, E_2)$ to be (K_1, D_1) and (D_2, K_2)-biregular near-Ramanujan graphs from Theorem 12.0.1 respectively, i.e., $\lambda_2(G_n^{(1)}) \leqslant (\sqrt{K_1 - 1} + \sqrt{D_1 - 1})(1 + 0.01/D_1)$ and $\lambda_2(G_n^{(2)}) \leqslant (\sqrt{K_2 - 1} + \sqrt{D_2 - 1})(1 + 0.01/D_2)$, along with the guarantee that no small bicycles exist.

Unique-neighbor expansion. Next, we analyze the vertex expansion of a subset $S \subseteq L(Z_n)$ in the product graph Z_n. Recall that $L(Z_n) = L$. Let $U := N_{G_n}(S) \subseteq M$ be the neighbors of S in $G_n^{(1)}$, and we partition U into $U_\ell := \{v \in U : |E_1(v, S)| \leqslant \sqrt{D}\}$ (the "low S-degree" vertices) and $U_h := U \setminus U_\ell$ (the "high S-degree" vertices). Consider the bipartite subgraph induced by $S \cup U_h$. By definition, the average right-degree in $G_n^{(1)}[S \cup U_h]$ is at least $\sqrt{D}$. By the upper bound on $\lambda_2(G_n^{(1)})$, we can apply Theorem 3.2.3 and bound the average left-degree by

$$d_{\text{left}}(S, U_h) \leqslant 1 + \frac{\sqrt{(K_1 - 1)(D_1 - 1)}}{\sqrt{D} - 1} \cdot 1.1 \leqslant 1 + 1.2\sqrt{K},$$

as long as $|S| \leqslant \mu |L|$ for some $\mu = \mu(K, D_1) > 0$ (depending only on $\varepsilon, \beta, d_1, d_2$). For any $K \geqslant 100$, the above is at most $0.2K$. Thus, we know that $|E_1(S, U_\ell)| \geqslant 0.8K|S|$, i.e., a constant fraction of edges incident to S go to U_ℓ. This also implies that $|U_h| \leqslant 0.2|U|$.

For each $v \in U$, let $S_v \subseteq S$ be the vertices in S incident to v. Consider the gadget H placed on v, and let $T_v \subseteq R$ be the set of unique-neighbors of S_v in the gadget. Further, let $\widetilde{T} := \bigcup_{v \in U} T_v$. Note that each vertex in $\widetilde{T}$ is a unique-neighbor *within some gadget*, but there may be edges coming from other gadgets, so not all of $\widetilde{T}$ are unique-neighbors of S in the final product graph. Our goal is to show that a large fraction of $\widetilde{T}$ are unique-neighbors of S.

We will analyze the induced subgraph $G_n^{(2)}[U \cup \widetilde{T}]$, and we claim that a large fraction of $\widetilde{T}$ are unique-neighbors of U in $G_n^{(2)}$, thus are also unique-neighbors of S in Z_n. We first lower bound the left average degree of $G_n^{(2)}[U \cup \widetilde{T}]$. For each $v \in U_\ell$, we have $1 \leqslant |S_v| \leqslant \sqrt{D}$, and by the expansion profile of the gadget (Eq. (12.1)), v has degree at least

$$|S_v| \cdot 0.9 \cdot \widetilde{d_1} \cdot \exp(-\theta |S_v|/\sqrt{D}) \geqslant 0.9 \cdot \widetilde{d_1} \cdot \min\left\{e^{-\theta/\sqrt{D}}, \sqrt{D} e^{-\theta}\right\}$$

in $G_n^{(2)}[U \cup \widetilde{T}]$. Since θ depends only on ε, β, we choose $C_0 = C_0(\varepsilon, \beta)$ to be large enough (thus also D) such that the above is at least $0.8 \cdot \widetilde{d_1}$.

Next, for $v \in U_h$, we have no control over its degree in $G_n^{(2)}[U \cup \widetilde{T}]$. However, since $|U_h| \leqslant 0.2|U| \leqslant \frac{1}{4}|U_\ell|$, we have

$$d_{\text{left}}(U, \widetilde{T}) \geqslant \frac{0.8\widetilde{d_1} \cdot |U_\ell|}{|U_\ell| + |U_h|} \geqslant 0.64 \cdot \widetilde{d_1}\,.$$

Then, for $|S| \leqslant \mu |L|$ where μ is small enough, we have $|U| \leqslant \mu' |M|$ where μ' (depending on $\varepsilon, \beta, K, D_2$) is small enough to apply Theorem 3.2.3 and conclude that the right average degree

$$d_{\text{right}}(U, \widetilde{T}) \leqslant 1 + \frac{\sqrt{(K_2 - 1)(D_2 - 1)}}{0.64 \cdot \widetilde{d_1} - 1} \cdot O(1) \leqslant 1.1\,,$$

since $\widetilde{d_1} + \widetilde{d_2} \leqslant \frac{1}{\beta}\widetilde{d_1}$ and $K_2 D_2 \leqslant \widetilde{d_1}^2 / C$ with some large C by our choice of θ and D_2. This implies that 0.9 fraction of $\widetilde{T}$ are unique-neighbors of S.

Finally, we lower bound $|E_2(U, \widetilde{T})|$. Again by Eq. (12.1),

$$\begin{aligned} |E_2(U, \widetilde{T})| &\geqslant \sum_{v \in U_\ell} |S_v| \cdot 0.9 \cdot \widetilde{d_1} \cdot \exp(-\theta |S_v|/\sqrt{D}) \geqslant 0.9 \cdot \widetilde{d_1} \cdot \exp(-\theta) \sum_{v \in U_\ell} |S_v| \\ &= 2\delta \cdot \widetilde{d_1} \cdot |E_1(S, U_\ell)| \geqslant 1.6\delta d_1 |S|\,, \end{aligned}$$

where $\delta = \delta(\varepsilon, \beta) > 0$. The last inequality uses the fact that $|E_1(S, U_\ell)| \geqslant 0.8K|S|$ and $d_1 = K\widetilde{d_1}$. With $d_{\text{right}}(U, \widetilde{T}) \leqslant 1.1$, it follows that

$$|\text{UN}_{Z_n}(S)| \geqslant \delta d_1 |S|\,.$$

For $S \subseteq R(Z_n)$, the analysis is completely symmetric. Indeed, we have $K_1 = K_2$, and we can verify that $K_1 D_1 \leqslant \widetilde{d_2}^2/C$. Since $d_1 \leqslant d_2$, the unique-neighbor lower bound holds for all $S \subseteq V(Z_n)$ with $|S| \leqslant \mu |V(Z_n)|$.

Small set lossless expansion. We now turn to the expansion of small subsets $S \subseteq L(Z_n)$. Let $U := N_{G_n^{(1)}}(S) \subseteq M$ and $T := N_{Z_n}(S) \subseteq R$. By assumption, $G_n^{(1)}$ has no bicycle of size at most g_n, thus Lemma 12.1.1 states that $|U| \geqslant (1-\varepsilon/2)K|S|$ (i.e., S expands losslessly in $G_n^{(1)}$) assuming that

$$|S| \leqslant \frac{1}{2(K+1)} \left(\frac{\varepsilon}{2}(K-1)\right)^{\frac{1}{6}g_n - \frac{1}{2}} .$$

With our choice of K and $g_n = \omega_n(1)$, it suffices that $|S| \leqslant \exp(g_n)$.

Next, as each gadget on $v \in U$ expands with a factor of at least $\widetilde{d_1}$, we can lower bound $|E_2(U,T)|$ by $\widetilde{d_1} \cdot |U|$. Moreover, the left average degree of the induced subgraph $G_n^{(2)}[U \cup T]$ is at least $\widetilde{d_1}$. Then, by Theorem 3.2.3, the right average degree of $G_n^{(2)}[U \cup T]$ is

$$d_{\text{right}}(U,T) \leqslant 1 + \frac{\sqrt{(K_2-1)(D_2-1)}}{\widetilde{d_1}-1} \cdot O(1) \leqslant 1 + \frac{\varepsilon}{2},$$

given that $K_2 D_2 \leqslant \varepsilon^2 \widetilde{d_1}^2/C$ for a large enough constant C. Thus, since $|E_2(U,T)| \geqslant \widetilde{d_1} \cdot |U| \geqslant (1-\varepsilon/2) \cdot K\widetilde{d_1} \cdot |S| = (1-\varepsilon/2)d_1|S|$,

$$|T| \geqslant \frac{1}{1+\varepsilon/2}|E_2(U,T)| \geqslant \frac{1-\varepsilon/2}{1+\varepsilon/2} \cdot d_1|S| \geqslant (1-\varepsilon)d_1|S| .$$

For $S \subseteq R(Z_n)$, the analysis is symmetric with $d_1, \widetilde{d_1}$ replaced by $d_2, \widetilde{d_2}$. □

Chapter 13

Explicit Lossless Vertex Expanders

In this chapter, we present our construction of constant-degree lossless expanders with arbitrary constant imbalance.

Theorem (Restatement of Theorem 10.2.2)**.** *For every $\varepsilon, \beta \in (0, 1]$, there exist $k = k(\varepsilon), d_0 = d_0(\varepsilon, \beta) \in \mathbb{N}$ such that for any $d_L, d_R \geqslant d_0$ for which $\beta \leqslant d_L/d_R \leqslant \beta + \varepsilon$, there is an infinite family of (kd_L, kd_R)-biregular bipartite graphs $(Z_n)_{n\geqslant 1}$ for which Z_n is a two-sided $(1-\varepsilon)$-vertex expander on $\Theta(n)$ vertices. Additionally, there is an algorithm that takes in a positive integer n as input, and in $\mathrm{poly}(n)$-time outputs Z_n.*

As in Chapter 12, our construction uses the tripartite line product (Chapter 11). For the base graphs (i.e., the two bipartite graphs), we replace the Ramanujan graphs used in Chapter 12 with graphs that have additional high-dimensional structure. In [HLMOZ25], we use Ramanujan high-dimensional expanders of [LSV05a, LSV05b], which yields $(3/5)$-vertex expanders. To get lossless expanders, we use Ramanujan *cubical complexes*. It turns out that cubical complexes give us extra symmetry that Ramanujan complexes lack.

Organization. We begin with a technical overview in Section 13.1, which includes an overview of cubical complexes, the associated incidence graphs, and an overview of the analysis. In Section 13.2, we introduce the notion of *structured bipartite graph* (Definition 13.2.1 and Lemma 13.2.4) and show that, when used as the base graph in the tripartite line product, they yield lossless expanders.

Next, in Section 13.3, we formally define expanding cubical complexes (Definition 13.3.4). Moreover, we show that the *coded incidence* graphs (Definition 13.3.7) of such complexes give structured bipartite graphs with the desired parameters. A crucial component of the proof is an upper bound on *small-set subcube density* in expanding cubical complexes (Lemma 13.3.11), which is proved in Section 13.3.2.

In Section 13.4, we describe the construction of expanding cubical complexes based on the LPS Ramanujan graphs [LPS88]. This section provides an exposition of LPS

graphs and self-contained proofs of the properties we need, while we will only use the fact that they are Ramanujan as a black box.

Finally, as noted in Remark 10.2.3, our construction admits a free group action, which leads to new families of quantum LDPC codes by [LH22b] with linear time decoding algorithms. The details are given in Section 13.5.

13.1 Technical overview

Our construction is obtained as the tripartite line product (Chapter 11) of bipartite graphs derived from *Ramanujan cubical complexes* with a constant-sized gadget graph, which can be thought of as a random graph. In Section 13.1.1, we first give an overview of expanding cubical complexes, deferring technical details and the construction to Sections 13.3 and 13.4. Next, in Section 13.1.2, we describe the *coded incidence* graphs associated with such complexes, which we will use as the bipartite base graphs in the tripartite line product. Finally, in Section 13.1.3, we give an overview of the analysis.

13.1.1 Cubical complexes

Our construction of lossless expanders relies on expanding cubical complexes. Here, we give a brief overview; see Section 13.3 for more definitions and properties, and Section 13.4 for an explicit construction using LPS Ramanujan graphs [LPS88].

The theory of expanding cubical complexes was first studied by Jordan and Livné [JL00] as a high-dimensional generalization of Ramanujan graphs, where it was shown that infinite families of such complexes exist but no explicit construction was given. Later, explicit constructions were presented in [RSV19] (in a slightly different form), where more general cases were also treated. Recently, cubical complexes were used in [DLV24] to construct quantum locally testable codes, and they instantiated the complexes using abelian lifts of expanders [JMOPT22].

Earlier, a 2-dimensional version of the cubical complexes, dubbed *left-right Cayley complexes*, was an important ingredient in the constructions of locally testable codes with constant rate, distance and locality, as well as good quantum LDPC codes by [DELLM22, PK22]. For our purposes, we will need higher-dimensional cubical complexes with constant degree and good expansion; notably, these can only be constructed over non-abelian groups.

Cayley cubical complex. A k-dimensional cubical complex[1] can be constructed from a finite group Γ and generating sets $A_1, A_2, \ldots, A_k \subseteq \Gamma$ that satisfy

(1) $A_i \cdot A_j = A_j \cdot A_i$ for all $i \neq j$, and

[1]One can define cubical complexes from any set Γ and sets of permutations of Γ. For simplicity, we restrict to Cayley cubical complexes.

(2) $|A_1 \cdots A_k| = |A_1| \cdots |A_k|$.

Here, we denote $A \cdot B = \{ab : a \in A,\ b \in B\}$. We call any collection of sets $A_1, \ldots, A_k$ satisfying the above *cubical generating sets*. Note that we require $A_1, \ldots, A_k$ to commute as sets while the elements do not necessarily commute. In particular, for any $a_1 \in A_1$ and $a_2 \in A_2$, there exist unique $b_1 \in A_1$ and $b_2 \in A_2$ such that $a_1 a_2 = b_2 b_1$. More generally, for any $\{a_i \in A_i\}_{i \in [k]}$ and any permutation $\pi \in S_k$, there exist unique $\{b_i \in A_i\}_{i \in [k]}$ such that $a_1 a_2 \cdots a_k = b_{\pi(1)} b_{\pi(2)} \cdots b_{\pi(k)}$.

Given a group Γ and cubical generating sets $A_1, \ldots, A_k \subseteq \Gamma$, the *decorated*[2] *cubical complex*, denoted $X = \mathrm{Cay}(\Gamma; (A_1, \ldots, A_k))$, is the complex with vertex set $X(0) = \Gamma \times \{0,1\}^k$, edges of the form $\{(g, x), (g a_i, x \oplus e_i)\}$ where $g \in \Gamma$ and $a_i \in A_i$, and k-faces (or cubes) $X(k)$ of the form $f = \{(f_x, x)\}_{x \in \{0,1\}^k}$ where $f_x^{-1} f_{x \oplus e_i} \in A_i$ for each $i \in [k]$ and $x \in \{0,1\}^k$. It is easy to verify that the requirements of cubical generating sets imply that each cube is uniquely specified by a group element $g \in \Gamma$ and $\{a_i \in A_i\}_{i \in [k]}$. See Definition 13.3.2 for a formal definition and Figure 13.1 for an illustration.

We note that it is straightforward to construct cubical complexes using abelian groups since all elements commute. However, we need the complex to exhibit strong expansion, and it is well known that constant-degree abelian Cayley graphs cannot be expanders [AR94].

We construct cubical complexes based on the LPS Ramanujan graphs [LPS88]. Section 13.4 contains an exposition and self-contained proofs of the properties we need. Here, we briefly recall that given primes $p, q \equiv 1 \pmod 4$, the LPS graphs $X(p;q)$ are Cayley graphs over $\Gamma = \mathrm{PSL}(2, \mathbb{F}_q)$ with $p+1$ generators $A(p)$. The Ramanujan cubical complex we construct is simply $\mathrm{Cay}(\Gamma; A(p_1), A(p_2), \ldots, A(p_k))$ for distinct primes $p_1, \ldots, p_k$. It is a remarkable fact that $A(p_1), \ldots, A(p_k)$ indeed form cubical generating sets as defined above (Lemma 13.4.8). Moreover, since each Cayley graph $\mathrm{Cay}(\Gamma; A(p_i))$ is Ramanujan (a fact that we will only use as a black box), the resulting Ramanujan cubical complexes also inherit strong expansion properties.

Remark 13.1.1. By substituting the (arguably more elementary) cubical complex from [DLV24, Section 3.5.2]—derived from abelian lifts of $\Theta(\log n)$-sized Ramanujan Cayley graphs [JMOPT22]—into our construction, one obtains constant-degree n-vertex graphs in which every subset of size $O(n/\mathrm{polylog}\, n)$ has lossless vertex expansion, and which supports a free group action by a $\Theta(n/\mathrm{polylog}\, n)$-sized group.

13.1.2 Coded incidence graphs

We now describe the base graph (i.e., the two bipartite graphs G_L, G_R) used in the tripartite line product. These graphs are constructed using a k-dimensional Ramanujan

[2] We use the word "decorated" since the vertex set $X(0)$ comprises 2^k copies of Γ, unlike traditional Cayley graphs that have only one copy of Γ.

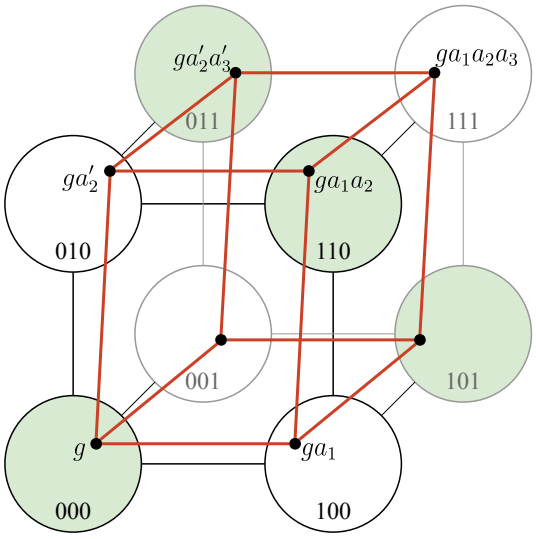


Figure 13.1: A 3-dimensional (decorated) cubical complex $X = \mathrm{Cay}(\Gamma; (A_1, A_2, A_3))$, where the vertex set $X(0) = \Gamma \times \mathbb{F}_2^3$. An element $g \in \Gamma$ and $a_1 \in A_1$, $a_2 \in A_2$, $a_3 \in A_3$ uniquely specify a face (or cube) $f \in X(3)$, as depicted in the figure. Note that by the properties of A_1, A_2, A_3, there exist unique $a_1' \in A_1$, $a_2' \in A_2$ and $a_3' \in A_3$ such that $a_1a_2a_3 = a_2'a_3'a_1'$.

The vertex-face incidence graph we need for our base graph construction will be restricted to a linear code $\mathcal{C} \subseteq \mathbb{F}_2^k$ of large distance — the bipartite graph between $X(k)$ and $\Gamma \times \mathcal{C} \subseteq X(0)$ where edges indicate containment. Here, a code $\{000, 011, 110, 101\}$ is highlighted.

cubical complex X and the Hadamard code $\mathcal{C} \subseteq \{0,1\}^k$ (with $|\mathcal{C}| = k = 2^r$ for some $r \in \mathbb{N}$). We set $L = X(k)$, the k-faces of X, and $M = \Gamma \times \mathcal{C}$, a subset of vertices $X(0)$ according to the code $\mathcal{C}$. A k-face $f \in L$ and a vertex $(g, x) \in M$ are connected in G_L if and only if $(g, x) \in f$. Thus, each $f \in L$ has degree $|\mathcal{C}| = k$, and each vertex in M has degree $D_L = \prod_{i=1}^{k} |A_i|$. The other bipartite graph G_R is defined the same way.

Remark 13.1.2. Restricting the vertices according to the Hadamard code $\mathcal{C}$ provides crucial symmetry in our construction. In particular, for two vertices (g, x) and (h, y) with $x \neq y \in \mathcal{C}$, their common neighborhood (i.e., the set of k-faces containing them) is either empty or all possible completions to a full cube. Since $\mathrm{dist}(x, y) = k/2$ for all $x \neq y \in \mathcal{C}$,[3] the common neighborhoods are all roughly the same structure (by choosing $|A_1|, \ldots, |A_k|$ to be a constant factor away from each other). We believe that this is one key improvement over [HLMOZ25] which is based on Ramanujan simplicial complexes, where M is also k-partite but the common neighborhoods (a.k.a. links) of two

[3]We expect that any δ-balanced linear code with a small enough constant δ will work as well; see Remark 13.3.9.

vertices differ drastically depending on which parts they are in.

13.1.3 Overview of the analysis

Our analysis follows the same outline as [HMMP24, HLMOZ25]. To bound the expansion of a set $S \subseteq L$ (sets on the right follow the same analysis), we split into two parts: the *left-to-middle* and the *middle-to-right* analysis. Fix a (small) subset $S \subseteq L$, and consider the neighbors $U = N_{G_L}(S) \subseteq M$. For each $u \in U$, as long as $\deg_S(u) := |S \cap N_{G_L}(u)|$ is sufficiently small, we will have lossless expansion within the gadget placed on u (since the gadget is random-like). On the other hand, if $\deg_S(u)$ is too large, then the gadget cannot experience lossless expansion because the number of right vertices in the gadget is much smaller than the number of edges in the gadget arising from $N_{G_L}(u)$. Thus, we split U into U_ℓ (low-degree) and U_h (high-degree), and we need to show that most elements of S partake in many U_ℓ gadgets and few U_h gadgets: precisely, we need to show that $e_{G_L}(S, U_h)$ is small such that $1 - \varepsilon$ fraction of edges from S go to U_ℓ.

Left-to-middle analysis: small-set subcube density. We bound the *small-set subcube density* of the cubical complex, similar to the triangle density bound of the Ramanujan simplicial complexes needed in [HLMOZ25]. Our goal is to show that there are not too many k-faces that have many vertices in U_h. More specifically, we upper bound the size of $\{f \in X(k) : |f \cap U_h| \geqslant 2\sqrt{k}\}$ by $O_k(1) \cdot D_L^{5/8}|U_h|$. This is proved in Section 13.3.2 using the structure and expansion of X. More specifically, whereas [HLMOZ25] used spectral properties within the links of the high dimensional expander to obtain their bounds, our complex notably is not a high dimensional expander as the links are disconnected. Instead, we rely on the Hadamard structure of the links along with a variant of the Loomis–Whitney inequality [LW49] to argue that U_h contains few subcubes.

To demonstrate the key ideas, we focus on the simple case of $k = 3$ —- subcube density of 3-dimensional *expanding* cubical complexes with a code $\mathcal{C} = \{000, 011, 110, 101\} \subseteq \mathbb{F}_2^3$, as depicted in Figure 13.1. For any small subset $U \subseteq \Gamma \times \mathcal{C}$, we will show an upper bound on the size of $\{f \in X(3) : |f \cap U| = 4\}$. For simplicity, assume that $|A_1| = |A_2| = |A_3| = p$ (this is true in our construction up to absolute constants), and denote $U_x := U \cap (\Gamma \times \{x\})$ for $x \in \mathcal{C}$.

First, we use the expansion property of the cubical complex. Consider the bipartite graph between $\Gamma \times \{000\}$ and $\Gamma \times \{110\}$, where $(g, 000)$ and $(ga_1a_2, 110)$ are connected for $a_1 \in A_1$ and $a_2 \in A_2$. This bipartite graph has degree $|A_1| \cdot |A_2| = p^2$ and has second eigenvalue $O(p)$, which implies that the subgraph induced by $U_{000} \cup U_{110}$ has average degree $O(p)$. Thus, a typical element $(g, 000) \in U_{000}$ has at most $O(p)$ neighbors in U_{110}, U_{101} and U_{011} respectively.

The next crucial property we use is the fact that any cube f is uniquely identified by any 3 points in $f \cap (\Gamma \times \mathcal{C})$. For example, $(g, 000)$, $(ga_1a_2, 110)$ and $(ga_1a_3, 101)$ uniquely specifies a cube $f \in X(3)$, and in particular, there exist unique $a_2' \in A_2$ and $a_3' \in A_3$

such that $(ga'_2a'_3, 011) \in f$. For simplicity, let us assume that $a'_2 = a_2$ and $a'_3 = a_3$. Then, the key question is:

> For a set of 3-tuples T, suppose $N_{12} = |\{(a_1, a_2) : (a_1, a_2, a_3) \in T \text{ for some } a_3\}|$ and N_{13}, N_{23} defined similarly, how large can T be?

The answer is $|T| \leqslant \sqrt{N_{12}N_{23}N_{13}}$. This is in fact a special case of the *Loomis–Whitney inequality*. Here, we give a simple proof using an entropic argument. For the uniform distribution over T, we have $H(a_1, a_2, a_3) = \log |T|$, while by assumption $H(a_i, a_j) \leqslant \log N_{ij}$ for $i < j$. The well-known Shearer's inequality states that

$$H(a_1, a_2, a_3) \leqslant \frac{1}{2} \sum_{i<j} H(a_i, a_j) \,,$$

which completes the proof.

Our argument for general k follows the same idea. The reason that $2\sqrt{k}$ is relevant is because for any subset $B \subseteq \mathcal{C}$ of a *linear* code $\mathcal{C} \subseteq \mathbb{F}_2^k$ with $|B| \geqslant 2\sqrt{|\mathcal{C}|}$, there exist four distinct elements $\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in B$ such that $\sigma_1 + \sigma_2 + \sigma_3 + \sigma_4 = 0$ (Lemma 13.3.13). This, at a high level, reduces to the 3-dimensional case. We are able to show that $|\{f \in X(k) : |f \cap U| \geqslant 2\sqrt{k}\}| \leqslant O_k(1) \cdot D_L^{5/8}|U|$. Thus, by setting the threshold for U_ℓ and U_h to be larger than $D_L^{5/8}$ and $k = O(1/\varepsilon^2)$, we have that most vertices in $S \subseteq L$ have at least $1 - \frac{2\sqrt{k}}{k} \geqslant 1 - \varepsilon$ fraction of edges going to U_ℓ. This completes the left-to-middle analysis.

Middle-to-right analysis. Having established that most vertices of S participate in many low-degree gadgets, it remains to show that these different gadgets do not have too many collisions in G_R. Our proof of this part closely follows the *middle-to-right analysis* in [HLMOZ25]. In fact, the common neighborhood structure of G_R is the key improvement over Chapter 12 which uses Ramanujan bipartite graphs.

It is convenient to view the expansion of each gadget H_u, for $u \in U$, as "red" edges going from u to vertices in $N_{G_R}(u) \subseteq R$. The neighbors of S in the final product Z are exactly the vertices incident to any red edge. See Figure 13.2 for an example. The red edges form a subgraph of G_R, denoted RED, and we need to show that there are very few collisions on the right.

To this end, we define a *collision* (multi-)graph C on U, where we place an edge $\{u, v\}$ for each $u \neq v \in U$ and $r \in R$ such that $\{u, r\}, \{v, r\} \in \text{RED}$ (see e.g. Figure 13.2b). We need to show an upper bound on $e(C)$. Let $\underline{C}$ be the simple graph obtained by removing duplicated edges from C. Moreover, let $\widetilde{G}_R$ be the simple graph on M where $u \neq v \in M$ are connected if they have a common neighbor in R. Observe that $\underline{C}$ is a subgraph of $\widetilde{G}_R$. Then, the natural idea to bound $e(C)$ is to use the expansion of $\widetilde{G}_R$, which we call *skeleton expansion* (Definition 13.2.3).

If G_R is chosen to be a Ramanujan bipartite graph (as in Chapter 12), then most pairs of vertices in M have few common neighbors, and $\widetilde{G}_R$ has degree $O(D)$ and sec-

ond eigenvalue $O(\sqrt{D})$. In our case, due to the structure of the cubical complexes, every pair of vertices in M has either zero or $\approx \sqrt{D}$ common neighbors, and thus $\widetilde{G}_R$ has degree $O(\sqrt{D})$ and second eigenvalue $O(D^{1/4})$. This is the key improvement over Chapter 12. Of course, now the collision graph C may have large multiplicities, which complicate the analysis. We handle this by using the spreadness of the "random" gadget H (Lemma 11.1.1), and crucially this requires us to place the gadget in the same way for every $u \in M$ (as opposed to arbitrarily). See Section 13.2.2 for more details.

13.2 Construction of lossless vertex expanders

13.2.1 Base graph construction

In this section, we describe the precise properties we will need from the bipartite graphs and the gadget graph.

Notation, terminology, and parameters. Given a graph G and $S, T \subseteq V(G)$, we use $G[S]$ to refer to the induced subgraph of G on S, and $G[S, T]$ as the induced bipartite subgraph of G between S and T. Given a bipartite graph (U, V, E), we denote an edge between a vertex $u \in U$ and $v \in V$ by the ordered tuple (u, v).

In our construction, the parameters k, D_L, D_R, d_L, d_R are all constants (large enough depending on ε, β) compared to the size of the base graphs. However, it is convenient to treat $k \approx \varepsilon^{-2}$ as fixed while d_L, d_R and $D := D_L + D_R$ grow (as we want constructions for infinitely many degrees), and we will use $o_D(1)$ to denote a quantity that can be made smaller than any constant by making D a large enough constant.

Base graph construction. We introduce the notion of a *structured bipartite graph*.

Definition 13.2.1 (Structured bipartite graph)**.** A (k, D)-biregular bipartite graph G between vertex sets V and M is a *structured bipartite graph* if:

(1) For each vertex $u \in M$, there is an injective function $\mathrm{Nbr}_u : [D] \to V$ that specifies an ordering of the D neighbors of u.

(2) The set M can be expressed as a disjoint union $\sqcup_{a \in [k]} M_a$ such that each $v \in V$ has exactly one neighbor in each M_a.

(3) There is an $s \in \mathbb{N}$ such that the following holds: for each pair of distinct $a, b \in [k]$, there are $r(a, b)$ *special sets* $\{Q_i^{a,b} \subseteq [D]\}_{i \in [r(a,b)]}$ that partition $[D]$ (abbreviated to r and Q_i), each $|Q_i| \in [\frac{D}{2s}, \frac{2D}{s}]$, such that for every $u \in M_a$, there are distinct $v_1, \ldots, v_r \in M_b$ with $N(u) \cap N(v_i) = \mathrm{Nbr}_u(Q_i)$ for each $i \in [r]$ and $N(u) \cap N(v') = \varnothing$ for all other $v' \in M$.

Intuitively, Item (3) of Definition 13.2.1 means that for every $u \in M_a$, there are $r(a, b)$ vertices in M_b that have common neighbors with u, and the common neighborhoods

form a specific structure. See Figure 13.2a for an illustration. For our construction, it is important that this structure is the same across all $u \in M_a$ — the special sets $\{Q_i \subseteq [D]\}$ are independent of u (but can depend on $a, b \in [k]$).

Henceforth, we fix G as a structured (k, D)-biregular graph between V and M.

Definition 13.2.2 (Small-set j-neighbor expansion)**.** We say G is a *τ-small-set j-neighbor expander* if for some small constant $\eta > 0$, and for every $U \subseteq M$ such that $|U| \leqslant \eta|M|$, the number of vertices in V with at least j neighbors in U is bounded by $\tau \cdot |U|$.

Definition 13.2.3 (Small-set skeleton expansion)**.** Let $\widetilde{G}$ be the simple graph on M obtained by placing an edge between $u, u' \in M$ if there exists a length-2 path between u and u'. We say G is a *λ-small-set skeleton expander* if for some small constant $\eta > 0$, and for every $U \subseteq M$ such that $|U| \leqslant \eta|M|$, the largest eigenvalue of the adjacency matrix of the graph $\widetilde{G}[U]$ is at most λ.

We now state the guarantees we can achieve in a structured bipartite graph, which we prove in Section 13.3.2.

Lemma 13.2.4. *For every k that is a power of 2, and large enough $D \in \mathbb{N}$, there is an algorithm that takes in $n, D_L, D_R \in \mathbb{N}$ as input where $D_L, D_R \leqslant D$, and constructs vertex sets L, M, R such that $|M| = \Theta(n)$ and $|R| = |L| \cdot D_L/D_R$ along with structured bipartite graphs G_L on (L, M), G_R on (R, M), where G_L is (k, D_L)-biregular and G_R is (k, D_R)-biregular, with the following properties:*

- *$s = \Theta(\sqrt{D})$ for the special set structure.*
- *G_L and G_R are $O\big(D^{5/8}\big)$-small-set $2\sqrt{k}$-neighbor expanders.*
- *G_L and G_R are $O\big(D^{1/4}\big)$-small-set skeleton expanders.*

13.2.2 Proof of Theorem 10.2.2

We are now ready to use the above ingredients to prove Theorem 10.2.2 on the explicit construction of 2-sided lossless vertex expanders. Given ε, d_L and d_R, we choose parameters $D, D_L, D_R, k \in \mathbb{N}$ and $\delta \in (0, 1)$ such that the following relations hold.

- $D_L \cdot d_L = D_R \cdot d_R$.
- $D = D_L + D_R$.
- $k \geqslant 16/\varepsilon^2$ and is a power of 2.
- $D^{-1/16} \leqslant \delta \leqslant o_D(1) \cdot \dfrac{1}{k^2}$.
- $\dfrac{D^{1/4} \log^2 D}{\delta} \leqslant d_L, d_R \leqslant \dfrac{\delta D^{3/8}}{\log D}$.

Here, we assume $d_L, d_R \geqslant d_0(\varepsilon, \beta)$ for a large enough $d_0(\varepsilon, \beta)$ such that any $o_D(1)$ term

is sufficiently small.

Let $G_L = (L, M, E_L)$ and $G_R = (R, M, E_R)$ be the structured bipartite graphs constructed from the algorithm in Lemma 13.2.4 with parameters k, D, n, D_L, D_R. Recall that G_L and G_R are structured bipartite graphs with $s = \Theta(\sqrt{D})$ for the special set structure and are $O(D^{5/8})$-small-set $2\sqrt{k}$-neighbor expanders, and $O(D^{1/4})$-small-set skeleton expanders. In this proof, we will use $\tau = O(D^{5/8})$ to denote the small-set $2\sqrt{k}$-neighbor expansion, and $\lambda = O(D^{1/4})$ to denote the small-set skeleton expansion.

Let H be a (d_L, d_R)-biregular bipartite graph on $[D_L] \cup [D_R]$ whose special subsets of $[D_R]$ are identical to the special subsets associated to G_R, and whose special subsets of $[D_L]$ are identical to the special subsets associated to G_L.

Looking ahead, we will need that

- $\tau \leqslant o_D(\delta) \cdot \frac{D_R}{d_L}$ and similarly $\tau \leqslant o_D(\delta) \cdot \frac{D_L}{d_R}$.
- $\lambda \leqslant s\delta$,
- $d_L, d_R \geqslant \frac{1}{\delta} \max\{\lambda, \sqrt{s}\} \log D$.

One can verify that with parameters $\tau = O(D^{5/8})$, $\lambda = O(D^{1/4})$ and $s = \Theta(\sqrt{D})$ from Lemma 13.2.4, our choice for δ and D_L, D_R listed above satisfy all requirements.

We output the tripartite line product $Z = (L, R, E_Z)$ of (G_L, G_R) with H. We will establish vertex expansion of small subsets of L; the analysis of the vertex expansion of small subsets of R is similar.

Left-to-middle analysis. Let $S \subseteq L$ such that $|S| \leqslant \eta|L|$. Let $U \subseteq M$ be the neighbors of S in G_L. We split U into its "high-degree" part $U_h := \left\{v \in U : \deg_{G_L[S,U]}(v) \geqslant \frac{\tau}{\delta}\right\}$, and "low-degree" part $U_\ell := U \setminus U_h$.

Our first step is to prove that most edges from S to U point to U_ℓ.

Claim 13.2.5. *The number of edges in $G_L[S, U]$ incident to U_ℓ is at least $\left(1 - \sqrt{\delta} - 2k^{-1/2}\right) \cdot k|S|$.*

Proof. By definition, the number of edges incident to U_h in $G_L[S, U]$ is at least $\frac{\tau}{\delta}|U_h|$. On the other hand, denoting $S_{\geqslant 2\sqrt{k}}$ to be the set of vertices in S with at least $2\sqrt{k}$ neighbors in U_h, by small-set $2\sqrt{k}$-neighbor expansion of G_L, we have $|S_{\geqslant 2\sqrt{k}}| \leqslant \tau|U_h|$. Consequently, the number of edges from $S_{\geqslant 2\sqrt{k}}$ into U_h satisfies:

$$e\left(S_{\geqslant 2\sqrt{k}}, U_h\right) \leqslant k\left|S_{\geqslant 2\sqrt{k}}\right| \leqslant k\tau|U_h| = k\delta \cdot \frac{\tau}{\delta}|U_h| \leqslant k\delta \cdot e(S, U_h) \leqslant \sqrt{\delta} \cdot k|S|\,.$$

Here, we use $k \leqslant 1/\sqrt{\delta}$. Thus, we have:

$$\begin{aligned} e(S, U_\ell) &= e(S, U) - e(S, U_h) \\ &= k|S| - e\left(S_{\geqslant 2\sqrt{k}}, U_h\right) - e\left(S_{<2\sqrt{k}}, U_h\right) \\ &\geqslant k|S| - \sqrt{\delta} \cdot k|S| - 2\sqrt{k}|S| \\ &= \left(1 - \sqrt{\delta} - \frac{2}{\sqrt{k}}\right) \cdot k|S| \,. \end{aligned}$$

□

Middle-to-right analysis. We have proved that most edges from S to U touch low-degree vertices, which the reader should think of as gadgets through which the expansion into R is lossless. We make this formal below.

Definition 13.2.6. For $S \subseteq L$ and $U = N_{G_L}(S) \subseteq M$, if a vertex $v \in R$ is a neighbor of S in the final product due to connections from the gadget H_u for $u \in U$, then we color the edge (u, v) red. The red edges form a subgraph of G_R, which we denote as $\mathsf{RED}(S)$ or simply RED when S is clear from context. Figure 13.2a contains an example of the subgraph RED.[4]

By the choice of the threshold, we have $\frac{\tau}{\delta} \leqslant o_D(1) \cdot D_R/d_L$, and hence, by Lemma 11.1.1, each vertex in U_ℓ expands by at least a $(1 - o_D(1))d_L$ factor. In particular, we have,

$$e(\mathsf{RED}) \geqslant \sum_{u \in U_\ell} (1 - o_D(1))d_L \cdot \deg_S(u) = (1 - o_D(1))d_L \cdot e_{G_L}(S, U_\ell) \,. \tag{13.1}$$

In the remainder of the argument, we prove that the collisions between neighborhoods of different gadgets inflict negligible damage on expansion.

We next show that the red edges have few collisions in R. We will crucially use the small-set skeleton expansion with $\lambda = O(D^{1/4})$ and the special set structure of G_R with $s = \Theta(\sqrt{D})$ (Definition 13.2.1 and Lemma 13.2.4).

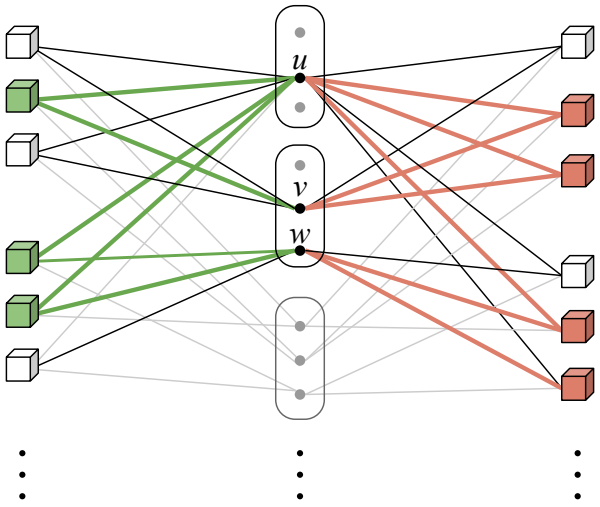
We construct the *collision graph* C — the multi-graph C on vertex set $U \subseteq M$ by placing a copy of the edge $\{u, v\}$ for each $u \neq v \in U$, and $r \in R$ such that $\{u, r\}$ and $\{v, r\}$ are red edges in RED. See Figure 13.2 for an example. The number of neighbors of S in the final product Z is at least

$$e(\mathsf{RED}) - e(C) \,,$$

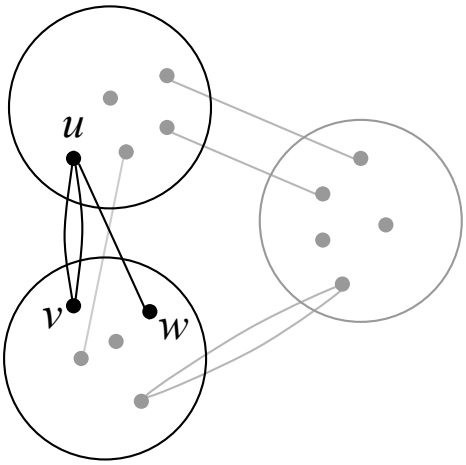
since a vertex $v \in R$ with degree d_v in RED contributes one neighbor, but it is counted d_v times in $e(\mathsf{RED})$ and $\binom{d_v}{2}$ times in $e(C)$, and $d_v - \binom{d_v}{2} \leqslant 1$ for all $d_v \in \mathbb{N}$.

We will need the following folklore fact.

[4]We note that in [HLMOZ25], they need to define "blue" and "red" edges to prove *unique-neighbor* expansion. In our case, since we will show lossless expansion, we do not need to make this distinction.



(a) Let $S \subseteq L$ consist of the cubes colored green, and the cubes on the right incident to red edges are the neighbors of S in the final product Z.



(b) The collision multi-graph C on M. Removing parallel edges gives the simple graph $\underline{C}$, which is a subgraph of $\widetilde{G}_R$.

Figure 13.2: The two bipartite base graphs G_L, G_R have the structure that M has k parts, and for $u \in M$ and $v, w \in M$ from a different part, the common neighborhoods $N_{G_R}(u) \cap N_{G_R}(v)$ and $N_{G_R}(u) \cap N_{G_R}(w) \subseteq R$ are disjoint, each corresponding to a *special set* in $[D_R]$, i.e., $N_{G_R}(u) \cap N_{G_R}(v) = \mathrm{Nbr}_u(Q_i)$ for some special set $Q_i \subseteq [D_R]$.

Figure 13.2a shows an example of $\mathrm{RED}(S)$, a subgraph of G_R. The middle-to-right analysis involves upper bounding the collisions of the red edges on the right. Here, u has collisions with v and w, represented as edges in the collision graph C in Figure 13.2b. We will show that this cannot happen too often by upper bounding $e(C)$.

Lemma 13.2.7 ([HLMOZ25, Lemma 2.17]). *Given a graph G whose adjacency matrix has maximum eigenvalue λ, then there is an orientation of the edges in G such that all vertices have out-degree at most λ.*

Claim 13.2.8. *Suppose $k\delta^2 \leqslant o_D(1)$, $\lambda \leqslant s\delta$, and $d_L \geqslant \frac{1}{\delta}\max\{\lambda, \sqrt{s}\}\log D$. Then, $e(C) \leqslant o_D(1) \cdot kd_L|S|$.*

Proof. Let $\underline{C}$ be the simple graph obtained by removing duplicate edges from C. Moreover, let $\widetilde{G}_R$ be the simple graph on M where $u \neq v \in M$ are connected if they have a common neighbor in R in the graph G_R. Clearly, $\underline{C}$ is a subgraph of $\widetilde{G}_R$. Moreover, recall from Definition 13.2.1 that M is a union of k vertex sets, and thus $\widetilde{G}_R$ is k-partite. Let us now restrict C to edges between two parts $a, b \in [k]$. We will write $r = r(a, b)$ and the special sets $Q_i = Q_i^{a,b}$ for simplicity.

By the λ-small set skeleton expansion, we have that $\underline{C}$ has largest eigenvalue at most λ. This intuitively means that $\underline{C}$ contains very few edges. Next, we need to upper

bound the multiplicities of edges in C. The main observation is that if $u \in M_a$ and $v \in M_b$ are connected in $\widetilde{G}_R$, then u, v in fact have many common neighbors in G_R. More specifically, u has neighbors $v_1, v_2, \ldots, v_r$ in $\widetilde{G}_R$, and each common neighborhood $N_{G_R}(u) \cap N_{G_R}(v_i) \subseteq R$ corresponds to a special set as in Definition 13.2.1. On the other hand, the pseudorandomness of the gadget H implies that the red edges coming out of u must be evenly spread among the special sets. In the following, we make this intuition formal.

The largest eigenvalue of $\underline{C}$ is at most λ. Thus, by Lemma 13.2.7, there is an orientation of the edges of $\underline{C}$ such that all vertices have out-degree at most λ. Pick such an orientation, and let $\mathrm{Out}(u)$ be the set of out-going edges incident to u. Then,

$$e(C) = \sum_{u \in U} \sum_{e \in \mathrm{Out}(u)} \mathrm{multiplicity}(e)\,.$$

Due to the special set structure of G_R (Definition 13.2.1), for any $u \in M_a$ and $v_1, \ldots, v_r$ (potentially) connected in $\underline{C}$, their common neighborhoods within G_R are exactly special sets in the gadget H_u — that is, $N_{G_R}(u) \cap N_{G_R}(v_i) = \mathrm{RNbr}_u(Q_i)$, and each $|Q_i| \in \left[\frac{D_R}{2s}, \frac{2D_R}{s}\right]$ where $s = \Theta(\sqrt{D})$ from Lemma 13.2.4.

Thus, we can upper bound $\sum_{e \in \mathrm{Out}(v)} \mathrm{multiplicity}(e)$ by the number of red edges that land in any $|\mathrm{Out}(v)|$ of the special sets. Denote $\deg_S(v) := \deg_{G_L[S,U]}(v)$. By Lemma 11.1.1, applying the bound with $|W| = \max\left\{|\mathrm{Out}(v)|, \frac{s \log D}{d_L}\right\} \leqslant \max\left\{\lambda, \frac{s \log D}{d_L}\right\}$ and $|A| = \deg_S(v)$, we get

$$\begin{aligned}
\sum_{e \in \mathrm{Out}(v)} \mathrm{multiplicity}(e) &\leqslant O(1) \cdot \max\left\{\lambda,\ \frac{s \log D}{d_L}\right\} \cdot \max\left\{\frac{d_L}{s} \cdot \deg_S(v),\ \log D\right\} \\
&\leqslant O(1) \cdot \max\left\{\frac{\lambda}{s},\ \frac{\lambda \log D}{d_L \deg_S(v)},\ \frac{\log D}{d_L},\ \frac{s \log^2 D}{d_L^2 \deg_S(v)}\right\} \cdot d_L \cdot \deg_S(v) \\
&\leqslant O(\delta) \cdot d_L \cdot \deg_S(v)\,.
\end{aligned}$$

Here, we use the assumptions on the parameters: $\lambda \leqslant \delta s$, and $d_L \geqslant \frac{1}{\delta} \max\{\lambda, \sqrt{s}\} \log D \geqslant \frac{1}{\delta} \log D$.

Summing over $v \in U$, we get

$$e(C) \leqslant O(\delta) \cdot d_L \sum_{v \in U} \deg_S(v) \leqslant O(\delta) \cdot k d_L |S|\,.$$

The above is restricted to one pair $a, b \in [k]$. For the final bound, we multiply the above by k^2. Since $k^2\delta \leqslant o_D(1)$, we get $e(C) \leqslant o_D(1) \cdot k d_L |S|$. □

Finally, we combine the above to finish the proof of Theorem 10.2.2. With $\delta \leqslant o_D(1) \cdot \frac{1}{k^2}$ and $k \geqslant 16/\varepsilon^2$, Claim 13.2.5 and Eq. (13.1) imply that

$$e(\mathsf{RED}) \geqslant (1 - o_D(1)) \cdot d_L \cdot \left(1 - \sqrt{\delta} - 2k^{-1/2}\right) k|S| \geqslant (1 - \varepsilon/2) k d_L |S|\,.$$

The number of neighbors of S in the final product Z is at least $e(\mathsf{RED}) - e(C)$, and by Claim 13.2.8 we have $e(C) \leqslant o_D(1) \cdot kd_L|S|$. Thus, choosing D large enough,

$$|N_Z(S)| \geqslant (1-\varepsilon)kd_L|S|\,.$$

The analysis for the expansion of any $T \subseteq R$ is identical. This finishes the proof.

13.3 Cubical complexes and coded incidence graphs

Notation and terminology. Given subsets A, B of a group Γ with multiplication operation $\cdot$, we define $A \cdot B$ to refer to the product set $\{a \cdot b : a \in A, b \in B\}$.

We start with the definition of cubical generating sets.

Definition 13.3.1 (Cubical generating set)**.** Let Γ be a finite group and $k \in \mathbb{N}$. We say $A_1, A_2, \ldots, A_k \subseteq \Gamma$ are *cubical generating sets* if they are closed under inverses, and

- $A_i \cdot A_j = A_j \cdot A_i$ for all $i \neq j$,
- $|A_1 \cdots A_k| = |A_1| \cdots |A_k|$.

Definition 13.3.2 (Decorated Cayley cubical complex)**.** Given a finite group Γ and cubical generating sets $\mathcal{A} = (A_1, \ldots, A_k)$, the *(decorated) Cayley cubical complex* $X = \mathrm{Cay}(\Gamma; \mathcal{A})$ is defined by:

- its vertex set $X(0) = \Gamma \times \{0,1\}^k$,
- its k-face set $X(k)$ consisting of all 2^k-sized subsets of $X(0)$ of the form $f = \{(f_x, x)\}_{x \in \{0,1\}^k}$ such that for every edge $\{x, x \oplus e_i\}$ of the hypercube, $f_x^{-1} f_{x \oplus e_i} \in A_i$.
- For $I \subseteq [k]$, we define an I-subcube to be all $\{0,1\}^k$ strings of the form $y \oplus \bigoplus_{i \in I} b_i e_i$, where $b_i \in \{0,1\}$ and e_i denotes the vector with a 1 in the i'th index. The dimension of an I-subcube is $|I|$.
- For a subcube C of $\{0,1\}^k$, we define the set of C-faces $X(C)$ as:

$$X(C) := \left\{ \{(f_x, x)\}_{x \in C} : f \in X(k) \right\}.$$

  We define the set of i-faces as $X(i) := \bigcup_{C:\dim(C)=i} X(C)$.

We use the word "decorated" since the vertex set $X(0)$ consists of 2^k copies of Γ, as opposed to the usual way of Cayley graphs on Γ.

Henceforth, we fix a group Γ along with cubical generating sets $A_1, \ldots, A_k$, and let $X = \mathrm{Cay}(\Gamma; (A_1, \ldots, A_k))$.

One important property of cubical complexes is that for any two points $(g, \vec{0})$ and $(g', \vec{1})$ in opposite corners, there is at most one k-face $f \in X(k)$ that contains the two points. More generally, given $U = \{(g_1, \vec{0}), (g_2, x_2), \ldots, (g_m, x_m)\}$, any face restricted to

the subcube of the coordinates $\bigcup_{t>1} \mathrm{supp}(x_t)$ is uniquely identified (if exists). An example is given in Figure 13.1. The points $(g, 000)$ and $(ga_1a_2a_3, 111)$ uniquely identify a 3-face. Moreover, the points $(g, 000)$, $(ga_1a_2, 110)$ and $(ga_1a_3', 101)$ also uniquely identify a 3-face, since $\mathrm{supp}(110) \cup \mathrm{supp}(101) = [3]$.

This property is crucial in our construction, and a more general form is formalized in the following lemma.

Lemma 13.3.3. *For any $U \subseteq X(0)$ where $U = \{(g_1, x_1), \ldots, (g_m, x_m)\}$, define*

$$S(U) = \{i \in [k] : \exists s, t \in [m] \text{ s.t. } x_s[i] \neq x_t[i]\} = \bigcup_{t>1} \mathrm{supp}(x_t \oplus x_1)\,,$$

and subcube

$$C(U) = x_1 \oplus \bigoplus_{i \in S(U)} \{0,1\} \cdot e_i\,.$$

There is at most one $C(U)$-face containing U, and if such an $C(U)$-face exists, the number of k-faces containing U is equal to $\prod_{i \notin S(U)} |A_i|$.

Proof. We will first prove that there is at most one $C(U)$-face containing U, and then prove that if nonzero, the number of k-faces containing U is equal to $\prod_{i \notin S(U)} |A_i|$.

Proof that there is at most one $C(U)$-face containing U. Define $B_r^S(x)$ as the set of all vectors y in $\{0,1\}^k$ such that the Hamming weight of $x \oplus y$ is at most r and $\mathrm{supp}(x \oplus y) \subseteq S$. We will prove for every $r \geqslant 1$ and each $y \in B_r^{S(U)}(x_1)$, there exists an element $g_y \in \Gamma$ such that $f_y = g_y$ for every face f containing U. Indeed, this claim implies that there can be at most one $C(U)$-face containing U.

We start by proving the claim for $r = 1$. Let $y = x_1 \oplus e_i \in B_1^{S(U)}(x_1)$ where $i \in S(U)$. Note that $i \in S(U)$ means that there is a $t \in [m]$ such that $x_1[i] \neq x_t[i]$. We will prove that the points (g_1, x_1) and (g_t, x_t) uniquely identify (f_y, y). Equivalently, any pair of faces f and f' containing U must have $f_y = f_y'$.

Define $a_i = g_1^{-1} f_y$ and $a_i' = g_1^{-1} f_y'$. Note that both a_i and a_i' must be in A_i. Pick an arbitrary order $j_1, \ldots, j_\ell$ for the coordinates in $\mathrm{supp}(x_1 \oplus x_t) \setminus \{i\}$. Next, observe that the sets $E := a_i \cdot A_{j_1} \cdots A_{j_\ell}$ and $E' := a_i' \cdot A_{j_1} \cdots A_{j_\ell}$, which both have size $\left|A_{j_1}\right| \cdots \left|A_{j_\ell}\right|$, must have a nonempty intersection since they both must contain $g_1^{-1} g_t$. Now, $\left|A_i \cdot A_{j_1} \cdots A_{j_\ell}\right| = |A_i| \cdot \left|A_{j_1}\right| \cdots \left|A_{j_\ell}\right|$, and thus if $a_i \neq a_i'$, then E and E' must be disjoint. Therefore, $a_i = a_i'$ and $f_y = f_y'$.

For the inductive step, assume that for some $r \geqslant 2$, the uniqueness statement holds for all $y \in B_{r-1}^{S(U)}(x_1)$. Let f be any face containing U and let $y \in B_r^{S(U)}(x_1)$. We will prove that f_y is uniquely determined. Define $U' := U \cup \left\{(g_x, x) : x \in B_{r-1}^{S(U)}(x_1)\right\}$ where g_x is the unique value of f_x for any face f containing U. Note that $S(U') = S(U)$. Observe that $\mathrm{supp}(y \oplus x_1)$ is nonempty by the assumption that $r \geqslant 2$, and let i be an arbitrary

element contained within. This means that $y \oplus e_i \in B_{r-1}^{S(U)}(x_1)$. Since $S(U') = S(U)$, the conclusion that f_y is uniquely determined follows by applying the statement we established for $r = 1$ to U' in place of U and $y \oplus e_i$ in place of x_1.

On number of ways to extend a $C(U)$-face to a k-face. It remains to prove that the number of ways to extend a $C(U)$-face to a full k-face is equal to $\prod_{i \notin S(U)} |A_i|$. To this end, fix an order $i_1, \ldots, i_\ell$ of coordinates in $\overline{S(U)}$ arbitrarily. For each choice of $(a_i \in A_i)_{i \notin S(U)}$, we will prove that there is a unique k-face f containing $U \cup \left\{ \left(g_1 \cdot a_{i_1} \cdots a_{i_\ell}, x_1 \oplus 1_{\overline{S(U)}} \right) \right\}$. The conclusion will follow from the fact that there are $\prod_{i \notin S(U)} |A_i|$ many choices for $(a_i)_{i \notin S(U)}$.

We will construct this face f by describing f_y for each $y \in \{0,1\}^k$. We will first treat the case of y of the form $x_1 \oplus \Delta$ for Δ supported on coordinates outside $S(U)$. Let $j_1, \ldots, j_s$ be the coordinates in the support of Δ, and let $j'_1, \ldots, j'_{\ell - s}$ be an arbitrary order for coordinates in $\{i_1, \ldots, i_\ell\} \setminus \{j_1, \ldots, j_s\}$. Now, by the property that $A_i \cdot A_j = A_j \cdot A_i$ for every i, j, we have:

$$g_1 \cdot a_{i_1} \cdots a_{i_\ell} = g_1 \cdot a'_{j_1} \cdots a'_{j_s} \cdot a'_{j'_1} \cdots a'_{j'_{\ell - s}}$$

where $a'_j \in A_j$. We define f_y as $g_1 \cdot a'_{j_1} \cdots a'_{j_s}$.

We now construct f_y for general $y \in \{0,1\}^k$. Observe that y can be written as $z \oplus \Delta$ for $z \in C(U)$ and Δ supported only on coordinates outside $S(U)$. Let $j_1, \ldots, j_s$ be the coordinates in the support of Δ, and let $j'_1, \ldots, j'_{s'}$ be the coordinates in the support of $x_1 \oplus z$. Now, we can write:

$$\begin{aligned} f_{x_1 \oplus \Delta} &= g_1 \cdot a'_{j_1} \cdots a'_{j_s} \\ &= g_z \cdot a'_{j'_1} \cdots a'_{j'_{s'}} \cdot a'_{j_1} \cdots a'_{j_s} \\ &= g_z \cdot a''_{j_1} \cdots a''_{j_s} \cdot a''_{j'_1} \cdots a''_{j'_{s'}}, \end{aligned}$$

where $a''_j \in A_j$. In the above, we used the construction of $f_{x_1 \oplus \Delta}$ from earlier in the first equality, the fact that there is a $C(U)$-face containing (g_z, z) and (g_1, x_1) in the second equality, and $A_i \cdot A_j = A_j \cdot A_i$ in the third equality. Finally, we set f_y as $g_z \cdot a''_{j_1} \cdots a''_{j_s}$. It can easily be checked using the set-commuting relation that f is indeed a valid k-face. Finally, f is the unique face containing $\widetilde{U} := U \cup \left\{ \left(g_1 \cdot a_{i_1} \cdots a_{i_\ell}, x_1 \oplus 1_{\overline{S(U)}} \right) \right\}$ since $S(\widetilde{U}) = [k]$, which completes the proof. □

Finally, we define a natural notion of expansion in a cubical complex that is useful for our purposes.

Definition 13.3.4 (Expanding cubical complex)**.** We say that a cubical complex $X = \mathrm{Cay}(\Gamma; (A_1, \ldots, A_k))$ is α-expanding if for any $x, y \in \{0,1\}^k$, the bipartite graph $\mathcal{I}_{y, y \oplus x}$

with edge set $\left\{\left\{(g,y),(g\cdot\prod_{i=1}^k a_i^{x_i}, y\oplus x)\right\} : g\in\Gamma, a_i\in A_i\right\}$, which has degree $d_x(X) = \prod_{i=1}^k |A_i|^{x_i}$, has second eigenvalue at most $\alpha\sqrt{d_x(X)}$. For $i\in[k]$, we define $d_i(X) := \max_{x\in\{0,1\}^k:|\mathrm{supp}(x)|=i} d_x(X)$.

The following theorem is essentially contained in [RSV19] in a different form. We provide a mostly self-contained proof in Section 13.4, assuming only that the expander graphs of Lubotzky–Phillips–Sarnak [LPS88] are Ramanujan.

Theorem 13.3.5. *Let $p_1 < \cdots < p_k$ and $q > 2\sqrt{\prod_{i=1}^k p_i}$ be any prime numbers congruent to 1 mod 4, and each p_i is a quadratic residue modulo q. There is an explicit choice of cubical generating sets $A_1,\dots,A_k$ on $\Gamma = \mathrm{PSL}_2(\mathbb{F}_q)$ such that $|A_i| = p_i+1$ and the cubical complex $X = \mathrm{Cay}(\Gamma;(A_1,\dots,A_k))$ is 2^k-expanding.*

Base graph construction. We will construct our bipartite base graph based on a cubical complex X and a code $\mathcal{C}\subseteq\{0,1\}^k$. To do so, we first introduce the notion of the "signature" of a cube.

Definition 13.3.6 (Signature of cube)**.** Given a k-face $f\in X(k)$, its *signature* is the following labeling of the directed edges of the k-dimensional hypercube with elements of Γ: for every $x\in\{0,1\}^k$ and every $i\in[k]$, we label the directed edge $(x, x\oplus e_i)$ with $f_x^{-1}f_{x\oplus e_i}$.

Definition 13.3.7 (Coded cubical incidence graph)**.** Given a code $\mathcal{C}\subseteq\{0,1\}^k$, the *$\mathcal{C}$-cubical incidence graph* of a cubical complex X is the edge-labeled bipartite graph (V_1,V_2,E) such that $V_1 = X(k)$, $V_2 = \Gamma\times\mathcal{C}\subseteq X(0)$, and $f\in X(k)$ and $(g,x)\in V_2$ are connected iff $(g,x)\in f$. Further, an edge between f and (g,x) is labeled with the signature of f.

Our construction uses the cubical incidence graph arising from the Hadamard code, of which we use minimal properties.

Fact 13.3.8. *Let k be a power of 2. The k-th Hadamard code $\mathcal{H}_k$ is a linear code in $\mathbb{F}_2^k$ of dimension $\log_2 k$ where for all distinct $x,y\in\mathcal{H}_k$, the Hamming distance between x and y is exactly $k/2$.*

Remark 13.3.9. For our purposes, any linear code with dimension growing in k and pairwise distance between $\frac{2}{5}+\delta$ and $\frac{3}{5}-\delta$ would suffice. The rate and distance of the chosen code determine the trade-off between the degree d and the parameter ε in the $(1-\varepsilon)$-vertex expansion. However, we do not optimize this dependence and use the Hadamard code for simplicity.

13.3.1 Proof of Lemma 13.2.4: structured bipartite graph construction

We now construct structured bipartite graphs (Definition 13.2.1) with the parameters specified in Lemma 13.2.4. It is quite straightforward to see that the $\mathcal{C}$-cubical incidence graph of a cubical complex from Theorem 13.3.5 has the desired special set structure and small-set skeleton expansion, while we defer the proof of small-set $2\sqrt{k}$-neighbor expansion to Section 13.3.2. However, since the construction from Theorem 13.3.5 restricts the degrees to be products of primes, we must remove some faces according to their signatures to get the desired degrees D_L, D_R.

We will need the following folklore fact (see, e.g., [HLMOZ25, Lemma 3.13] for a proof).

Lemma 13.3.10. *For any n-vertex d-regular graph G with largest nontrivial eigenvalue λ, and any subgraph H of G incident to at most δn vertices, the largest eigenvalue of H is at most $\lambda + \delta d$.*

Let $p_1, \dots, p_k$ and $p'_1, \dots, p'_k$ be $2k$ distinct primes congruent to 1 mod 4 such that each $D^{1/k} \leqslant p_i \leqslant 2D^{1/k}$, and let q be a prime of the form $1 + 4\ell \prod_{i=1}^k p_i p'_i$ for $\ell \in \mathbb{N}$. These primes exist due to Fact 13.4.11. Let X be the cubical complex given by Theorem 13.3.5 for $p_1, \dots, p_k$ and q, and let X' be the corresponding cubical complex for $p'_1, \dots, p'_k$ and q. Let $\mathcal{C} = \mathcal{H}_k \subseteq \mathbb{F}_2^k$ be the Hadamard code, let $\underline{D}_L := \prod_{i=1}^k (p_i + 1)$, and let $\underline{D}_R := \prod_{i=1}^k (p'_i + 1)$. Finally, let $\underline{G}_L = (\underline{L}, M, E_{\underline{L}})$ and $\underline{G}_R = (\underline{R}, M, E_{\underline{R}})$ be the $\mathcal{C}$-cubical incidence graphs (Definition 13.3.7) of X and X' respectively.

We first prove the desired properties for $\underline{G}_L$ and $\underline{G}_R$, and then show how to construct G_L and G_R from them, which inherit the desired properties and additionally are (k, D_L)-biregular and (k, D_R)-biregular respectively.

Small-set skeleton expansion. Recall that $M = \Gamma \times \mathcal{C}$ has $|\mathcal{C}| = k$ parts, and the skeleton of X (Definition 13.2.3) is the simple graph on M where vertices $(g, x), (h, y) \in M$ are connected if they are contained in some face $f \in X(k)$. Thus, the skeleton of X is the union of bipartite graphs over each pair $x \neq y \in \mathcal{C}$ with edges $\{(g, x), (g \cdot \prod_{i=1}^k a_i^{x_i \oplus y_i}, y)\}$ for $g \in \Gamma$ and $a_i \in A_i$. Since x, y have distance exactly $k/2$, the degree of the bipartite graph is $d_{x \oplus y} = \prod_{i=1}^k |A_i|^{x_i \oplus y_i} = O(\sqrt{D})$. By the fact that X is 2^k-expanding (from Theorem 13.3.5), its second eigenvalue is at most $2^k \sqrt{d_{x \oplus y}} \leqslant O(D^{1/4})$. By Lemma 13.3.10, we get that $\underline{G}_L$ is an $O\big(D^{1/4}\big)$-skeleton expander. The same argument applies for $\underline{G}_R$.

Bound on the number of special sets. For every $x, y \in \mathcal{C}$, along with any signature σ on the subcube given by $C_{x,y} := \{x \oplus z : \mathrm{supp}(z) \subseteq \mathrm{supp}(x \oplus y)\}$, let Q_σ be the set of all signatures τ of the hypercube that extend σ. The number of choices of x, y and signature σ on the subcube is at most $k^2 \cdot \sqrt{D}$. It can be verified that for any pair of vertices u, v, either the neighborhoods are empty, or are described by one of the sets Q_σ.

Small-set $2\sqrt{k}$-neighbor expansion. The precise statement from which our bounds on small-set $2\sqrt{k}$-neighbor expansion follows is given below.

Lemma 13.3.11. *For any subset of vertices $U \subseteq M$ of size at most $D^{-1}|M|$, we have that the number of vertices in $\underline{L}$ and $\underline{R}$ with more than $2\sqrt{k}$ neighbors in U is at most $O(D^{5/8})|U|$.*

We defer the proof of Lemma 13.3.11 to Section 13.3.2, and describe how to construct G_L and G_R.

Satisfying degree constraints. There is a collection $\underline{\mathcal{S}}_L$ of $\underline{D}_L$ distinct signatures τ such that every $m \in M$ is incident to exactly one element of $\underline{L}$ with signature τ in $\underline{G}_L$. Likewise, there is a collection $\underline{\mathcal{S}}_R$ of $\underline{D}_R$ distinct signatures τ such that every $m \in M$ is incident to exactly one element of $\underline{R}$ with signature τ in $\underline{G}_R$.

We pick an arbitrary D_L-sized subcollection $\mathcal{S}_L$ of $\underline{\mathcal{S}}_L$ and an arbitrary D_R-sized subcollection $\mathcal{S}_R$ of $\underline{\mathcal{S}}_R$, and define L and R as:

$$L \coloneqq \{v \in \underline{L} : \text{Signature}(v) \in \mathcal{S}_L\}, \qquad R \coloneqq \{v \in \underline{R} : \text{Signature}(v) \in \mathcal{S}_R\}.$$

We now define G_L and G_R as the induced subgraphs $\underline{G}_L[L, M]$ and $\underline{G}_R[R, M]$ respectively. The graphs G_L and G_R are (k, D_L)- and (k, D_R)-biregular bipartite graphs, respectively, and each inherits the desired small-set skeleton expansion and small-set $2\sqrt{k}$-neighbor expansion properties from its parent graph.

Neighborhood functions. Arbitrarily order the D_L signatures in $\mathcal{S}_L$ as $\ell_1, \ldots, \ell_{D_L}$, and the D_R signatures in $\mathcal{S}_R$ as $r_1, \ldots, r_{D_R}$. For any vertex $u \in M$ and $i \in [D_L]$, the function $\text{LNbr}_u(i)$ maps to the neighbor of u in L with the signature ℓ_i, and similarly for $i \in [D_R]$, $\text{RNbr}_u(i)$ maps to the neighbor of u with signature r_i. □

13.3.2 Small-set subcube density in cubical complexes

In this section, we prove Lemma 13.3.11, which states that for any small enough subset $U \subseteq M = \Gamma \times \mathcal{H}_k$, there are at most $O_k(D^{5/8})|U|$ faces $f \in X(k)$ that contain at least $2\sqrt{k}$ vertices in U. Here, recall that $\mathcal{H}_k \subseteq \mathbb{F}_2^k$ is the k-th Hadamard code of distance $k/2$ (Fact 13.3.8). Thus, the following lemma directly implies Lemma 13.3.11.

Lemma 13.3.12. *Let Γ be a group with cubical generating sets $A_1, \ldots, A_k$ such that $\max_{i\in[k]}|A_i| \leqslant 2 \cdot \min_{i\in[k]}|A_i|$. Let $D \coloneqq \prod_{i\in[k]}|A_i|$, and let $X = \text{Cay}(\Gamma; (A_1, \ldots, A_k))$ be a 2^k-expanding cubical complex with vertex set $X(0) = \Gamma \times \mathbb{F}_2^k$. Then, for any $U \subseteq \Gamma \times \mathcal{H}_k$ where $|U| \leqslant D^{-1}|\Gamma \times \mathcal{H}_k|$, we have:*

$$\left|\left\{f \in X(k) : |f \cap U| \geqslant 2\sqrt{k}\right\}\right| \leqslant O_k\left(D^{5/8}\right) \cdot |U|.$$

Notations. For a vertex $(g, s) \in X(0)$, we say that it has *type* $s \in \mathbb{F}_2^k$. We use $F_k(U; \geqslant 2\sqrt{k})$ to denote the set of k-faces $\left\{f \in X(k) : |f \cap U| \geqslant 2\sqrt{k}\right\}$, which is what we will

bound in Lemma 13.3.12. More generally, for $\sigma \subseteq \mathcal{H}_k$, we define $F_k(U;\sigma)$ to be the set of all k-faces whose vertices with types in σ lie in U, i.e.,

$$F_k(U;\sigma) := \{f \in X(k) : (f_s, s) \in U,\ \forall s \in \sigma\}.$$

When restricted to a subcube $C \subseteq \mathbb{F}_2^k$, we use $F_C(U;\sigma)$ to denote the C-faces in $X(C)$ (recall Definition 13.3.2) whose vertices with types in σ lie in U.

Our first observation is that for any $f \in F_k(U; \geqslant 2\sqrt{k})$, $f \cap U$ must contain four vertices whose types sum to 0.

Lemma 13.3.13. *Let $S \subseteq \mathcal{H}_k$ be of size $\geqslant 2\sqrt{k}$. Then there exists a four-tuple of distinct elements $\sigma \in S^4$ for which $\sigma_1 \oplus \sigma_2 \oplus \sigma_3 \oplus \sigma_4 = 0$.*

Proof. Consider the set of sums of two distinct elements of S. Since there are $\binom{|S|}{2} \geqslant \binom{2\sqrt{k}}{2} > k$ such sums, whereas there are only $|\mathcal{H}_k| = k$ possible values for the sum, there must be two distinct pairs of elements that have the same sum. Namely, there are elements $\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in S$ for which $\sigma_1 + \sigma_2 = \sigma_3 + \sigma_4$. Note that $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ must be pairwise distinct: if for instance $\sigma_1 = \sigma_3$, then $\sigma_2 = \sigma_4$ also, which implies that the pair $\{\sigma_1, \sigma_2\}$ is equal to the pair $\{\sigma_3, \sigma_4\}$. □

We may therefore partition the set $F_k(U; \geqslant 2\sqrt{k})$ according to the value of the four vertex types that sum to 0. In particular, $F_k(U;\sigma)$ is the set of all k-faces that have four vertices of types $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ in U. Then

$$F_k(U; \geqslant 2\sqrt{k}) \subseteq \bigcup_{\sigma:\, \sigma_1 \oplus \sigma_2 \oplus \sigma_3 \oplus \sigma_4 = 0} F_k(U;\sigma),$$

which lets us bound $|F_k(U; \geqslant 2\sqrt{k})|$ by

$$|F_k(U; \geqslant 2\sqrt{k})| \leqslant \sum_{\sigma:\, \sigma_1 \oplus \sigma_2 \oplus \sigma_3 \oplus \sigma_4 = 0} |F_k(U;\sigma)|. \tag{13.2}$$

It therefore suffices to upper bound the size of each $F_k(U;\sigma)$ individually.

To this end, fix $\sigma \in \mathcal{H}_k^4$ for which $\sigma_1 \oplus \sigma_2 \oplus \sigma_3 \oplus \sigma_4 = 0$. The tuple σ determines a subcube

$$C_\sigma = \sigma_1 \oplus \bigoplus_{i \in \Delta(\sigma)} \{0,1\} \cdot e_i, \tag{13.3}$$

where

$$\Delta(\sigma) := \{i \in [k] : \exists j_1, j_2 \in [4] \text{ s.t. } \sigma_{j_1}[i] \neq \sigma_{j_2}[k]\} = \bigcup_{j \in \{2,3,4\}} \operatorname{supp}(\sigma_1 \oplus \sigma_j).$$

Let us establish some properties of $\Delta(\sigma)$.

Claim 13.3.14. *For any $(\sigma_1, \sigma_2, \sigma_3, \sigma_4) \in \mathcal{H}_k^4$ that sum to 0 over $\mathbb{F}_2^k$, there are three disjoint sets $a, b, c \subseteq [k]$, each of size $k/4$, for which*

$$\begin{aligned}\operatorname{supp}(\sigma_1 \oplus \sigma_2) &= a \cup b\\ \operatorname{supp}(\sigma_1 \oplus \sigma_3) &= a \cup c\\ \operatorname{supp}(\sigma_1 \oplus \sigma_4) &= b \cup c\,.\end{aligned}$$

In particular, $\Delta(\sigma) = a \cup b \cup c$ is of size $3k/4$.

Proof. Notice that $\sigma_2' := \sigma_2 \oplus \sigma_1$ and $\sigma_3' := \sigma_3 \oplus \sigma_1$ are distinct codewords of $\mathcal{H}_k$, and hence have weight $k/2$. Furthermore, the distance between σ_2' and σ_3' is also $k/2$. Define $a = \operatorname{supp}(\sigma_2') \cap \operatorname{supp}(\sigma_3')$. Then, the Hamming distance between σ_2' and σ_3', which is $k/2$, can also be written as $(k/2 - |a|) + (k/2 - |a|)$, implying that $|a| = k/4$. We can now define $b = \operatorname{supp}(\sigma_2') \backslash a$ and $c = \operatorname{supp}(\sigma_3') \backslash a$, which will both be of size $k/4$ as well. We simply need to check that $\operatorname{supp}(\sigma_4 \oplus \sigma_1) = b \cup c$, which we do as follows: $\sigma_4 \oplus \sigma_1 = \sigma_2 \oplus \sigma_3 = \sigma_2' \oplus \sigma_3'$ implies that $\operatorname{supp}(\sigma_4 \oplus \sigma_1) = \operatorname{supp}(\sigma_2' \oplus \sigma_3') = b \cup c$. □

For any element $x \in \mathcal{H}_k$, we use U_x to denote $U \cap (\Gamma \times \{x\})$. For a subcube C of $\{0,1\}^k$, recall that $F_C(U;\sigma)$ is all C-faces with a vertex in each U_{σ_i} for $\sigma_i \in \sigma$. By Lemma 13.3.3, each $f' \in F_{C_\sigma}(U;\sigma)$ can be extended to a k-face $f \in F_k(U;\sigma)$ in $\prod_{i \notin \Delta(\sigma)} |A_i|$ ways.

In the remainder of this section, we will use C to refer to C_σ. We can further partition $F_C(U;\sigma)$ based on the value of its type-σ_1 vertex. That is, for $u \in U_{\sigma_1}$, define

$$F_C(u;U;\sigma) := \{f \in F_C(U;\sigma) : u \in f\}\,.$$

We will bound the size of $F_C(u;U;\sigma)$ in the following lemma.

In order to state the bound, we define the s-neighborhood $N_s(u)$ of $u \in U_{\sigma_1}$, for $s \in \mathbb{F}_2^k$, as all the neighbors of u in the bipartite graph $\mathcal{I}_{\sigma_1,s}$ between $\Gamma \times \{\sigma_1\}$ and $\Gamma \times \{s\}$ (recall Definition 13.3.4).

Lemma 13.3.15. *Suppose that $u \in U_{\sigma_1}$ is such that $|N_s(u) \cap U| \leqslant \nu$ for $s \in \{\sigma_2, \sigma_3, \sigma_4\}$. Then*

$$|F_C(u;U;\sigma)| \leqslant \nu^{3/2}\,.$$

Proof. Let a, b, c be the partition of $\Delta(\sigma) \subseteq [k]$ given by Claim 13.3.14. Define $A^{(a)} = \prod_{i \in a} A_i$, $A^{(b)} = \prod_{i \in b} A_i$, and $A^{(c)} = \prod_{i \in c} A_i$. There is a one-to-one correspondence between $N_{\sigma_2}(u)$ and $A^{(a)}A^{(b)} = A^{(b)}A^{(a)}$, $N_{\sigma_3}(u)$ and $A^{(a)}A^{(c)} = A^{(c)}A^{(a)}$, and $N_{\sigma_3}(u)$ and $A^{(b)}A^{(c)} = A^{(c)}A^{(b)}$. For instance, we can view N_{σ_2} as the set of vertices obtained by starting from $u = (g_1, \sigma_1)$, and then multiplying g_1 first by an $A^{(a)}$ element and then an $A^{(b)}$ element to obtain a type-σ_2 vertex.

By Lemma 13.3.3, any C-face containing $u = (g_1, \sigma_1)$ can be uniquely specified by choosing one element each from $A^{(a)}$, $A^{(b)}$, and $A^{(c)}$. Concretely, Lemma 13.3.3 implies

that for $\overline{a} \in A^{(a)}, \overline{b} \in A^{(b)}, \overline{c} \in A^{(c)}$, and $g_2 = g_1\,\overline{a}\,\overline{b}$, $g_3 = g_1\,\overline{a}\,\overline{c}$, there is a unique C-face f containing $(g_1, \sigma_1), (g_2, \sigma_2), (g_3, \sigma_3)$, where for $f_{\sigma_4} = (g_4, \sigma_4)$ we have $g_4 = g_1\,\overline{b}'\,\overline{c}'$ for some $\overline{b}' \in A^{(b)}$ and $\overline{c}' \in A^{(c)}$. Similarly, f is also uniquely determined by the choice of $\overline{a}, \overline{b}'$, and $\overline{c}'$.

Let $H(\cdot)$ be the entropy function, and let $\boldsymbol{f}$ denote the random variable obtained by sampling a uniformly random C-face in $F_C(u; U; \sigma)$, and let $\overline{a}, \overline{b}, \overline{c}, \overline{b}', \overline{c}'$ denote the corresponding group elements. Then,

$$\begin{aligned}
\log_2 |F_C(u;U;\sigma)| &= H(\boldsymbol{f}) \\
&= \frac{1}{2} \cdot H(\overline{a}, \overline{b}, \overline{c}) + \frac{1}{2} \cdot H(\overline{a}, \overline{b}', \overline{c}') \\
&= \frac{1}{2} \cdot \Big(H(\overline{a}, \overline{b}) + H(\overline{c} \mid \overline{a}, \overline{b}) \Big) + \frac{1}{2} \cdot \Big(H(\overline{a}) + H(\overline{b}', \overline{c}' \mid \overline{a}) \Big) \\
&\leqslant \frac{1}{2} \cdot \Big(H(\overline{a}, \overline{b}) + H(\overline{c} \mid \overline{a}) + H(\overline{a}) + H(\overline{b}', \overline{c}') \Big) \\
&= \frac{1}{2} \cdot \Big(H(\overline{a}, \overline{b}) + H(\overline{a}, \overline{c}) + H(\overline{b}', \overline{c}') \Big) \\
&\leqslant \frac{1}{2} \cdot \left(\log_2 |N_{s_2}(u) \cap U| + \log_2 |N_{s_3}(u) \cap U| + \log_2 |N_{s_4}(u) \cap U| \right) ,
\end{aligned}$$

or equivalently,

$$|F_C(u;U;\sigma)| \leqslant \sqrt{|N_{\sigma_2}(u) \cap U| \cdot |N_{\sigma_3}(u) \cap U| \cdot |N_{\sigma_4}(u) \cap U|} \,. \qquad \square$$

In Lemma 13.3.15, we bounded the size of $F_C(u; U; \sigma)$ in terms of $\max_{s \in \{\sigma_2, \sigma_3, \sigma_4\}} |N_s(u) \cap U|$. We also need to establish an upper bound on the number of $u \in U_{\sigma_1}$ with a given value of $\max_{s \in \{\sigma_2, \sigma_3, \sigma_4\}} |N_s(u) \cap U|$. To do this, we use the fact that our cubical complex X is 2^k-expanding, i.e., each bipartite graph $\mathcal{I}_{\sigma_1, s}$ has second eigenvalue at most $2^k \sqrt{d_{\sigma_1 \oplus s}(X)} \leqslant 2^k \sqrt{d_{k/2}(X)}$ (Definition 13.3.4 and Theorem 13.3.5). Here, we use that $\sigma_1 \oplus s$ has weight $k/2$ for $s \in \{\sigma_2, \sigma_3, \sigma_4\}$.

By our assumption that $\max_{i \in [k]} |A_i| \leqslant 2 \cdot \min_{i \in [k]} |A_i|$ and $D = \prod_{i \in [k]} |A_i|$, we have $d_{k/2} := d_{k/2}(X) \leqslant \sqrt{2^k D} = O_k(1) \cdot \sqrt{D}$. For $1 \leqslant \alpha \leqslant 1 + \log_2 d_{k/2}$, define

$$U_{\sigma_1}(\alpha) := \left\{ u \in U_{\sigma_1} : \max_{s \in \{\sigma_2, \sigma_3, \sigma_4\}} |N_s(u) \cap U| \in [2^{\alpha - 1}, 2^\alpha) \right\} .$$

Lemma 13.3.16. *For any $\sigma \in \mathcal{H}_k^4$ with $\sigma_1 \oplus \sigma_2 \oplus \sigma_3 \oplus \sigma_4 = 0$, it holds that*

$$|U_{\sigma_1}(\alpha)| \leqslant O_k(1) \cdot \min \left\{ 1, \frac{\sqrt{D}}{2^{2\alpha}} \right\} \cdot |U| \,.$$

Proof. For $s \in \{\sigma_2, \sigma_3, \sigma_4\}$ and integer $\alpha \leqslant 1 + \log d_{k/2}$, let us define

$$U_{\sigma_1, s}(\alpha) := \left\{ u \in U_{\sigma_1} : 2^{\alpha - 1} \leqslant |N_s(u) \cap U| < 2^\alpha \right\} .$$

Note that

$$|U_{\sigma_1}(\alpha)| \leqslant \sum_{s\in\{\sigma_2,\sigma_3,\sigma_4\}} |U_{\sigma_1,s}(\alpha)|\,,$$

so it suffices to bound each $|U_{\sigma_1,s}(\alpha)|$ separately.

To do this, we count the number of edges between $U_{\sigma_1,s}(\alpha)$ and U_s in $\mathcal{I}_{\sigma_1,s}$ in two different ways. First, by definition each $u \in U_{\sigma_1,s}(\alpha)$ has at least $2^{\alpha-1}$ neighbors within U_s, so we have that

$$|E(U_{\sigma_1,s}(\alpha), U_s)| \geqslant 2^{\alpha-1} \cdot |U_{\sigma_1,s}(\alpha)|\,. \tag{13.4}$$

Second, by the expander mixing lemma on the graph $\mathcal{I}_{\sigma_1,s}$ and using that X is 2^k-expanding and that $d_{k/2}$ is an upper bound on the degree of $\mathcal{I}_{\sigma_1,s}$,

$$\begin{aligned} E(U_{\sigma_1,s}(\alpha), U_s) &\leqslant \frac{d_{k/2}\cdot|U_{\sigma_1,s}(\alpha)|\cdot|U_s|}{|\Gamma|} + 2^k\cdot\sqrt{d_{k/2}}\cdot\sqrt{|U_{\sigma_1,s}(\alpha)|\cdot|U_s|} \\ &\leqslant \left(d_{k/2}\cdot kD^{-1} + 2^k\cdot\sqrt{d_{k/2}}\right)\cdot\sqrt{|U_{\sigma_1,s}(\alpha)|\cdot|U_s|} \\ &\leqslant O_k(1)\cdot D^{1/4}\cdot\sqrt{|U_{\sigma_1,s}(\alpha)|\cdot|U_s|}\,, \end{aligned} \tag{13.5}$$

where in the second line we use that $|U| \leqslant D^{-1}\cdot|\Gamma\times\mathcal{H}_k|$ and in the last line we use that $d_{k/2} = O_k(1)\cdot\sqrt{D}$. Combining Eq. (13.4) and (13.5), this gives that

$$2^{\alpha-1}\cdot|U_{\sigma_1,s}(\alpha)| \leqslant O_k(1)\cdot D^{1/4}\cdot\sqrt{|U_{\sigma_1,s}(\alpha)|\cdot|U_s|}\,,$$

which rearranges to give

$$|U_{\sigma_1,s}(\alpha)| \leqslant O_k(1)\cdot\frac{D^{1/2}}{2^{2\alpha}}\cdot|U_s| \leqslant O_k(1)\cdot\frac{D^{1/2}}{2^{2\alpha}}\cdot|U|\,.$$

Thus,

$$|U_{\sigma_1}(\alpha)| \leqslant \sum_{s\in\{\sigma_2,\sigma_3,\sigma_4\}} |U_{\sigma_1,s}(\alpha)| \leqslant O_k(1)\cdot\frac{D^{1/2}}{2^{2\alpha}}\cdot|U|\,. \tag{13.6}$$

Finally, we obtain the lemma statement by combining Eq. (13.6) with the fact that $|U_{s_1}(\alpha)| \leqslant |U|$. □

We are now ready to prove Lemma 13.3.12.

Proof of Lemma 13.3.12. We first prove $|F_k(U;\sigma)| \leqslant O_k(1)\cdot D^{5/8}\cdot|U|$ for $\sigma = (\sigma_1,\sigma_2,\sigma_3,\sigma_4) \in \mathcal{H}_k^4$ that sums up to 0 over $\mathbb{F}_2^k$.

For the subcube $C = C_\sigma = \sigma_1 \oplus \bigoplus_{i \in \Delta(\sigma)} \{0,1\} \cdot e_i$ (Eq. (13.3)), we can write

$$\begin{aligned}
|F_C(U;\sigma)| &= \sum_{u \in U_{\sigma_1}} |F_C(u;U;\sigma)| \\
&= \sum_{\alpha=1}^{1+\log d_{k/2}} \sum_{u \in U_{\sigma_1}(\alpha)} |F_C(u;U;\sigma)| \\
&\leqslant \sum_{\alpha=1}^{1+\log d_{k/2}} |U_{\sigma_1}(\alpha)| \cdot 2^{3\alpha/2} \\
&\leqslant \sum_{\alpha=1}^{1+\log d_{k/2}} O_k(1) \cdot \min\left\{1, \frac{D^{1/2}}{2^{2\alpha}}\right\} \cdot |U| \cdot 2^{3\alpha/2} \\
&= O_k(1) \sum_{\alpha=1}^{(\log D)/4} 2^{3\alpha/2} \cdot |U| + O_k(1) \sum_{\alpha=1+(\log D)/4}^{1+\log d_{k/2}} \frac{D^{1/2}}{2^{\alpha/2}} \cdot |U| \\
&\leqslant O_k(1) \cdot D^{3/8} \cdot |U|\,,
\end{aligned}$$

where the first inequality follows from Lemma 13.3.15 (since every $u \in U_{\sigma_1}(\alpha)$ satisfies $|N_s(u) \cap U| \leqslant 2^\alpha$ for $s \in \{\sigma_2, \sigma_3, \sigma_4\}$ by definition), and the second inequality follows from Lemma 13.3.16.

Next, by Lemma 13.3.3, each $f \in F_C(U;\sigma)$ can be extended to $f \in F_k(U;\sigma)$ in $\prod_{i \notin \Delta(\sigma)} |A_i| \leqslant O_k(1) \cdot D^{1/4}$ ways, so

$$|F_k(U;\sigma)| \leqslant |F_C(U;\sigma)| \cdot O_k(1) \cdot D^{1/4} \leqslant O_k(1) \cdot D^{5/8} \cdot |U|\,.$$

Finally, by plugging in the above into Eq. (13.2), we obtain the desired inequality:

$$|F_k(U;\geqslant 2\sqrt{k})| \leqslant O_k(1) \cdot D^{5/8} \cdot |U|\,. \qquad \square$$

13.4 Ramanujan cubical complexes

In this section, we give a proof of Theorem 13.3.5, which is essentially contained in [RSV19]. In particular, we describe the construction of expanding cubical complexes (Definition 13.3.4) based on the LPS Ramanujan graphs [LPS88]. For our purposes, we only need basic properties of the generating sets of these Cayley graphs, while using the (highly non-trivial) fact that they are Ramanujan as a black box.

13.4.1 LPS Ramanujan graphs

In this section, we give a brief overview of the LPS Ramanujan graphs [LPS88] (see also [Lub94]).

Notation. For any $n \in \mathbb{N}$, let $r_4(n) := |\{(a,b,c,d) \in \mathbb{Z}^4 : a^2+b^2+c^2+d^2 = n\}|$.

We start with a standard fact.

Fact 13.4.1 (Jacobi's four-square theorem)**.** *For any odd n, $r_4(n) = 8\sum_{m|n} m$. In particular, if $n = p_1 p_2 \cdots p_k$ for distinct odd primes $p_1, \ldots, p_k$, then $r_4(n) = 8\prod_{i=1}^{k}(p_i+1)$.*

Let us start with the definition of quaternions. We will restrict our attention to integral quaternions (a.k.a. Lipschitz quaternions).

Definition 13.4.2 (Integral quaternions)**.** Define $\mathcal{H}(\mathbb{Z}) = \{a\,\mathrm{id} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} : a,b,c,d \in \mathbb{Z}\}$ where

$$\mathbf{i} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \qquad \mathbf{j} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \qquad \mathbf{k} = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \quad \in \mathbb{C}^{2\times 2}.$$

For $\alpha = a\mathrm{id} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathcal{H}(\mathbb{Z})$, we define its norm $N(\alpha)$ as $\det(\alpha) = a^2+b^2+c^2+d^2$, and we define the (normalized) trace $\mathrm{tr}(\alpha) = a$.

Remark 13.4.3. It can be verified that $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ in Definition 13.4.2 satisfy the following relations:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -\mathrm{id}\,.$$

The quaternions are traditionally defined according to these relations. Definition 13.4.2 is a *matrix representation* of quaternions in $\mathbb{C}^{2\times 2}$.

Note that the norm is a multiplicative map: $N(\alpha\beta) = \det(\alpha\beta) = N(\alpha)N(\beta)$. Thus, for integral quaternions, the group of units is

$$\mathcal{H}(\mathbb{Z})^\times = \{\pm\mathrm{id}, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}\}\,.$$

We now formulate the "unique factorization" theorem for $\mathcal{H}(\mathbb{Z})$. This is a key property that we will need later to construct the Ramanujan cubical complexes (see Section 13.4).

Fact 13.4.4 (Unique factorization [Dic22, Theorem 8])**.** *Let $\alpha \in \mathcal{H}(\mathbb{Z})$ such that $N(\alpha)$ is odd.*[5] *Let $N(\alpha) = p_1 p_2 \cdots p_k$ be the factorization of the norm into primes, arranged in an arbitrary but definite order. Then, there is a decomposition $\alpha = \alpha_1\alpha_2\cdots\alpha_k$ where $N(\alpha_i) = p_i$ for each $i \in [k]$. Moreover, the decomposition is unique up to "unit migration", where $\alpha_1\alpha_2\cdots\alpha_k$ and $(\alpha_1 u_1)(\overline{u}_1\alpha_2\overline{u}_2)\cdots(\overline{u}_{k-1}\alpha_k)$ for any $u_1,\ldots,u_{k-1} \in \mathcal{H}(\mathbb{Z})^\times$ are considered the same decomposition.*

Note that factorization can only be unique up to unit migration simply because $\alpha\beta = (\alpha\overline{u})(u\beta)$ for any unit $u \in \mathcal{H}(\mathbb{Z})^\times$.[6]

[5] $N(\alpha)$ being odd is necessary because $2 = (1+\mathbf{i})(1-\mathbf{i}) = (1+\mathbf{j})(1-\mathbf{j})$, which is not unique up to unit migration. One can extend $\mathcal{H}(\mathbb{Z})$ to the *Hurwitz* quaternions to handle this case (see, e.g., [Pal40, CS03]).

[6] This is similar for integers $\mathbb{Z}$ where factorization is unique up to the association $a \sim -a$.

Next, we define the following, which will later give us the generators of the LPS graphs.

Definition 13.4.5. For $n \in \mathbb{N}$, define

$$A(n) \coloneqq \{\alpha \in \mathcal{H}(\mathbb{Z}) : N(\alpha) = n,\ \mathrm{tr}(\alpha) \text{ is odd}\} \,/\, \{\mathrm{id}, -\mathrm{id}\}\,.$$

It is convenient to view this quotient as the set of odd-trace quaternions where α and $-\alpha$ are considered to be identical.

The following fact is a simple consequence of Jacobi's four-square theorem (Fact 13.4.1). We will prove a generalization later (Lemma 13.4.8).

Fact 13.4.6. *For a prime p congruent to 1 modulo 4, $|A(p)| = p + 1$.*

LPS Ramanujan graphs. We now describe the LPS Ramanujan graphs $X(p;q)$, where

- $p < q$ are primes congruent to 1 modulo 4,
- p is a quadratic residue modulo q — that is, there exists $x \in \mathbb{Z}$ such that $p \equiv x^2 \pmod q$.[7]

The graph is a Cayley graph over the group $\mathrm{PSL}(2, \mathbb{F}_q)$ with $p+1$ generators defined by $A(p)$ (Definition 13.4.5). Here, $\mathrm{PSL}(2, \mathbb{F}_q)$ is the *projective special linear group*: it is a subgroup of 2×2 matrices in $\mathbb{F}_q$ of determinant 1 modulo scalar multiplication, i.e., $\widetilde{\alpha}$ belongs to the equivalence class $[c\widetilde{\alpha}]$ if $\det(c\widetilde{\alpha}) = c^2 \det(\widetilde{\alpha}) = 1$ (in $\mathbb{F}_q$). It is easy to check that $|\mathrm{PSL}(2, \mathbb{F}_q)| = q(q^2-1)/2$.

We first need to map a quaternion $\alpha \in A(p)$ to an element in $\mathrm{PSL}(2, \mathbb{F}_q)$. To do so, we need an element $j \in \mathbb{F}_q$ such that $j^2 = -1$ (thus behaving like the imaginary unit i). This requires $q \equiv 1 \pmod 4$, in which case it is well known (by Euler's criterion) that -1 is a quadratic residue mod q, i.e., there exists $y \in \mathbb{Z}$ such that $y^2 \equiv -1 \pmod q$.

Moreover, each $\alpha \in A(p)$ has $\det(\alpha) = p$. We need that there exists $c \in \mathbb{Z}$ such that $\det(c\alpha) = c^2 p \equiv 1 \pmod q$ to get an element in $\mathrm{PSL}(2, \mathbb{F}_q)$. Thus, choosing p such that $p \equiv x^2 \pmod q$ for some $x \in \mathbb{Z}$, since there always exists $c \in \mathbb{Z}$ such that $cx \equiv 1 \pmod q$, we have that $c^2 p \equiv c^2 x^2 \equiv 1 \pmod q$.

This gives a natural map $\alpha \in A(p)$ to $\widetilde{\alpha} \in \mathrm{PSL}(2, \mathbb{F}_q)$ by simply replacing i with $j \in \mathbb{F}_q$ with $j^2 = -1$. We denote

$$\widetilde{A}(p) \coloneqq \{\widetilde{\alpha} : \alpha \in A(p)\}\,.$$

Note that $|\widetilde{A}(p)| = |A(p)| = p + 1$, since no distinct $\alpha, \beta \in A(p)$ are scalar multiples of each other.

The following is the main theorem of [LPS88] whose proof is out of the scope of this paper.

[7] [LPS88] also defined Cayley graphs when p is *not* a quadratic residue. In this case, the graphs are over $\mathrm{PGL}(2, \mathbb{F}_q)$ and they are bipartite. We will not consider this case.

Theorem 13.4.7 ([LPS88]). *Suppose $p < q$ are primes congruent to 1 modulo 4, and p is a quadratic residue modulo q. Let $\Gamma = \mathrm{PSL}(2, \mathbb{F}_q)$. Then, the Cayley graph $\mathrm{Cay}(\Gamma; \widetilde{A}(p))$ is a $(p+1)$-regular graph on $q(q^2-1)/2$ vertices with all non-trivial eigenvalues at most $2\sqrt{p}$.*

13.4.2 Construction of Ramanujan Cayley cubical complexes

The following is an important lemma that allows us to construct cubical complexes. The proof is straightforward given Facts 13.4.1 and 13.4.4.

Lemma 13.4.8. *For any $k \in \mathbb{N}$ and distinct primes $p_1, p_2, \ldots, p_k$ congruent to 1 modulo 4,*

(1) $|A(p_1 p_2 \cdots p_k)| = \prod_{i=1}^{k}(p_i + 1)$.

(2) $A(p_1) \cdot A(p_2) \cdots A(p_k) = A(p_1 p_2 \cdots p_k)$.

Proof. First, note that any number x has $x^2 \equiv 1 \pmod 4$ if x is odd, and 0 otherwise. Thus, $p \equiv 1 \pmod 4$ implies that for $a_0^2 + a_1^2 + a_2^2 + a_3^2 = p$, the set a_0, a_1, a_2, a_3 must have exactly one odd and three even integers. Note also that $p_i \equiv 1 \pmod 4$ implies that $p_1 p_2 \cdots p_k \equiv 1 \pmod 4$.

With a slight abuse of notation, we will view an element α of $A(n)$ as a quaternion even though it is technically a coset $\{\alpha, -\alpha\}$, since $N(\alpha) = N(-\alpha)$ and $\mathrm{tr}(\alpha), \mathrm{tr}(-\alpha)$ have the same parity.

For (1), let $n = p_1 p_2 \cdots p_k$. By Jacobi's four-square theorem (Fact 13.4.1), $r_4(n) = 8\prod_{i=1}^{k}(p_i + 1)$. Since $A(n)$ has the restriction that $\mathrm{tr}(\alpha)$ is odd, each element in $A(n)$ gives rise to 8 distinct 4-tuples of integers whose squares sum up to n (by specifying the position of the odd integer and its sign). This shows that $|A(n)| = \frac{1}{8} r_4(n) = \prod_{i=1}^{k}(p_i + 1)$.

For (2), we first show that for any $n_1 \neq n_2$ congruent to 1 modulo 4, we have $A(n_1) \cdot A(n_2) \subseteq A(n_1 n_2)$. This implies that $A(p_1) \cdot A(p_2) \cdots A(p_k) \subseteq A(p_1 p_2 \cdots p_k)$ as all $p_i \equiv 1 \pmod 4$. For any $\alpha = a_0 \mathrm{id} + a_1 \mathbf{i} + a_2 \mathbf{j} + a_3 \mathbf{k} \in A(n_1)$ and $\beta = b_0 \mathrm{id} + b_1 \mathbf{i} + b_2 \mathbf{j} + b_3 \mathbf{k} \in A(n_2)$, we have that $N(\alpha\beta) = N(\alpha)N(\beta) = n_1 n_2$. Moreover, we know that a_0, b_0 are odd and the rest are even, thus $\mathrm{tr}(\alpha\beta) = a_0 b_0 - a_1 b_1 - a_2 b_2 - a_3 b_3$ is odd. This implies that $\alpha\beta \in A(n_1 n_2)$.

On the other hand, $A(p_1 p_2 \cdots p_k) \subseteq A(p_1) \cdot A(p_2) \cdots A(p_k)$ follows directly from unique factorization (Fact 13.4.4). □

The next lemma follows almost immediately from Theorem 13.4.7 and Lemma 13.4.8.

Lemma 13.4.9. *Let $p_1, p_2, \ldots, p_k$ and q be distinct primes congruent to 1 modulo 4, and suppose each p_i is a quadratic residue modulo q. Let $\Gamma = \mathrm{PSL}(2, \mathbb{F}_q)$. Consider the bipartite graph G defined on $\Gamma \times \{0, 1\}$ where $(g, 0)$ and $(h, 1)$ are connected if and only if $g^{-1}h \in \widetilde{A}(p_1 p_2 \cdots p_k)$. Then, G has degree $d = \prod_{i=1}^{k} |\widetilde{A}(p_i)| = \prod_{i=1}^{k}(p_i + 1)$ and second eigenvalue at most $2^k\sqrt{d}$.*

Proof. By Lemma 13.4.8, we have that $A(p_1) \cdot A(p_2) \cdots A(p_k) = A(p_1 p_2 \cdots p_k)$ and that $|A(p_1 p_2 \cdots p_k)| = \prod_{i=1}^{k}(p_i + 1)$. Thus, the degree $d = \prod_{i=1}^{k}(p_i + 1)$. The adjacency matrix of G is the (bipartite form of) product of adjacency matrices of $\mathrm{Cay}(\Gamma; \widetilde{A}(p_i))$. The trivial eigenvector is the all-ones vector for all these graphs, and thus, by submultiplicativity of the spectral norm, the second eigenvalue of G is at most the product of the second eigenvalues of $\mathrm{Cay}(\Gamma; \widetilde{A}(p_i))$, which is $\prod_{i=1}^{k}(2\sqrt{p_i}) \leqslant 2^k\sqrt{d}$ by Theorem 13.4.7. □

Infinite family of cubical complexes. For any distinct primes $p_1, p_2, \ldots, p_k$, we need to show that there are infinitely many desirable primes q: congruent to 1 modulo 4 and that each p_i is a quadratic residue modulo q. This is standard and follows directly from the law of quadratic reciprocity and the Dirichlet prime number theorem.

Lemma 13.4.10. *Let $p_1, p_2, \ldots, p_k$ be distinct primes congruent to 1 modulo 4. There are infinitely many primes q such that $q \equiv 1 \pmod 4$ and that each p_i is a quadratic residue modulo q.*

Proof. Let $n = p_1 p_2 \cdots p_k$, and consider the arithmetic progression $\{1 + 4n\ell\}_{\ell \in \mathbb{N}}$. The Dirichlet prime number theorem states that this sequence contains infinitely many prime numbers (since 1 and $4n$ are coprime). For any such prime q, we have $q \equiv 1 \pmod 4$ and $q \equiv 1 \pmod{p_i}$ for each i, which also means that q is a quadratic residue modulo p_i. Then, quadratic reciprocity implies that each p_i is a quadratic residue modulo q. □

We also need to argue that there exist such primes that are all within a constant factor apart. This follows from standard facts about the density of primes in arithmetic progressions (see e.g., [BMOR18]).

Fact 13.4.11. *For any $k \in \mathbb{N}$ and $B > 1$, there exists $x_0 = x_0(k, B)$ such that for any $x \geqslant x_0$, there are distinct primes $p_1, p_2, \ldots, p_k \in [x, Bx]$ congruent to 1 modulo 4.*

13.5 Free group action and good quantum LDPC codes

The main result of [LH22b] is a construction of good quantum low density parity check (qLDPC) codes with a linear time decoding algorithm, assuming the existence of two-sided lossless expanders with a free group action, which they left as a conjecture. We state their conjecture below.

Conjecture 13.5.1 ([LH22b], Conjecture 10)**.** *For any $\epsilon > 0$, and for any $\beta \in (0, 1]$ and $\epsilon_0 > 0$, there are $d_L, d_R \in \mathbb{N}$ satisfying $\frac{d_R}{d_L} \in [\beta, \beta + \epsilon_0]$, a constant $\eta > 0$, and an infinite family of (d_L, d_R)-biregular bipartite graphs $\{Z_i = (L_i, R_i, E_i)\}$ and groups $\{G_i\}$, satisfying the following properties:*

(I) Z_i *is a two-sided* $(1-\epsilon)$*-vertex expander. Namely, any* $S \subseteq L_i$ *with* $|S| \leqslant \eta \cdot |L_i|$ *has* $\geqslant (1-\epsilon)d_L \cdot |S|$ *neighbors on the right, and any* $S \subseteq R_i$ *with* $|S| \leqslant \eta \cdot |R_i|$ *has* $\geqslant (1-\epsilon)d_R \cdot |S|$ *neighbors on the left.*

(II) $|G_i| = \Theta(|Z_i|)$*, and* Z_i *has a free* G_i*-action.*

Lin and M. Hsieh used such two-sided lossless expanders to construct good qLDPC codes.

Theorem 13.5.2 ([LH22b], Theorem 9 and Theorem 14)**.** *Assuming Conjecture 13.5.1, then for all* $r \in (0,1)$*, there exists* $\delta > 0$*,* $w \in \mathbb{N}$ *and a infinite family of quantum error-correcting codes* $C = \{C_i\}_{i \in \mathbb{N}}$ *with parameters* $[[n_i, k_i, d_i]]$*, such that* $k_i/n_i > r$*,* $d_i/n_i > \delta$*, and all stabilizers of* C_i *have weight* w*. Furthermore,* C *has a linear time decoding algorithm.*

In what follows, we show that the graphs we construct in Section 13.2 resolve Conjecture 13.5.1, thereby giving a new instantiation of qLDPC codes via the framework of [LH22b]. We have already proved Condition (I) in Theorem 10.2.2. It remains simply to check that the groups G_i satisfying Condition (II) exist.

Proposition 13.5.3. *The graph* Z *constructed in Section 13.2, using Cayley cubical complexes over* $\Gamma = \mathrm{PSL}(2, \mathbb{F}_q)$*, has a free* Γ*-action.*

Proof. We begin by recalling some notation. Let $X = \mathrm{Cay}(\Gamma, \mathcal{A})$ be a cubical complex over Γ, where $\mathcal{A} = \{A_1, \ldots, A_k\}$ are k sets of Cayley cubical generators. The graphs $G_L = (L, M, E_L)$ and $G_R = (M, R, E_R)$ are defined as follows:

- $L = \{v \in X(k) : \mathrm{Signature}(v) \in \mathcal{S}_L\}$, where $\mathcal{S}_L \subseteq \underline{\mathcal{S}_L}$ is a D_L-sized collection of signatures,
- $R = \{v \in X(k) : \mathrm{Signature}(v) \in \mathcal{S}_R\}$, where $\mathcal{S}_R \subseteq \underline{\mathcal{S}_R}$ is a D_R-sized collection of signatures (see Section 13.3.1),
- $M = \Gamma \times \mathcal{H}_k$,
- $(f, u) \in E_L$ if $u \in f$, and $(u, f) \in E_R$ if $u \in f$.

Then, the graph Z was constructed by placing a copy of the gadget graph H on the left and right neighbors of each $u \in M$. Precisely, for each edge $(i, j) \in H$, we place an edge between $\mathrm{LNbr}_u(i)$ and $\mathrm{RNbr}_u(j)$.

We claim that Z has a free left Γ-action. This will essentially follow from the observations that G_L and G_R permit a free left Γ-action, and the placement of the gadget H respects the group structure.

More concretely, let us define the left Γ-action on $u = (g, x) \in M$ as follows:

$$\gamma u := (\gamma g, x).$$

We can also define a left Γ-action on $\underline{L} = \underline{R} = X(k)$: for $f = \{(f_x, x)\}_{x \in \{0,1\}^k}$, we define

$$\gamma f := \{(\gamma f_x, x)\}_{x \in \mathcal{H}_k}.$$

It turns out that because the cubical generating sets A_i all act on the right, this defines a legal action on $X(k)$ as well, which we check by verifying $\gamma f \in X(k)$:

$$(\gamma f)_x^{-1}(\gamma f)_{x+e_i} = f_x^{-1}\gamma^{-1}\gamma f_{x+e_i} = f_x^{-1} f_{x+e_i} \in A_i. \tag{13.7}$$

Both the above actions are free because Γ acting on itself is free. This will imply that the left Γ-action on Z, which has vertex set a subset of $X(k)$, is free as well.

Eq. (13.7) actually implies something even stronger: acting on the left by γ preserves the signature of the cube. It follows that the subsets $L \subseteq \underline{L}$ and $R \subseteq \underline{R}$ also permit a free left Γ-action, since L and R consist of all cubes with a certain collection of signatures. Now looking at the base graph G_L, we define for $(f, u) \in E_L$

$$\gamma(f, u) := (\gamma f, \gamma u).$$

This defines a valid left Γ-action on E_L, since if $u \in f$ then $\gamma u \in \gamma f$. Similarly, we can define for $(u, f) \in E_R$

$$\gamma(u, f) := (\gamma u, \gamma f).$$

Note in particular that if f is the neighbor of u with a given signature σ, then γf is the neighbor of γu with signature σ.

Next, we show that the placement of the gadget graph H respects the left Γ action. Recall that in Section 13.3.1, LNbr_u (similarly, RNbr_u) were defined so that $\mathrm{Signature}(\mathrm{LNbr}_u(i)) = \mathrm{Signature}(\mathrm{LNbr}_{u'}(i))$ for any $u, u' \in \Gamma \times \{\sigma\}$, $\sigma \in \mathcal{H}_k$. From the above discussion, this implies that

$$\gamma \mathrm{LNbr}_u(i) = \mathrm{LNbr}_{\gamma u}(i).$$

In particular, under a left Γ-action, an edge $(\mathrm{LNbr}_u(i), \mathrm{RNbr}_u(j)) \in E$ gets sent to

$$\gamma(\mathrm{LNbr}_u(i), \mathrm{RNbr}_u(j)) := (\gamma \mathrm{LNbr}_u(i), \gamma \mathrm{RNbr}_u(j)) = (\mathrm{LNbr}_{\gamma u}(i), \mathrm{RNbr}_{\gamma u}(j)) \in E.$$

Finally, we check that Γ has linear size:

$$|Z| \leqslant 2|X(k)| = 2|\Gamma| \cdot \prod_{i=1}^{k} |A_i| = O_k(1) \cdot |\Gamma|. \qquad \square$$

Bibliography

[ABS15] Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential Algorithms for Unique Games and Related Problems. *Journal of the ACM (JACM)*, 62(5):1–25, 2015. 1.2, 6.3

[AC88] Noga Alon and Fan RK Chung. Explicit construction of linear sized tolerant networks. *Discrete Mathematics*, 72(1-3):15–19, 1988. 1, 1.1, 2.1.1

[AC02] Noga Alon and Michael Capalbo. Explicit unique-neighbor expanders. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 73–79. IEEE, 2002. 10.1, 11, 11.0.2

[ACC06] Sanjeev Arora, Eden Chlamtac, and Moses Charikar. New approximation guarantee for chromatic number. In *38th Annual ACM Symposium on Theory of Computing, STOC'06*, pages 215–224, 2006. 6.3

[ACIM01] Dimitris Achlioptas, Arthur Chtcherba, Gabriel Istrate, and Cristopher Moore. The phase transition in 1-in-k SAT and NAE 3-SAT. In *Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms*, pages 721–722, 2001. 8.2.3

[AD24] Ron Asherov and Irit Dinur. Bipartite unique neighbour expanders via ramanujan graphs. *Entropy*, 26(4):348, 2024. 1.1, (2), 3.2.2, 3.2.2, 10.1, 11.0.2

[AE98] Gunnar Andersson and Lars Engebretsen. Better approximation algorithms for Set splitting and Not-All-Equal SAT. *Information Processing Letters*, 65(6):305–311, 1998. 8.2.3

[AF09] Noga Alon and Uriel Feige. On the power of two, three and four probes. In *Proceedings of the twentieth annual ACM-SIAM symposium on Discrete algorithms*, pages 346–354. SIAM, 2009. 3.1.2

[AG11] Sanjeev Arora and Rong Ge. New Tools for Graph Coloring. In *International Workshop on Approximation Algorithms for Combinatorial Optimization*, pages 1–12. Springer, 2011. 6.3

[AGK21] Jackson Abascal, Venkatesan Guruswami, and Pravesh K Kothari. Strongly refuting all semi-random Boolean CSPs. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 454–

472. SIAM, 2021. 6.1, 6.2, 6.2, 8, 8.1.2

[AHL02] Noga Alon, Shlomo Hoory, and Nathan Linial. The Moore bound for irregular graphs. *Graphs and Combinatorics*, 18(1):53–57, 2002. 1.1, (1), 3.1, 3.1.1, 4

[AK97] Noga Alon and Nabil Kahale. A Spectral Technique for Coloring Random 3-Colorable Graphs. *SIAM Journal on Computing*, 26(6):1733–1748, 1997. 6.3

[AK98] Noga Alon and Nabil Kahale. Approximating the independence number via the ϑ-function. *Mathematical Programming*, 80(3):253–264, 1998. 6.3

[AKK95] Sanjeev Arora, David R. Karger, and Marek Karpinski. Polynomial time approximation schemes for dense instances of *NP*-hard problems. In *Proceedings of the Twenty-Seventh Annual ACM Symposium on Theory of Computing, 29 May-1 June 1995, Las Vegas, Nevada, USA*, pages 284–293. ACM, 1995. 6.1

[AKK+08] Sanjeev Arora, Subhash A Khot, Alexandra Kolla, David Steurer, Madhur Tulsiani, and Nisheeth K Vishnoi. Unique games on expanding constraint graphs are easy. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 21–28, 2008. 1.2, 6.3

[AKS98] Noga Alon, Michael Krivelevich, and Benny Sudakov. Finding a large hidden clique in a random graph. *Random Structures & Algorithms*, 13(3-4):457–466, 1998. 1

[ALM96] Sanjeev Arora, FT Leighton, and Bruce M Maggs. On-Line Algorithms for Path Selection in a Nonblocking Network. *SIAM Journal on Computing*, 25(3):600–625, 1996. 10.1

[Alo86] Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986. 1

[AOW15] Sarah R Allen, Ryan O'Donnell, and David Witmer. How to refute a random CSP. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 689–708. IEEE, 2015. 1.2, 6.1, 6.2, 7, 8

[App16] Benny Applebaum. Cryptographic Hardness of Random Local Functions: Survey. *Computational complexity*, 25:667–722, 2016. 6.2

[AR94] Noga Alon and Yuval Roichman. Random Cayley graphs and expanders. *Random Structures & Algorithms*, 5(2):271–284, 1994. 13.1.1

[Bas92] Hyman Bass. The Ihara-Selberg zeta function of a tree lattice. *International Journal of Mathematics*, 3(06):717–797, 1992. 2.2, 5

[BBKSS21] Mitali Bafna, Boaz Barak, Pravesh K Kothari, Tselil Schramm, and David

Steurer. Playing Unique Games on Certified Small-Set Expanders. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1629–1642, 2021. 6.3, 9.1.1

[BCK15] Boaz Barak, Siu On Chan, and Pravesh K. Kothari. Sum of Squares Lower Bounds from Pairwise Independence. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 97–106. ACM, 2015. 6.1

[BFSU98] Andrei Z Broder, Alan M Frieze, Stephen Suen, and Eli Upfal. Optimal construction of edge-disjoint paths in random graphs. *SIAM Journal on Computing*, 28(2):541–573, 1998. 10.1

[BGIKS08] Radu Berinde, Anna C Gilbert, Piotr Indyk, Howard Karloff, and Martin J Strauss. Combining geometry and combinatorics: A unified approach to sparse signal recovery. In *2008 46th Annual Allerton Conference on Communication, Control, and Computing*, pages 798–805. IEEE, 2008. 10.1

[BGLR93] Mihir Bellare, Shafi Goldwasser, Carsten Lund, and Alexander Russell. Efficient probabilistically checkable proofs and applications to approximations. In *Proceedings of the twenty-fifth annual ACM symposium on Theory of computing*, pages 294–304, 1993. 8.1.2

[BGMT12] Siavosh Benabbas, Konstantinos Georgiou, Avner Magen, and Madhur Tulsiani. SDP gaps from pairwise independence. *Theory of Computing*, 8(1):269–289, 2012. 6.1

[BH92] Ravi Boppana and Magnús M Halldórsson. Approximating maximum independent sets by excluding subgraphs. *BIT Numerical Mathematics*, 32(2):180–196, 1992. 6.3

[BHK25] Mitali Bafna, Jun-Ting Hsieh, and Pravesh K Kothari. Rounding Large Independent Sets on Expanders. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 631–642, 2025. 1.4, (3)

[BHL+02] Wolfgang Barthel, Alexander K Hartmann, Michele Leone, Federico Ricci-Tersenghi, Martin Weigt, and Riccardo Zecchina. Hiding solutions in random satisfiability problems: A statistical mechanics approach. *Physical review letters*, 88(18):188701, 2002. 6.2, 8

[BHSV25] Rares-Darius Buhai, Yiding Hua, David Steurer, and Andor Vári-Kakas. Finding Colorings in One-Sided Expanders. In *2025 IEEE 66th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2025. 6.3

[Big93] Norman Biggs. *Algebraic graph theory*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 1993. 3.1

[BK97] Avrim Blum and David Karger. An $\widetilde{O}(n^{3/14})$-coloring algorithm for 3-

colorable graphs. *Information processing letters*, 61(1):49–53, 1997. 6.3

[BK09] Nikhil Bansal and Subhash Khot. Optimal Long Code Test with One Free Bit. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 453–462. IEEE, 2009. 6.3, 6.3, 9.4, 9.4.1

[BKHL99] Claudia Bertram-Kretzberg, Thomas Hofmeister, and Hanno Lefmann. Sparse 0- 1 matrices and forbidden hypergraphs. *Combinatorics, Probability and Computing*, 8(5):417–427, 1999. 3.1.2

[Blu94] Avrim Blum. New approximation algorithms for graph coloring. *Journal of the ACM (JACM)*, 41(3):470–516, 1994. 6.3

[BM16] Boaz Barak and Ankur Moitra. Noisy Tensor Completion via the Sum-of-Squares Hierarchy. In *Proceedings of the 29th Conference on Learning Theory, COLT 2016, New York, USA, June 23-26, 2016*, volume 49 of *JMLR Workshop and Conference Proceedings*, pages 417–445. JMLR.org, 2016. 6.1

[BMOR18] Michael A Bennett, Greg Martin, Kevin O'Bryant, and Andrew Rechnitzer. Explicit bounds for primes in arithmetic progressions. *Illinois Journal of Mathematics*, 62(1-4):427–532, 2018. 13.4.2

[BMS08] Louay Bazzi, Mohammad Mahdian, and Daniel A Spielman. The minimum distance of turbo-like codes. *IEEE Transactions on Information Theory*, 55(1):6–15, 2008. 3.1.2

[Bol78] Béla Bollobás. *Extremal graph theory*, volume 11 of *London Mathematical Society Monographs*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London-New York, 1978. 3.1

[Bor20] Charles Bordenave. A new proof of Friedman's second eigenvalue Theorem and its extension to random lifts. In *Annales Scientifiques de l'École Normale Supérieure*, volume 4, pages 1393–1439, 2020. 1.3, 3.1.1

[BQ09] Andrej Bogdanov and Youming Qiao. On the security of Goldreich's one-way function. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques: 12th International Workshop, APPROX 2009*, pages 392–405. Springer, 2009. 6.2

[BR14] S Ajesh Babu and Jaikumar Radhakrishnan. An entropy-based proof for the Moore bound for irregular graphs. In *Perspectives in Computational Complexity*, pages 173–181. Springer, 2014. 3.1

[BRS11] Boaz Barak, Prasad Raghavendra, and David Steurer. Rounding Semidefinite Programming Hierarchies via Global Correlation. In *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011*, pages 472–481. IEEE, 2011. 2.5.2, 9.1.1

[BS95] Avrim Blum and Joel Spencer. Coloring Random and Semi-Random k-Colorable Graphs. *J. Algorithms*, 19(2):204–234, 1995. 1.2, 6.1, 6.3

[BS14] Boaz Barak and David Steurer. Sum-of-squares proofs and the quest toward optimal algorithms. *arXiv preprint arXiv:1404.5236*, 2014. 6.2

[BS16] Boaz Barak and David Steurer. Proofs, beliefs, and algorithms through the lens of sum-of-squares. *Course notes:* http://www.sumofsquares.org/public/index.html, 2016. 2.5

[BV09] Eli Ben-Sasson and Michael Viderman. Tensor products of weakly smooth codes are robust. *Theory of Computing*, 5(1):239–255, 2009. 10.1

[CCF10] Amin Coja-Oghlan, Colin Cooper, and Alan Frieze. An efficient sparse regularity concept. *SIAM Journal on Discrete Mathematics*, 23(4):2000–2034, 2010. 6.2

[CGL07] Amin Coja-Oghlan, Andreas Goerdt, and André Lanka. Strong refutation heuristics for random k-SAT. *Combinatorics, Probability & Computing*, 16(1):5, 2007. 6.1, 8.1.5

[CGRZ24] Eshan Chattopadhyay, Mohit Gurumukhani, Noam Ringach, and Yunya Zhao. Two-Sided Lossless Expanders in the Unbalanced Setting. *arXiv preprint arXiv:2409.04549*, 2024. 10.1

[Che25] Yeyuan Chen. Unique-neighbor Expanders with Better Expansion for Polynomial-sized Sets. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 3335–3362. SIAM, 2025. 12.1

[Chl09] Eden Chlamtac. *Non-local analysis of SDP-based approximation algorithms*. Princeton University, 2009. 6.3

[CRT23] Itay Cohen, Roy Roth, and Amnon Ta-Shma. HDX condensers. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1649–1664. IEEE, 2023. 10.1

[CRVW02] Michael Capalbo, Omer Reingold, Salil Vadhan, and Avi Wigderson. Randomness conductors and constant-degree lossless expanders. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pages 659–668, 2002. 1.3, 10.1

[CS03] John H Conway and Derek A Smith. *On quaternions and octonions*. AK Peters/CRC Press, 2003. 5

[DELLM22] Irit Dinur, Shai Evra, Ron Livne, Alexander Lubotzky, and Shahar Mozes. Locally Testable Codes with constant rate, distance, and locality. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 357–374, 2022. 10.1, 13.1.1

[DF16] Roee David and Uriel Feige. On the effect of randomness on planted 3-coloring models. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 77–90, 2016. 6.3, 6.3, 5

[DHV16] Amit Deshpande, Prahladh Harsha, and Rakesh Venkat. Embedding Approximately Low-Dimensional ℓ_2^2 Metrics into ℓ_1. In *36th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2016)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016. 6.3

[Dic22] Leonard E Dickson. Arithmetic of quaternions. *Proceedings of the London Mathematical Society*, 2(1):225–232, 1922. 13.4.4

[Din24] Irit Dinur. Expanders and PCPs: Emergence from Local to Global. FOCS 2024 Plenary Talk, YouTube video, 2024. https://www.youtube.com/watch?v=5eGoy6NfkZE. 10.2

[DKPS10] Irit Dinur, Subhash Khot, Will Perkins, and Muli Safra. Hardness of Finding Independent Sets in Almost 3-Colorable Graphs. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 212–221. IEEE, 2010. 6.3

[DLV24] Irit Dinur, Ting-Chun Lin, and Thomas Vidick. Expansion of High-Dimensional Cubical Complexes: with Application to Quantum Locally Testable Codes. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 379–385. IEEE, 2024. 13.1.1, 13.1.1

[DMR06] Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional Hardness for Approximate Coloring. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of Computing*, pages 344–353, 2006. 6.3, 9, 9.4, 9.4.4, 9.4

[DS05] Irit Dinur and Samuel Safra. On the hardness of approximating minimum vertex cover. *Annals of mathematics*, pages 439–485, 2005. 6.3, 9.4

[DSS14] Jian Ding, Allan Sly, and Nike Sun. Satisfiability threshold for random regular NAE-SAT. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 814–822, 2014. 8.2.3

[DSW06] Irit Dinur, Madhu Sudan, and Avi Wigderson. Robust local testability of tensor products of LDPC codes. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques.*, pages 304–315. Springer, 2006. 10.1

[Fei02] Uriel Feige. Relations between average case complexity and approximation complexity. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 534–543, 2002. 6.1, 1, 7

[Fei04] Uriel Feige. Approximating maximum clique by removing subgraphs.

SIAM Journal on Discrete Mathematics, 18(2):219–225, 2004. 6.3

[Fei07] Uriel Feige. Refuting Smoothed 3CNF Formulas. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2007), October 20-23, 2007, Providence, RI, USA, Proceedings*, pages 407–417. IEEE Computer Society, 2007. 6.1

[Fei08] Uriel Feige. Small linear dependencies for binary vectors of low weight. In *Building Bridges: Between Mathematics and Computer Science*, pages 283–307. Springer, 2008. 1.1, 1.2, (1), 3.1.2, 3.1.5, 6.1

[FK01] Uriel Feige and Joe Kilian. Heuristics for Semirandom Graph Problems. *J. Comput. Syst. Sci.*, 63(4):639–671, 2001. 6.1

[FKP19] Noah Fleming, Pravesh Kothari, and Toniann Pitassi. Semialgebraic Proofs and Efficient Algorithm Design. *Foundations and Trends® in Theoretical Computer Science*, 14(1-2):1–221, 2019. 2.5, 6.2

[FLP16] Dimitris Fotakis, Michael Lampis, and Vangelis Th. Paschos. Sub-exponential Approximation Schemes for CSPs: From Dense to Almost Sparse. In *33rd Symposium on Theoretical Aspects of Computer Science, STACS 2016, February 17-20, 2016, Orléans, France*, volume 47 of *LIPIcs*, pages 37:1–37:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016. 6.1, 6.1, 6.2

[FM17] Zhou Fan and Andrea Montanari. How well do local algorithms solve semidefinite programs? In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 604–614, 2017. 2.2

[FPV15] Vitaly Feldman, Will Perkins, and Santosh S. Vempala. Subsampled Power Iteration: a Unified Algorithm for Block Models and Planted CSP's. In *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015, Montreal, Quebec, Canada*, pages 2836–2844, 2015. 1.2, 6.2, 6.2, 8, 8, 8.1, 8.2, 8.2, 4

[FPV18] Vitaly Feldman, Will Perkins, and Santosh S. Vempala. On the Complexity of Random Satisfiability Problems with Planted Solutions. *SIAM Journal on Computing*, 47(4):1294–1338, 2018. 6.1

[Fri08] Joel Friedman. *A proof of Alon's second eigenvalue conjecture and related problems*. American Mathematical Soc., 2008. 1.3

[GHKM23] Venkatesan Guruswami, Jun-Ting Hsieh, Pravesh K Kothari, and Peter Manohar. Efficient Algorithms for Semirandom Planted CSPs at the Refutation Threshold. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 307–327. IEEE, 2023. 1.4, (2), 8.2.3

[GKM22] Venkatesan Guruswami, Pravesh K Kothari, and Peter Manohar. Algorithms and certificates for Boolean CSP refutation: smoothed is no harder than random. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 678–689, 2022. 1.1, 1.2, (1), 3.1.2, 3.1.2, 4.2, 4.2, 4.3, 4.4, 6.1, 6.1, 6.2, 7, 7.0.2, 8, 8.1.5, 8.3

[GL03] Andreas Goerdt and André Lanka. Recognizing more random unsatisfiable 3-sat instances efficiently. *Electron. Notes Discret. Math.*, 16:21–46, 2003. 6.1

[GLR10] Venkatesan Guruswami, James R Lee, and Alexander Razborov. Almost Euclidean subspaces of ℓ_1^N via expander codes. *Combinatorica*, 30(1):47–68, 2010. 10.1

[GM88] Chris D Godsil and Bojan Mohar. Walk generating functions and spectral measures of infinite graphs. *Linear Algebra and its Applications*, 107:191–206, 1988. 2.1

[GMM22] Venkatesan Guruswami, Peter Manohar, and Jonathan Mosheiff. ℓ_p-Spread and Restricted Isometry Properties of Sparse Random Matrices. In *37th Computational Complexity Conference (CCC 2022)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2022. 10.1

[Gol00] Oded Goldreich. Candidate One-Way Functions Based on Expander Graphs. *Electron. Colloquium Comput. Complex.*, 2000. 6.2

[Gol24] Louis Golowich. New explicit constant-degree lossless expanders. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 4963–4971. SIAM, 2024. 10.1, 11.0.2

[GS20] Venkatesan Guruswami and Sai Sandeep. d-to-1 hardness of coloring 3-colorable graphs with $O(1)$ colors. In *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020. 6.3

[GUV09] Venkatesan Guruswami, Christopher Umans, and Salil Vadhan. Unbalanced expanders and randomness extractors from Parvaresh–Vardy codes. *Journal of the ACM (JACM)*, 56(4):1–34, 2009. 10.1

[Hae95] Willem H Haemers. Interlacing eigenvalues and graphs. *Linear Algebra and its applications*, 226:593–616, 1995. 2.1

[Hås01] Johan Håstad. Some optimal inapproximability results. *Journal of the ACM (JACM)*, 48(4):798–859, 2001. 6.1

[HKM23] Jun-Ting Hsieh, Pravesh K. Kothari, and Sidhanth Mohanty. A simple and sharper proof of the hypergraph Moore bound. In *Proceedings of the 2023 ACM-SIAM Symposium on Discrete Algorithms, SODA 2023, Florence, Italy,*

January 22-25, 2023, pages 2324–2344. SIAM, 2023. 1.4, (1), (1), 6.2, 8

[HKMMS25] Jun-Ting Hsieh, Pravesh K Kothari, Sidhanth Mohanty, David Munhá Correia, and Benny Sudakov. Small Even Covers, Locally Decodable Codes and Restricted Subgraphs of Edge-Colored Kikuchi Graphs. *International Mathematics Research Notices*, 2025(5):rnaf045, 2025. 3.1.2

[HLMOZ25] Jun-Ting Hsieh, Ting-Chun Lin, Sidhanth Mohanty, Ryan O'Donnell, and Rachel Yun Zhang. Explicit Two-Sided Vertex Expanders Beyond the Spectral Barrier. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, 2025. 1.3, 1.4, 10, 10.2, 10.2, 11.1, 11.1.1, 13, 13.1.2, 13.1.3, 4, 13.2.7, 13.3.1

[HLMRZ25] Jun-Ting Hsieh, Alexander Lubotzky, Sidhanth Mohanty, Assaf Reiner, and Rachel Yun Zhang. Explicit Lossless Vertex Expanders. In *2025 IEEE 66th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2025. 1.3, 1.4, 10, 10.2, 10.2

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006. 1, 1.3, 3.2, 10

[HMMP24] Jun-Ting Hsieh, Theo McKenzie, Sidhanth Mohanty, and Pedro Paredes. Explicit two-sided unique-neighbor expanders. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 788–799, 2024. 1.3, 1.4, (2), 10, 10.2, 10.2, 11.1, 12.2.1, 13.1.3

[HMP06] Shlomo Hoory, Avner Magen, and Toniann Pitassi. Monotone circuits for the majority function. In *International Workshop on Approximation Algorithms for Combinatorial Optimization*, pages 410–425. Springer, 2006. 10.1

[HMY24] Jiaoyang Huang, Theo McKenzie, and Horng-Tzer Yau. Ramanujan Property and Edge Universality of Random Regular Graphs. *arXiv preprint arXiv:2412.20263*, 2024. 2

[Hof70] Alan J Hoffman. On eigenvalues and colorings of graphs. In *Graph Theory and its Applications*, pages 79–91. Acad. Press, 1970. 6.3

[Hoo02] Shlomo Hoory. The size of bipartite graphs with a given girth. *Journal of Combinatorial Theory, Series B*, 86(2):215–220, 2002. 3.1

[Hsi25] Jun-Ting Hsieh. Coloring 3-Colorable Graphs with Low Threshold Rank. *arXiv preprint arXiv:2508.03093*, 2025. 6.3

[Iha66] Yasutaka Ihara. On discrete subgroups of the two by two projective linear group over p-adic fields. *Journal of the Mathematical Society of Japan*, 18(3):219–235, 1966. 2.2, 5

[IP01] Russell Impagliazzo and Ramamohan Paturi. On the Complexity of k-

SAT. *J. Comput. Syst. Sci.*, 62(2):367–375, 2001. 6.1

[Jer92] Mark Jerrum. Large cliques elude the Metropolis process. *Random Structures & Algorithms*, 3(4):347–359, 1992. 6.3

[JL00] Bruce W Jordan and Ron Livné. The Ramanujan property for regular cubical complexes. *Duke Math. J.*, 104(1):85–103, 2000. 13.1.1

[JMOPT22] Fernando Granha Jeronimo, Tushant Mittal, Ryan O'Donnell, Pedro Paredes, and Madhur Tulsiani. Explicit Abelian Lifts and Quantum LDPC Codes. In *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, pages 88–1. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2022. 13.1.1, 13.1.1

[JMS07] Haixia Jia, Cristopher Moore, and Doug Strain. Generating Hard Satisfiable Formulas by Hiding Solutions Deceptively. *Journal of Artificial Intelligence Research*, 28:107–118, 2007. 6.2, 8

[Kah95] Nabil Kahale. Eigenvalues and expansion of regular graphs. *Journal of the ACM (JACM)*, 42(5):1091–1106, 1995. 1.1, 1.3, (2), 3.2.1, 3.2.1, 10, 10.1, 12, 12.1

[Kar72] Richard M Karp. *Reducibility among combinatorial problems*. Springer, 1972. 6.3

[Kar94] David R Karger. Random sampling in cut, flow, and network design problems. In *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*, pages 648–657, 1994. 8.1.3, 8.1.3

[Kar11] Zohar S Karnin. Deterministic construction of a high dimensional ℓ_p section in ℓ_1^n for any $p < 2$. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 645–654, 2011. 10.1

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 767–775, 2002. 9, 9.4.3

[KK22] Amitay Kamber and Tali Kaufman. Combinatorics via closed orbits: number theoretic Ramanujan graphs are not unique neighbor expanders. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 426–435, 2022. 3.2.1, 10.1

[KLT18] Akash Kumar, Anand Louis, and Madhur Tulsiani. Finding Pseudorandom Colorings of Pseudorandom Graphs. In *37th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018. 6.3

[KMOW17] Pravesh K. Kothari, Ryuhei Mori, Ryan O'Donnell, and David Witmer.

Sum of squares lower bounds for refuting any CSP. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 132–145. ACM, 2017. 6.1

[KMS98] David Karger, Rajeev Motwani, and Madhu Sudan. Approximate Graph Coloring by Semidefinite Programming. *Journal of the ACM (JACM)*, 45(2):246–265, 1998. 6.3, 6.3, 9, 9.5

[KR08] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within 2-epsilon. *J. Comput. Syst. Sci.*, 74(3):335–349, 2008. 6.3, 9.4

[KRS23] Swastik Kopparty, Noga Ron-Zewi, and Shubhangi Saraf. Simple Constructions of Unique Neighbor Expanders from Error-correcting Codes. *arXiv preprint arXiv:2310.19149*, 2023. 10.1

[KS12] Subhash Khot and Rishi Saket. Hardness of Finding Independent Sets in Almost q-Colorable Graphs. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 380–389. IEEE, 2012. 6.3

[KT17] Ken-ichi Kawarabayashi and Mikkel Thorup. Coloring 3-Colorable Graphs with Less than $n^{1/5}$ Colors. *Journal of the ACM (JACM)*, 64(1):1–23, 2017. 6.3

[KT22] Itay Kalev and Amnon Ta-Shma. Unbalanced Expanders from Multiplicity Codes. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2022)*, pages 12–1. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2022. 10.1

[KTY24] Ken-ichi Kawarabayashi, Mikkel Thorup, and Hirotaka Yoneda. Better coloring of 3-Colorable graphs. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 331–339, 2024. 6.3

[Kuč95] Luděk Kučera. Expected complexity of graph partitioning problems. *Discrete Applied Mathematics*, 57(2-3):193–212, 1995. 6.3

[KVV04] Ravi Kannan, Santosh Vempala, and Adrian Vetta. On clusterings: Good, bad and spectral. *Journal of the ACM (JACM)*, 51(3):497–515, 2004. 2.3, 8.1.4

[KY24] Dmitriy Kunisky and Xifan Yu. Computational hardness of detecting graph lifts and certifying lift-monotone properties of random regular graphs. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1621–1633. IEEE, 2024. 10.1

[Las01] Jean B. Lasserre. Global Optimization with Polynomials and the Problem of Moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001. 2.5.1

[LH22a] Ting-Chun Lin and Min-Hsiu Hsieh. c^3-Locally Testable Codes from Lossless Expanders. In *2022 IEEE International Symposium on Information Theory*

(ISIT), pages 1175–1180. IEEE, 2022. 10.1

[LH22b] Ting-Chun Lin and Min-Hsiu Hsieh. Good quantum LDPC codes with linear time decoder from lossless expanders. *arXiv preprint arXiv:2203.03581*, 2022. 10.1, 10.2.3, 13, 13.5, 13.5.1, 13.5.2, 13.5

[LPS88] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988. 1.3, 3.2, 10.2, 12, 13, 13.1.1, 13.1.1, 13.3, 13.4, 13.4.1, 13.4.1, 7, 13.4.7

[LS96] Wen-Ch'ing Winnie Li and Patrick Solé. Spectra of regular graphs and hypergraphs and orthogonal polynomials. *European Journal of Combinatorics*, 17(5):461–477, 1996. 2.1

[LSV05a] Alexander Lubotzky, Beth Samuels, and Uzi Vishne. Explicit constructions of Ramanujan complexes of type $\widetilde{A}_d$. *European Journal of Combinatorics*, 26(6):965–993, 2005. 10.2, 13

[LSV05b] Alexander Lubotzky, Beth Samuels, and Uzi Vishne. Ramanujan complexes of type $\widetilde{A}_d$. *Israel Journal of Mathematics*, 149:267–299, 2005. Probability in mathematics. 10.2, 13

[Lub94] Alex Lubotzky. *Discrete groups, expanding graphs and invariant measures*, volume 125. Springer Science & Business Media, 1994. 13.4.1

[LW49] LH Loomis and H Whitney. An inequality related to the isoperimetric inequality. *Bulletin of the American Mathematical Society*, 55(10):961–962, 1949. 13.1.3

[Mar88] Grigorii Aleksandrovich Margulis. Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators. *Problemy peredachi informatsii*, 24(1):51–60, 1988. 1.3, 3.2

[MM11] Konstantin Makarychev and Yury Makarychev. How to play unique games on expanders. In *Approximation and Online Algorithms: 8th International Workshop, WAOA 2010, Liverpool, UK, September 9-10, 2010. Revised Papers 8*, pages 190–200. Springer, 2011. 1.2, 6.3

[MM21] Theo McKenzie and Sidhanth Mohanty. High-Girth Near-Ramanujan Graphs with Lossy Vertex Expansion. In *48th International Colloquium on Automata, Languages, and Programming (ICALP 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2021. 10.1

[MOP20] Sidhanth Mohanty, Ryan O'Donnell, and Pedro Paredes. Explicit near-Ramanujan graphs of every degree. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 510–523, 2020. 3.1.2, 3.2.2, 12, 12

[Mor94] Moshe Morgenstern. Existence and explicit constructions of q+ 1 regular Ramanujan graphs for every prime power q. *Journal of Combinatorial Theory, Series B*, 62(1):44–62, 1994. 1.3, 3.2, 12

[Mos15] Dana Moshkovitz. The Projection Games Conjecture and the NP-Hardness of ln n-Approximating Set-Cover. *Theory Comput.*, 11:221–235, 2015. 8.1.2

[MR10] Dana Moshkovitz and Ran Raz. Two-query PCP with subconstant error. *J. ACM*, 57(5):29:1–29:29, 2010. 8.1.2, 2

[MR17] Pasin Manurangsi and Prasad Raghavendra. A Birthday Repetition Theorem and Complexity of Approximating Dense CSPs. In *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017. 2.5.9

[MST06] Elchanan Mossel, Amir Shpilka, and Luca Trevisan. On ε-biased generators in NC0. *Random Structures & Algorithms*, 29(1):56–81, 2006. 6.2

[MW16] Ryuhei Mori and David Witmer. Lower Bounds for CSP Refutation by SDP Hierarchies. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2016, September 7-9, 2016, Paris, France*, volume 60 of *LIPIcs*, pages 41:1–41:30, 2016. 6.1

[Nil91] Alon Nilli. On the second eigenvalue of a graph. *Discrete Mathematics*, 91(2):207–210, 1991. 2.1

[NV08] Assaf Naor and Jacques Verstraëte. Parity check matrices and product representations of squares. *Combinatorica*, 28(2):163–185, 2008. 3.1.2

[OW14] Ryan O'Donnell and David Witmer. Goldreich's PRG: evidence for near-optimal polynomial stretch. In *2014 IEEE 29th Conference on Computational Complexity (CCC)*, pages 1–12. IEEE, 2014. 6.1

[OW20] Ryan O'Donnell and Xinyu Wu. Explicit near-fully X-Ramanujan graphs. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1045–1056. IEEE, 2020. 3.2.2, 10.2, 12, 12.0.1, 12

[Pal40] Gordon Pall. On the arithmetic of quaternions. *Transactions of the American Mathematical Society*, 47(3):487–500, 1940. 5

[Par00] Pablo A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, 2000. 2.5.1

[PD01] Alexander Prestel and Charles Delzell. *Positive Polynomials: From Hilbert's 17th Problem to Real Algebra*. Springer Science & Business Media, 2001. 2.5.1, 2.5.5

[Pip93] Nicholas Pippenger. Self-routing superconcentrators. In *Proceedings of the twenty-fifth annual ACM symposium on Theory of Computing*, pages 355–361, 1993. 10.1

[PK22] Pavel Panteleev and Gleb Kalachev. Asymptotically good quantum and locally testable classical LDPC codes. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 375–388, 2022. 10.1, 13.1.1

[RRS17] Prasad Raghavendra, Satish Rao, and Tselil Schramm. Strongly refuting random CSPs below the spectral threshold. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 121–131, 2017. 1.2, 6.1, 6.2, 7.0.2, 8

[RSV19] Nithi Rungtanapirom, Jakob Stix, and Alina Vdovina. Infinite series of quaternionic 1-vertex cube complexes, the doubling construction, and explicit cubical Ramanujan complexes. *International Journal of Algebra and Computation*, 29(06):951–1007, 2019. 10.2, 13.1.1, 13.3, 13.4

[RT12] Prasad Raghavendra and Ning Tan. Approximating CSPs with global cardinality constraints using SDP hierarchies. In *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*, pages 373–387. SIAM, 2012. 2.5.2, 2.5.7, 9.1.1

[RV17] Yuval Rabani and Rakesh Venkat. Approximating Sparsest Cut in Low Rank Graphs via Embeddings from Approximately Low Dimensional Spaces. *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, 2017. 6.3

[Sch04] Markus Schweighofer. On the complexity of Schmüdgen's Positivstellensatz. *Journal of Complexity*, 20(4):529–543, 2004. 2.5.1, 2.5.5

[SS96] Michael Sipser and Daniel Spielman. Expander Codes. *IEEE Trans. Inform. Theory*, 42(6, part 1):1710–1722, 1996. 10.1

[SS12] Warren Schudy and Maxim Sviridenko. Concentration and moment inequalities for polynomials of independent random variables. In *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*, pages 437–446. SIAM, 2012. 3.1.2

[ST11] Daniel A Spielman and Shang-Hua Teng. Spectral sparsification of graphs. *SIAM Journal on Computing*, 40(4):981–1025, 2011. 2.3, 8.1.4, 8.5

[SW19] Thatchaphol Saranurak and Di Wang. Expander decomposition and pruning: Faster, stronger, and simpler. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2616–2635. SIAM, 2019. 2.3, 8.1.4

[Tre08] Luca Trevisan. Approximation Algorithms for Unique Games. *Theory OF Computing*, 4:111–128, 2008. 1.2, 6.3

[Tro15] Joel A Tropp. An introduction to matrix concentration inequalities. *Foundations and Trends® in Machine Learning*, 8(1-2):1–230, 2015. 2.4.2

[TUZ07] Amnon Ta-Shma, Christopher Umans, and David Zuckerman. Lossless condensers, unbalanced expanders, and extractors. *Combinatorica*, 27:213–240, 2007. 10.1

[Vid13] Michael Viderman. Linear-time decoding of regular expander codes. *ACM Transactions on Computation Theory (TOCT)*, 5(3):1–25, 2013. 10.1

[WAM19] Alexander S Wein, Ahmed El Alaoui, and Cristopher Moore. The Kikuchi hierarchy and tensor PCA. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1446–1468. IEEE, 2019. 1, 1.1, 1.2, 3.1.2, 6.1

[Wig83] Avi Wigderson. Improving the performance guarantee for approximate graph coloring. *Journal of the ACM (JACM)*, 30(4):729–735, 1983. 6.3

[Wik22] Wikipedia contributors. Moore graph — Wikipedia, the free encyclopedia, 2022. [Online; accessed 12-July-2022]. 3.1

[Wul17] Christian Wulff-Nilsen. Fully-dynamic minimum spanning forest with improved worst-case update time. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1130–1143, 2017. 2.3, 8.1.4