Random Matrices and the Sum-of-Squares Hierarchy

Tselil Schramm

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2017-129
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-129.html

July 18, 2017

Random Matrices and the Sum-of-Squares Hierarchy

by

Tselil Schramm

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Associate Professor Prasad Raghavendra, Chair
Professor Satish Rao
Assistant Professor Nikhil Srivastava

Summer 2017

Random Matrices and the Sum-of-Squares Hierarchy

Abstract

Random Matrices and the Sum-of-Squares Hierarchy

by

Tselil Schramm

Doctor of Philosophy in Computer Science

University of California, Berkeley

Associate Professor Prasad Raghavendra, Chair

We study the Sum-of-Squares semidefinite programming hierarchy via the lens of average-case problems.

The Sum-of-Squares Hierarchy is a formulaic family of convex relaxations to polynomial optimization problems, which allows one to trade runtime for accuracy in a smooth manner. The Hierarchy has been studied since the early 2000's, both from the perspective of optimization and control and as a proof system. In the past five years, the Hierarchy has become a focus of intensive study in the theory of computation community. This is because recent results give us reason to hope that Sum-of-Squares algorithms may refute important conjectures on hardness of approximation. However, our understanding of the guarantees of the Hierarchy remains relatively incomplete.

In this dissertation, we present three results which make modest progress towards understanding the power and limitations of the Sum-of-Squares Hierarchy; all three works use average-case problems as a lens for the Sum-of-Squares algorithms, by enabling us to use random matrix theory as a tool in the analysis.

First, we analyze the performance of the Hierarchy for strongly refuting random constraint satisfaction problems (CSPs). We obtain a full characterization of the Sum-of-Squares Hierarchy for strong refutation of random CSPs, and give new subexponential-time strong refutation algorithms for CSPs with super-linear density.

Next, we give impossibility results for solving the planted clique problem via a Sum-of-Squares algorithm, demonstrating that the degree-4 Sum-of-Squares algorithm cannot distinguish graphs which contain a planted clique from uniformly random graphs.

Finally, even in the asymptotically polynomial-time regime, the Sum-of-Squares algorithm is often prohibitively slow. We show that for average-case problems, polynomial-time Sum-of-Squares algorithms can often be replaced with fast spectral algorithms, which run in linear or near-linear time in the input size.

To Mom and Dad: from life's first moments you welcomed my aspirations; you gave me courage to try and a soft place to fall.

And to my brother Pele: it's wonderful to laugh with someone who gets it.

Contents

Acknowledgments

Throughout my graduate studies, it has been my singular fortune to be surrounded and supported by mentors, collaborators, friends and family, and I offer my warmest thanks:

To my advisors, Prasad Raghavendra and Satish Rao. I remain in awe of their brilliant minds, their uncompromising aesthetic, and their loyalty to their personal identities outside of their work. I thank them both for their incredible generosity and kindness, and for their support, advice, and patience along my journey from novice to researcher.

To Elchanan Mossel, for advising me in my first year. I arrived at Berkeley green and confused—I thank Elchanan for guiding me through that that raw initial phase, and for his continuing friendship and genuine advice.

To my thesis and quals committee members, Nikhil Srivastava and Luca Trevisan. I thank them for serving on my committee. I thank them even more for serving as an inspiration to me throughout my studies, and I am greatful for all I have learned from them.

To Konstantin Makarychev, for hosting me as an intern at MSR Redmond during the summer of 2014. The internship was formative for me as a researcher, and I thank Kostya for that opportunity, and for our enjoyable collaboration that summer and in the years since.

To David Steurer, for his mentorship and for innumerable coffees.

To my collaborators and co-authors, Shuchi Chawla, Ronen Eldan, Sam Hopkins, Varun Kanade, Pravesh Kothari, Konstantin Makarychev, Elchanan Mossel, Aaron Potechin, Miki Rácz, Prasad Raghavendra, Satish Rao, Aviad Rubinstein, Jonathan Shi, David Steurer, Matt Weinberg, Benjamin Weitz, and Grigory Yaroslavtsev. I am extremely fortunate to have worked alongside such a creative and talented group of researchers, and I thank them all for their insight and their companionship in the pursuit of answers.

To the people that comprise the Berkeley theory group, who bring the discipline of Theory to life. I thank the theory group faculty for manning the helm with wisdom and style. In addition to those thanked above, I want to extend a special thanks to Umesh Vazirani for getting to the heart of the matter and for offering unabashed advice. My days were considerably sweetened by the Berkeley theory students (and by many of the Simons student visitors). I especially thank Ma'ayan Bresler, Anindya De, Kira Goldner, Fotis Iliopolus, Pasin Manurangsi, Alex Psomas, Miki Rácz, Aviad Rubinstein, Jarett Schwartz, Ning Tan, and Ben Weitz for their friendship and for making me laugh. I also thank Rishi Gupta, for changing my perspective; Jonah Brown-Cohen, for the contagious optimism and enthusiasm; and Sam Hopkins, for our moments spent together in the trenches.

To all my friends. Especially to Daniel, Olga, and Dabney for helping me look back; to Kiley, for the radiant warmth; to Alice, for the clarity and stability; to Evan, David, Coline, and Andrea, for creating the home that I lived in.

To Tynan, my partner, friend, and comrade—thank you for sharing in my highs and my lows, and for letting me share in yours.

And to my mother Avivit, my father Oded, and my brother and friend Pele; being born into our family has been the biggest privilege of all.

Chapter 1

Introduction

Often, the design of algorithms is an ad hoc endeavor: first a specific computational problem is formulated, and then an algorithm is tailor-made to solve the problem. This customized approach is natural because it allows for the exploitation of problem-specific structure. Driven by the goal of solving a particular algorithmic problem, one studies the problem at length, and then produces an algorithm based on this study.

Yet from a theoretical computer scientist's perspective, a menagerie of algorithms is not enough. As theorists, we wish not only to produce algorithms, but a *theory* of algorithms. We want to understand why our algorithms work. We want algorithm design to be based on broad principles, rather than guesswork and intuition. We want to classify problems based on the sorts of algorithms that can solve them, instead of thinking of algorithms as mere problem-solving tools.

From this desire arises an alternative approach to algorithms research: start by proposing an algorithm, and then discover which problems it can solve.

1.1 Sum-of-Squares Algorithms

This dissertation is concerned with understanding the power and limitations of a particular family of algorithms: the *Sum-of-Squares (SoS) Hierarchy*. Given any polynomial optimization problem, the SoS Hierarchy automatically generates a family of algorithms with increasing power (and computational demands), completely mechanizing the algorithms design process. In order to set the stage for the description of SoS algorithms, we'll begin by introducing the concepts of polynomial optimization and convex relaxations.

Polynomial Optimization

In polynomial optimization, we have some polynomial f that we would like to optimize over some region $C \subset \mathbb{R}^n$, where C is defined by polynomial constraints such as $g(x) = 0$ for some set of polynomials $\{g\}$ and $h(x) \geq 0$ for some set of polynomials $\{h\}$.

Polynomial optimization problems can be used to capture many combinatorial optimization problems. Consider for example MAX CUT:

Problem 1.1.1 (MAX CUT)**.** Suppose we are given a graph $G = (V, E)$. Find a bipartition of the vertices V that maximizes the number of edges whose endpoints are in different partitions.

In other words, we want to find a cut in the graph that maximizes the number of cut edges. We can express MAX CUT as a polynomial optimization problem:

Program 1.1.2 (Polynomial optimization formulation for MAX CUT)**.** Index the vertices of G by $[n]$ for $n = |V|$, and the edges $e \in E$ by pairs $(i, j) \in \binom{[n]}{2}$ corresponding to the endpoints of e.

$$\max_{x \in \mathbb{R}^n} \quad \sum_{(i,j) \in E} \frac{1}{2}(1 - x_i x_j) \tag{1.1.1}$$

$$s.t. \quad x_i^2 = 1 \quad \forall i \in [n]. \tag{1.1.2}$$

The intent is to assign every vertex $i \in [n]$ to either the $+1$ partition or the -1 partition. The constraints (1.1.2) ensure that the feasible $x \in \mathbb{R}^n$ for the program are assignments $x \in \{\pm 1\}^n$, since for $x_i \in \mathbb{R}$, the only solutions to $x_i^2 = 1$ are $x_i = \pm 1$. Since $x_i x_j \in \{\pm 1\}$ for feasible x, each term in the objective function (1.1.1) contributes 1 if the edge (i, j) is cut, and 0 otherwise.

For any combinatorial optimization problem, one can apply a similar transformation to obtain a corresponding polynomial optimization problem. However, the combinatorial optimization problems we are interested in are often **NP**-hard (MAX CUT is **NP**-hard [Kar10]), and so solving the polynomial optimization problem exactly is also **NP**-hard.

Convex relaxations

Suppose we have some polynomial optimization problem $\mathcal{P} = \min_{x \in C} f(x)$. Instead of solving $\mathcal{P}$ exactly, we can *relax* the problem to a convex optimization problem $\mathcal{Q}$. Given a convex body $K \subset \mathbb{R}^N$ with poly(N) constraints, there are efficient algorithms for minimizing a convex function f over K.[1] The idea of convex relaxation is that, rather than solving an **NP**-hard polynomial optimization problem $\mathcal{P}$, we identify some convex set $K \subset \mathbb{R}^N$ that contains representatives $R(x)$ all of the points $x \in C$ (the feasible region of $\mathcal{P}$), and a convex objective function $\tilde{f}$ that matches the value of the polynomial f on the representatives of c, so that $\tilde{f}(R(x)) = f(x)$ for $x \in C$.

In this way, the convex optimization problem $\min_{y \in K} \tilde{f}(y)$ is a *relaxation* of the polynomial optimization problem $\min_{x \in C} f(x)$, as its value can only be smaller than the value of $\mathcal{P}$,

$$\text{val}(\mathcal{Q}) = \min_{y \in K} \tilde{f}(y) \leq \min_{x \in C} f(x) = \text{val}(\mathcal{P}).$$

By solving the convex $\mathcal{Q}$, we can efficiently obtain a lower bound on the minimum of $\mathcal{P}$. A good relaxation $\mathcal{Q}$ has value as close to val$(\mathcal{P})$ as possible.

[1] This assumes that K and f have reasonable bit complexity so that they can be expressed with a polynomial number of bits, and that K has an efficient separation oracle, but we will state these assumptions precisely in Chapter 2.

Sum-of-Squares Relaxations

The SoS Hierarchy is a family of *convex relaxations* for polynomial optimization problems, independently first proposed by [Par00, Las01, Sho87]. For each even $d \in \mathbb{N}$, the *degree-d, n-variate SoS relaxation* is a convex relaxation of size $n^{O(d)}$. The formulation of these relaxations is simple—the degree-d SoS relaxation introduces a variable for each monomial of degree at most d, and then introduces affine constraints in these variables to mimic the polynomial constraints, as well as an eigenvalue constraint.[2]

For example, consider the degree-2 SoS relaxation for MAX CUT:

Example 1.1.3 (Degree-2 SoS relaxation for MAX CUT)**.** In the degree-2 SoS relaxation for MAX CUT, we introduce variables $X_\emptyset, \{X_{\{i\}}\}_{i\in[n]}, \{X_{\{i,j\}}\}_{i,j\in[n]}$, where X_S is a stand-in for the monomial $\prod_{i\in S} x_i$. We then write the program

$$\begin{aligned} \max \quad & \sum_{(i,j)\in E} \frac{1}{2}(1 - X_{ij}) && (1.1.3)\\ s.t. \quad & X_{ii} = 1 \quad \forall i \in [m], && (1.1.4)\\ & X_\emptyset = 1, && (1.1.5)\\ & X \succeq 0\,. && (1.1.6) \end{aligned}$$

The constraints (1.1.3) and (1.1.4) are affine analogs of the polynomial objective function $\sum_{(i,j)\in E} \frac{1}{2}(1 - x_i x_j)$ and the polynomial constraints $\{x_i^2 = 1\}_{i\in[n]}$. The constraint (1.1.5) is a normalization. In the final constraint (1.1.6), we define the matrix X to be the matrix with rows and columns indexed by $\emptyset \cup [n]$, so that for $S, T \in \emptyset \cup [n]$ the entry $X(S,T)$ is given by $X_{S\cup T}$. The constraint $X \succeq 0$ requires that X is positive-semidefinite; that is, that all eigenvalues of X are non-negative.

One can see that this is a relaxation, since for any solution $x \in \mathbb{R}^n$, the point $X_S = \prod_{i\in S} x_i$ is feasible (the matrix $[1\ x]^\top [1\ x]$ is also positive-semidefinite) and attains the same value as the polynomial program Program 1.1.2.

At degree-2, the SoS program for MAX CUT is identical to the Goemans-Williamson [GW94] relaxation. For SoS relaxations of degree $d > 2$, we simply add monomials for the variables of higher degrees, and then enforce the implied higher-degree constraints (for example, at SoS degree 4 we enforce the affine analogue of the constraint $(x_i^2-1)(x_j^2-1) = 0$). As the degree d of the SoS relaxation grows, more constraints are added, and so increasing the SoS degree clearly only strengthens the convex relaxation. One of our primary lines of inquiry will be: how much does increasing the SoS degree d strengthen the relaxation?

For now, we will content ourselves with the example of MAX CUT, leaving more formal, general definitions for Chapter 2.

[2] The eigenvalue constraint ensures that the relaxation treats squares of degree-$d/2$ polynomials as non-negative functions. We will treat this more thoroughly in Chapter 2.

Duality and SoS Proofs

Above, we have introduced the degree-d SoS relaxation in the *primal* view, in which we view SoS variables as relaxations of degree-d monomials. The degree-d SoS relaxation is a semidefinite program (SDP). If we instead consider the dual of the SoS SDP, we obtain another natural optimization problem, from which the Sum-of-Squares hierarchy derives its name.

Suppose that we would like to certify that a degree-k polynomial $f(x)$ in variables $x \in \mathbb{R}^n$ has value at least c over the domain defined by the constraints $C = \{g_i(x) = 0\}_{i \in [m]}$ for degree-k polynomials $g_1, \ldots, g_m$. One form that such a proof could take is a polynomial identity of the form

$$f(x) - c = \sum_{j=1}^{\ell} s_j(x)^2 + \sum_{i=1}^{m} h_i(x) \cdot g_i(x),$$

where $s_1, \ldots, s_\ell$ are real polynomials of degree at most $d/2$, and $h_1, \ldots, h_m$ are arbitrary real polynomials with the property that $\deg(h_i) + \deg(g_i) \leq d$. Since a sum of squares of polynomials cannot be negative over the reals, and since for any $x \in C$, the latter term on the right-hand-side is zero, this identity certifies that for any $x \in C$, $f(x) - c \geq 0$.

Given the formulation of the SoS hierarchy and standard duality arguments, it is not difficult to see that the dual of the degree-d SoS relaxation finds the best such polynomial identity of degree d.[3] That is, it finds such an identity with polynomials $s_1, \ldots, s_\ell$ and $g_1 \cdot h_1, \ldots, g_m \cdot h_m$ of degree at most d such that the lower bound c is as large as possible. This is a *degree-d Sum-of-Squares proof* that $f(x) \geq c$ on C. Here, it is even easier to see that as the degree d is allowed to grow, the tightness of the certifiable lower bound improves.

1.2 Why Study SoS Algorithms?

The SoS Hierarchy was first formulated independently by Parrilo, Lasserre, and Shor [Par00, Las01, Sho87] as a means for studying polynomial optimization. One reason that this study was initiated was purely mathematical inquiry: given a polynomial optimization problem $\mathcal{P}$, can it be captured by a convex relaxation of finite size? In his paper introducing the SoS hierarchy, Lasserre [Las01] showed that as the degree d of the relaxation is taken to infinity, $\mathcal{P}$ can be approximated arbitrarily well, and that in many cases (for example, if the domain of $\mathcal{P}$ is the hypercube), a finite d suffices.

Another reason was more practical, out of a desire to solve actual polynomial optimization problems. In control theory, one often wishes to solve multivariate partial differential inequalities; for example, the Lyapunov function for a nonlinear system. Applying the SoS relaxation to these polynomial optimization problems can sometimes give good solutions. For this reason, SoS algorithms have been widely studied in the control theory community since the effort was initiated by Parrilo in his PhD thesis [Par00, PP02, BMH12].

The degree-d SoS relaxation for an optimization problem with n variables has size $n^{O(d)}$, and for problems of interest can usually be optimized in time $n^{O(d)}$ (see [RW17, Wei17]). For applications in control, the optimization problem is often continuous, and the number

[3] We will argue this in Chapter 2.

of variables is often small, $n \leq 5$. In this setting, taking $d = \Omega(n)$ produces a large convex program, but one that is nevertheless possible to optimize with fast machines.

As theoretical computer scientists, we are interested in a different setting. Often, we have a combinatorial optimization problem $\mathcal{P}$ over variables $x \in \{0,1\}^n$, and we would like to design an algorithm to solve $\mathcal{P}$ that is as efficient as possible as a function of n, as $n \to \infty$. Having a convex relaxation of size $2^{\Omega(n)}$ which faithfully captures $\mathcal{P}$ is not useful for us, since optimizing such a relaxation requires essentially as much time as checking each point in $\{0,1\}^n$ by force.

Approximation Algorithms and the Unique Games Conjecture

In fact, the SoS Hierarchy has been studied extensively by theoretical computer scientists, for the special case when the degree $d = 2$. When $d = 2$, the SoS relaxation is a simple Semidefinite Program (SDP). In the early 90's, the seminal work of Goemans and Williamson showed that semidefinite programming could be used to obtain a better-than-random approximation to MAX CUT [GW94].

In the decades since, semidefinite programming has become the scaffolding for the theory of approximation algorithms, and researchers have developed an arsenal of tools for analyzing the performance of semidefinite programs for problems of varied structures. Semidefinite programs give the best known algorithms for constraint satisfaction problems [Rag08, RS09a], and for geometric problems such as SPARSEST CUT [ARV09]. SDPs have also been crucial in the discovery of new algorithms for domains in which only existential results were known from pure mathematical fields, such as discrepancy theory [Ban10]. And vice-versa, SDPs have been used to establish novel mathematical results, such as improving the upper bound on the Grothendieck constant [BMMN11]. At the same time, the limitations of a family of basic semidefinite programs have been used to build a beautiful (though still incomplete) theory of hardness of approximation around Khot's Unique Games Conjecture [Kho02, MOO05, KKMO04, Rag08].

While higher levels of the hierarchy, with $d \geq 4$, have been studied from the perspective of polynomial optimization [Par00] and proof complexity [Gri01b] for many years, the computational resources required by the Sum-of-Squares programs make them completely intractable in practice, and so historically the Sum-of-Squares Hierarchy has received little attention from the algorithms community. Five years ago, in a surprising work Barak et al. [BBH+12] showed that the degree-8 Sum-of-Squares Hierarchy can efficiently solve integrality gap instances of the Unique Games problem that have eluded other linear and semidefinite programs, establishing it as our most promising avenue for refuting the Unique Games Conjecture.

This discovery spurred a flurry of activity around the Sum-of-Squares Hierarchy, with progress made in both lower and upper bounds for a variety of problems, especially for average-case problems (e.g. [BBH+12, BKS15, BHK+16]).

A New Perspective on Spectral Algorithms

While studying SoS algorithms in pursuit of disproving the Unique Games Conjecture, researchers have uncovered new algorithms for a number of problems. One particularly fruitful

insight has been the discovery of a new family of *spectral algorithms* that are naturally implied by the analysis of the SoS relaxations.

A *spectral algorithm* is any algorithm in which a matrix is assembled from the input, after which the spectrum (or eigenvalues and eigenvectors) of the matrix are used to deduce information about the input. Spectral algorithms are an old and important algorithmic primitive, with applications to clustering [NJW01], computer vision [SM00], and learning latent variable models [HKZ12], to name just a few. This family of methods contains algorithms simple enough to be taught in undergraduate algorithms, such as Principle Components Analysis, as well as much more sophisticated graph partitioning methods on the bleeding edge of algorithmic research [ABS15].

One common tool in the study of SoS algorithms is primal-dual analysis; in order to bound the value of the polynomial optimization problem, one considers the dual formulation of the SoS program. Since the SoS primal and dual are both semidefinite programs, the primal and dual both optimize over the space of matrices. Because of the connection between the value of the SoS relaxation and the top eigenvalues of feasible dual matrices, often, the analysis of SoS algorithms boils down to characterizing the spectral information of a well-chosen dual matrix.

This in turn means that the analysis used to bound the SoS relaxation value can often be translated into a stand-alone efficient spectral algorithm! For algorithms design, this connection is one of the more salient contributions to come out of the study of SoS relaxations thus far. We will return to this theme throughout this document.

1.3 Average-Case Problems and Random Matrix Theory as a Lens

Though the Sum-of-Squares semidefinite programs are easy to formulate, as a community we are still far from understanding their guarantees. Our best insights into the Hierarchy have come from examining the performance of the SoS algorithms for *average-case problems* and *planted problems*.

Average-Case Problems

In an average-case problem, problem instances come from some previously specified distribution, and the goal is to design an algorithm that performs well for a typical instance from this distribution, rather than for *every* instance.

To illustrate these concepts, let us take for example the problem of finding a clique in a graph. The MAX CLIQUE problem is one of the most classical algorithmic problems in graph theory: given a graph G, the goal is to find the largest subset of vertices in which every pair of vertices share an edge. MAX CLIQUE is **NP**-hard to approximate within any polynomial factor [Hås96]. For this reason, it is natural to consider average-case versions of the problem, and ask whether they are still computationally intractable.

The following natural average-case variant was proposed by Karp in the 1970's [Kar76].

Problem 1.3.1 (Average-Case MAX CLIQUE)**.** Given a uniformly random graph, or equivalently, a graph from the Erdős-Rényi distribution $G(n, \frac{1}{2})$ in which every edge is included independently with probability $\frac{1}{2}$, can one find the largest clique?

A heuristic argument shows that with high probability over the choice of graph $G \sim G(n, \frac{1}{2})$, a greedy algorithm can find a clique of size $\approx \log n$ in polynomial time. On the other hand, a careful application of the second moment method can show that the largest clique in a graph from $G(n, \frac{1}{2})$, has size $2(1 \pm o(1)) \log n$ with high probability [GM75, Mat76, BE76].[4] Thus, while MAX CLIQUE is hard to approximate within any polynomial factor, the average-case version yields an easy 2-approximation. But no polynomial time algorithm is known for identifying cliques of size $(1 + \varepsilon) \log n$ for any constant $\varepsilon > 0$ in this average-case setting.

Planted Problems

In a planted (maximization) problem, we are given a quasi-random instance, in which a solution with large objective value has been artificially planted within a random instance.

Perhaps the most famous planted problem is the planted version of MAX CLIQUE.

Definition 1.3.2 ((n, ω)-PLANTED CLIQUE distribution)**.** The (n, ω)-PLANTED CLIQUE distribution is a distribution over n-vertex graphs. A uniform sample is generated by first sampling an Erdős-Rényi graph $G \sim G(n, \frac{1}{2})$, then a subset $\mathcal{S}$ of ω of the n vertices are chosen uniformly at random, and a clique is "planted" on $\mathcal{S}$ by including every edge (i, j) with both endpoints in $\mathcal{S}$.

Algorithmic Tasks for Average-Case Problems

For average case and planted problems, there are several different algorithmic questions. There are *search* problems, *distinguishing* problems, and *refutation* problems.

Definition 1.3.3 (Search problem)**.** Given an instance sampled from a planted distribution, find the planted solution.

For example, in the search version of (n, ω)-PLANTED CLIQUE, we wish to find the subset of ω vertices on which the clique has been planted.

Definition 1.3.4 (Distinguishing problem)**.** In a distinguishing problem, there are two distributions over instances: a *random distribution* ν and a *planted distribution* μ. Given a sample $\mathcal{I}$, we are asked to determine whether it was sampled from ν or μ with probability better than a random guess.

For example, in the distinguishing version of the (n, ω)-PLANTED CLIQUE problem, the *random* distribution is the Erdős-Rényi distribution $G(n, \frac{1}{2})$, the *planted* distribution is the (n, ω)-PLANTED CLIQUE distribution, and given a sample graph G, we want to decide which of the distributions it was sampled from.

[4] This gives an easy quasi-polynomial time algorithm for finding the largest clique—simply check every one of the $\binom{n}{2 \log n}$ subsets of vertices to determine whether it is a clique.

Because the maximum clique in $G \sim G(n, \frac{1}{2})$ has size $2 \log n$ with high probability, when $\omega \geq (2 + \varepsilon) \log n$, the task becomes information-theoretically possible.

Finally, there is the refutation problem:

Definition 1.3.5 (Refutation problem)**.** Given an optimization problem $\mathcal{P}$ and an average-case instance $\mathcal{I}$ sampled from a distribution $\mathcal{D}$, provide a certificate that $\mathcal{I}$ does not have a solution of value larger than ω, where ω is at least the expected value of $\mathcal{P}$ over $\mathcal{D}$.

For example, for the $\mathcal{G}(n, \frac{1}{2})$ distribution, one can ask for a refutation that instances sampled from $\mathcal{G}(n, \frac{1}{2})$ have no clique of size larger than $\omega \gg 2 \log n$.

Observation 1.3.6. We notice that the hardest algorithmic question is the search question, as a search algorithm can be used to solve both the distinguishing and refutation problems. Refutation algorithms can be used to solve the distinguishing problem, and so distinguishing is the easiest task. Nonetheless, when we are asking for convex-programming based algorithms the three notions are often roughly equivalent, and so we will often conflate the three.

We will return to a discussion of PLANTED CLIQUE later, in Chapter 4.

Average-Case Problems, Random Matrices, and the SoS Hierarchy

Average-case and planted problems are interesting in their own right, as an alternative to the more pessimistic worst-case analysis. But in studying a fixed, structured algorithm such as the SoS Hierarchy, they can be an invaluable tool for understanding the power and limitations of the algorithm.

The SoS SDP optimizes over matrices. In the case of SoS algorithms, average-case problems are particularly helpful, because they allow us to use tools from random matrix theory to analyze the performance of the algorithms.

To understand the limitations of the SoS SDP, we want to establish lower bounds—that is, we want to find problems, and instances of those problems, for which the SDP gives a poor approximation to the value of the instance. This usually amounts to demonstrating that for an instance $\mathcal{I}$ of the problem, there is matrix which is feasible for the SoS relaxation but not for the original polynomial optimization problem, and which gives a large objective value. For an average-case problem, if we create a mapping from problem instances to feasible SoS matrices with large objective values, we can use random matrix theory to argue that the matrices satisfy the SoS constraints, and in particular the positive-semidefiniteness constraint, with high probability.

On the other hand, in order to prove algorithmic results using SoS, one can again take advantage of average-case problems using primal-dual analysis. In order to analyze the degree-d SoS relaxation $\mathcal{Q}$ for an instance of a planted problem $\mathcal{P}$, our primal-dual analysis goes as follows: we carefully choose a feasible point X for the dual of $\mathcal{Q}$. Because $\mathcal{Q}$ and its dual are programs over symmetric $n^{O(d)} \times n^{O(d)}$ matrices, X itself is a symmetric matrix, and standard facts from the theory of convex duality imply that the maximum eigenvalue of X yields an upper bound on $\mathrm{opt}(\mathcal{P})$. Furthermore, by definition of the dual, X is a sum of the objective function and constraint matrices for $\mathcal{P}$. That is to say, the entries of X are functions of the input problem.

Now, we would show the following: if $\mathcal{P}$ is an instance drawn from the random distribution, then with high probability $\lambda_{\max}(X)$ is less than some value θ, and so $\text{obj}(\mathcal{Q}) \leq \theta$. Otherwise, if $\mathcal{P}$ comes from the planted distribution, then $\text{obj}(\mathcal{P}) \gg \theta$, and therefore $\lambda_{\max}(X) \geq \theta$ because $\mathcal{Q}$ is a relaxation for $\mathcal{P}$. This form of primal-dual analysis shows that the SoS relaxation $\mathcal{Q}$ distinguishes between planted and random instances. The fact that X is a random matrix in the random case often makes analyzing its eigenvalues tractable, in contrast to the dual matrices in worst-case problems.

1.4 Results

In this thesis, we will explore three instances in which average-case problems shed light on the power and limitations of the Sum-of-Squares Hierarchy. In each of these instances, we will see how average-case problems, by enabling us to use tools from random matrix theory, allow us to exploit the structured nature of SoS algorithms to make precise and general statements about their guarantees.

Strongly Refuting Random Constraint Satisfaction Problems

Random constraint satisfaction problems (CSPs) are perhaps the most canonical example of an average-case problem. Random CSPs have been studied deeply across several mathematical communities, including theoretical computer science, probability and statistical physics.

At the same time, the polynomial optimization formulation for maximizing CSPs is incredibly simple—for a Boolean CSP over n variables with predicates on k variables, the objective function is a degree-k polynomial, and the only constraints are constraints of the form $\{x_i^2 = 1\}_{i \in [n]}$. The simplicity of this polynomial formulation, as well as the wealth of literature concerned with random CSPs, make random CSPs an ideal starting point for studying SoS algorithms.

Our first result pertains to the *refutation* of random CSPs. To precisely explain the problem of refutation, we first introduce the problem of random 3-XOR.

Definition 1.4.1 (Random 3-XOR with density α)**.** A *random* 3-*XOR* instance Φ on n variables with density α is generated as follows: sample $m = \alpha n$ uniformly random triples $(i_1, i_2, i_3) \in [n]^3$ and add the constraint that $x_{i_1} x_{i_2} x_{i_3} = b_i$ for a uniformly random sign b_i.

It is known[5] that random CSPs, such as 3-XOR, exhibit threshold phenomena: there exists some constant α_s such that random 3-XOR instances Φ sampled at density $\alpha < \alpha_s$ are satisfiable with high probability, and Φ sampled at density $\alpha > \alpha_s$ are unsatisfiable with high probability. Below the satisfiability threshold, when $\alpha < \alpha_s$, the natural algorithmic question is finding a satisfying assignment. However, when $\alpha > \alpha_s$, instances are unsatisfiable with high probability, and the natural algorithmic problem is *refutation*—the task of proving that there are no satisfying assignments.

[5]This has been empirically verified for many CSPs, but has been rigorously established only in a few cases—we will give more details in Chapter 3.

Definition 1.4.2 (Refutation)**.** An algorithm $\mathcal{A}$ is a *refutation* algorithm for random 3-XOR if given a random instance Φ with density $\alpha > \alpha_s$, the algorithm $\mathcal{A}$:

- Outputs SAT if Φ is satisfiable.
- Outputs UNSAT with probability at least 0.9 over the choice of Φ.

Note that if the algorithm $\mathcal{A}$ outputs UNSAT on an instance Φ, it certifies or proves that the instance Φ is unsatisfiable.

Refutation is a well-studied problem with connections to myriad areas of theoretical computer science including proof complexity [BB02], inapproximability [Fei02], SAT solvers, cryptography [ABW10a], learning theory [DLS14b], statistical physics [CLP02] and complexity theory [BKS13]. We survey the prior work on refuting CSPs in Chapter 3.

At densities far above the satisfiability threshold $\alpha \gg \alpha_s$, a simple union bound argument can be used to show that a random 3-XOR instance Φ has no assignment satisfying more than a $\frac{1}{2} + o(1)$ fraction of constraints. In this regime, it is natural to ask for a *strong refutation*:

Definition 1.4.3 (Strong Refutation)**.** An algorithm $\mathcal{A}$ is a *strong refutation* algorithm for random 3-XOR if for a fixed constant $\delta > 0$, given a random instance Φ with density $\alpha \gg \alpha_s$, the algorithm $\mathcal{A}$:

- Outputs SAT if Φ has an assignment satisfying at least a $(1-\delta)$-fraction of clauses.
- Outputs UNSAT with probability at least 0.9 over the choice of Φ.

For the 3-XOR predicate, there is an efficient algorithm for refutation: since a 3-XOR instance is a system of linear equations over $\mathbb{F}_2$, one can simply perform Gaussian elimination to find a satisfying assignment, if one exists. Strong refutation, on the other hand, is a different matter. A natural spectral algorithm can efficiently strongly refute k-XOR at densities $m/n \geq n^{k/2-1}$ [CGL07, CGL07, AOW15, BM16]. However, strong refutation at any lower density is widely believed to be an intractable problem [ABW10a, BM16, DLS14a, Dan16]. We refer the reader to [Dan16] for a survey of the evidence pointing to the intractability of the problem.

The Sum-of-Squares proof system is a natural proof system for algorithmic refutations and strong refutations of CSPs. As alluded to above, given an instance Φ of a Boolean k-CSP, the fraction of constraints satisfied by an assignment x can be written as a polynomial $P_\Phi(x)$ of degree at most k in x, and the Booleanness constraints can be written as $\{x_i^2 = 1\}_{i\in[n]}$. Let $\mathrm{opt}(\Phi)$ denote the largest fraction of constraints satisfied by any assignment to the variables. The SoS-hierarchy offers a natural family of algorithms for producing algorithmic strong refutations: the primal degree-d SoS relaxation $\mathcal{Q}_d$ has value $\mathrm{opt}(\mathcal{Q}_d) \geq \mathrm{opt}(\Phi)$, and the degree-$d$ SoS dual provides a Sum-of-Squares proof of size at most $n^{O(d)}$ that $\mathrm{opt}(\Phi) \leq \mathrm{opt}(\mathcal{Q}_d)$.

At the same time, as a lens for understanding SoS relaxations, random CSPs offer all the benefits of an average-case problem, coupled with the simplicity of the CSP SDP constraints. By examining this problem with simple structure and a rich supporting literature, we can hope to learn how increasing the degree of the SoS relaxation can help to obtain upper bounds. In this exploration we obtain the following theorem, which gives new strong refutation algorithms in the subexponential regime for any Boolean CSP.

Theorem 1.4.4. *Let $P : \{\pm 1\}^k \to \{0,1\}$ be a predicate with expected value $\mathbb{E}[P]$ over a random assignment in $\{\pm 1\}^k$. For all $\delta \in (0,1]$, given an instance Φ of a random k-CSP with predicate P on n variables, the degree $O(n^\delta)$ SoS hierarchy strongly refutes Φ with high probability, certifying that*

$$\mathrm{opt}(\Phi) \leq \mathbb{E}[P] + \varepsilon\,,$$

for any constant $\varepsilon > 0$ so long as Φ has density at least $m/n \geq \widetilde{O}(n^{(k/2-1)(1-\delta)})$, where the $\widetilde{O}$ hides a dependence on a polylog factor, k and ε. Furthermore, there is a spectral algorithm achieving the same guarantees.

Notice that the result establishes a smooth trade-off between the clause density of Φ and the running time of the refutation algorithm. Specifically for all $\delta \in [0,1)$, the algorithm strongly refutes at density $m/n = \widetilde{O}(n^{(k/2-1)(1-\delta)})$ in time $\exp(\widetilde{O}(n^\delta))$, so that when $\delta = 0$ the result matches the performance of the best known polynomial-time algorithms, and at $\delta = 1$, the algorithm refutes instances just above the threshold of satisfiability in exponential time.

Further, this result is tight within the SoS framework—the lower bounds of Grigoriev [Gri01b], Schoenebeck [Sch08], and Kothari-Mori-O'Donnell-Witmer[6] [KMOW17] rule out strong SoS refutations with a better SoS-degree/density tradeoff.

Implications for Sum-of-Squares Algorithms

Theorem 1.4.4 was among the first results to give a tight characterization of the performance of the SoS algorithms for any degree d; previous works had focused on the regime where $d = O(1)$, or $d = \Theta(n)$. The primary technical contribution of this work was to understand how the basic SoS constraints, in combination with only the bare-bones constraints $\{x_i^2 = 1\}$, can be used to obtain better SoS and spectral algorithms when d grows as a function of n.

Lower Bounds for Planted Clique

We recall the (n,ω)-PLANTED CLIQUE distinguishing problem defined above (Section 1.3)—this is the problem of distinguishing graphs drawn from $G(n,\frac{1}{2})$ and graphs drawn from $G(n,\frac{1}{2})$ in which a clique of size ω has been planted. As ω increases, the distance between the two distribution grows, and the problem becomes easier. For $\omega \geq \sqrt{n}$, there is a polynomial-time spectral algorithm which solves PLANTED CLIQUE [AKS98]. But for $\omega < \sqrt{n}$, no algorithm which improves on the performance of this brute-force is known.

In lieu of algorithmic progress, PLANTED CLIQUE has become a benchmark for new algorithmic techniques, and there has been a lot of success in proving impossibility results for breaking the $\omega = \sqrt{n}$ barrier in polynomial time. Today, we have impossibility results (lower bounds) in several algorithmic frameworks: Markov Chain Monte Carlo [Jer92], the Lovász-Schrijver SDP hierarchy, [FK00, FK03], and "statistical algorithms" [FGR+12], to name a few.

[6] We can extend Theorem 1.4.4 so that the density/runtime trade-off depends on the *independence parameter* of the predicate P as defined by Allen, O'Donnell, and Witmer [AOW15], giving a better runtime/density tradeoff for some CSPs. These tradeoffs are tight with respect to the lower bounds of Kothari-Mori-O'Donnell and Witmer. We defer the details to Section 3.5.

Some of the initial successes of the algorithms community in studying the SoS hierarchy were new algorithms for average-case problems (e.g. [BBH+12, BKS15, HSS15]). For this reason, there was a focus in the community on understanding the performance of SoS on PLANTED CLIQUE. At degree-2, impossibility results for the SoS algorithm are already implied by [FK00, FK03]; however, the same results do not apply to degree larger than 2, and researchers hoped to break the $\omega = \sqrt{n}$ barrier with some constant degree d.

However, lower bounds came more readily than new upper bounds. First, the work of [MPW15] showed a $\omega = \tilde{\Omega}(n^{1/d})$-lower bound for the degree-d SoS SDP, by demonstrating a mapping from $G \sim G(n, \frac{1}{2})$ which with high probability yields a feasible SDP solution of large value. Then, the work of [DM15b] gave a tighter analysis of this same mapping to show a $\omega = \tilde{\Omega}(n^{1/3})$ lower bound, but only for $d = 4$; a counterexample of Kelner (which may be found in [Bar14]) demonstrates that the analysis of [DM15b] is tight for their construction within logarithmic factors. At this juncture, it was still unclear whether the degree-4 SoS relaxation could solve the (n, ω)-PLANTED CLIQUE problem for $\omega \ll \sqrt{n}$. We proved that this is in fact not the case:

Theorem 1.4.5. *Suppose that $G \sim \mathbb{G}(n, \frac{1}{2})$. Then with probability $1 - O(n^{-4})$, there exists a feasible solution to the SoS-SDP of degree $d = 4$ with objective value $\frac{\sqrt{n}}{\operatorname{polylog} n}$.*

Implications for Sum-of-Squares Algorithms

When our result was proven, it was unclear whether the degree-4 SoS relaxation could give improved results for PLANTED CLIQUE, and the opinions of experts were divided on this subject. Our result resolved this question. Since then, the work of [BHK+16] proved that the $(n, n^{1/2-\delta})$-PLANTED CLIQUE problem requires at least a degree-$d = O(\delta) \log n$ relaxation to distinguish planted and random instances.

Fast Spectral Algorithms from Sum-of-Squares Analyses

As mentioned above, many of the cases in which we currently know that SoS algorithms give good guarantees are *average-case* and *planted* problems. For instance, SoS has lead to advances in algorithms for planted sparse vector [BKS14], tensor completion [BM16], tensor principal components analysis [HSS15], tensor decomposition [BKS15, GM15], and dictionary learning [BKS15]. For a wide range of parameters of these problems, degree-d SoS achieves significantly stronger guarantees than other methods, for constant or logarithmic degree d.

Unfortunately, the computational cost grows rather steeply in terms of the parameter d: the running time is $n^{O(d)}$ where n is the number of variables (usually comparable to the instance size). Further, even when the SDP has size polynomial in the input (when $d = O(1)$), solving the underlying semidefinite programs is prohibitively slow for large instances.

As discussed above, for a planted problem $\mathcal{P}$ the analysis of these SoS algorithms usually proceeds through a spectral analysis of a well-chosen dual matrix X. We show that $\lambda_{\max}(X)$ is large in the planted case and small with high probability in the random case. This naturally yields a *spectral algorithm*—if one can write the dual point X in time n^d given the input

problem, then evaluating $\lambda_{\max}(X)$ gives an n^d-time spectral algorithm for solving the planted problem $\mathcal{P}$.

However, when n^d is significantly larger than the input size, this is still computationally intractable. In the following result, we introduce spectral algorithms for planted sparse vector and tensor decomposition that exploit the same high-degree information as the corresponding Sum-of-Squares algorithms without relying on semidefinite programming, and achieve the same (or close to the same) guarantees. That is, we show how to take the matrices used in the SoS primal-dual analysis for these problems and *compress* them into smaller matrices that have compact *factorizations*, which allows us to implement the spectral algorithms in near-linear time. The resulting algorithms are quite simple (a couple of lines of MATLAB code) and have considerably faster running times—quasi-linear or close to linear in the input size.

We first present our result for the planted sparse vector problem, which arose as a primitive in the dictionary learning problem [SWW12] and has since become a problem of mild interest in its own right (we give more background in Chapter 5).

Theorem 1.4.6 (Near-linear time algorithm for planted sparse vector)**.** *Suppose we are given an arbitrary orthogonal basis of a subspace V spanned by $v_0, \ldots, v_{d-1} \in \mathbb{R}^n$, where $v_1, \ldots, v_{d-1}$ are sampled independently from $\mathcal{N}(0, \mathrm{Id})$ and v_0 is a vector with at most εn nonzero entries for some $1/100 > \varepsilon > 0$.*

Then if $d \leq \sqrt{n}/\operatorname{polylog} n$, there is a spectral algorithm running in time $\widetilde{O}(nd)$ which recovers a unit vector u correlated with the sparse vector v_0, such that $\langle u, \frac{v_0}{\|v_0\|}\rangle^2 \geq 1 - O(\varepsilon^{1/4}) - o_n(1)$.

We remark that the parameter requirement we obtain here, that $d \leq \widetilde{O}(\sqrt{n})$, is within logarithmic factors of that obtained by the corresponding Sum-of-Squares algorithm [BKS14]. The best SoS algorithm solves an SDP over matrices of size $n^2 \times n^2$, while our runtime is linear in the input.

Tensor decomposition is an important algorithmic primitive in many unsupervised learning tasks [AGH+15]. Although the problem is **NP**-hard in general, under some assumptions it is known to be tractable. We consider the average-case variant for order-3 tensors in $(\mathbb{R}^d)^{\otimes 3}$, in which the tensor has rank $d^{1+\delta}$ for a constant $\delta \leq \frac{1}{3}$, and the components of the tensor are i.i.d. samples from $\mathcal{N}(0, \mathrm{Id})$. In this setting, we obtain subquadratic algorithms for tensor decomposition.

Theorem 1.4.7 (Fast random tensor decomposition)**.** *Suppose we are given an order-3 tensor $\mathbf{T}$ with n random components $a_1, \ldots, a_n \sim \mathcal{N}(0, \frac{1}{d}\,\mathrm{Id}_d)$ in $\mathbb{R}^d$, so that $\mathbf{T} = \sum_i a_i^{\otimes 3}$.*

Then so long as $d \leq n \leq n^{4/3}/\log^{O(1)} n$, there exists a randomized algorithm that finds a vector b which is close to a component of $\mathbf{T}$, so that for some $i \in [n]$ $\langle b, a_i\rangle \geq 1 - \eta$ for $\eta \leq \widetilde{O}(n^3/d^4)^{1/2}$ in time $\widetilde{O}(d^{1+\omega})$ (where ω is the matrix multiplication constant). Moreover, $\widetilde{O}(n)$ iterations will recover all of the a_i with high probability.

Our algorithms come close to the best parameter tradeoffs achieved by the SoS algorithms of [GM15] and [MSS16], with significantly faster runtimes.

Implications for Sum-of-Squares Algorithms

Our work demonstrates that polynomial-time Sum-of-Squares algorithms, even with degree parameter $d \geq 4$, can still yield valuable insights into devising truly efficient algorithms in the average case. This motivates the further study of SoS algorithms; while the SoS algorithm itself is inherently inefficient (even when it runs in polynomial time asymptotically), we show that it is sometimes possible to use insights from the SoS algorithm's analysis to obtain fast algorithms.

This approach has already been extended and applied to the SoS tensor completion algorithm of [BM16] in [MS16], and for the tensor decomposition results of [MSS16] in a different parameter regime [SS17].

1.5 Organization

In Chapter 2, we provide notation and technical preliminaries. Chapter 3 contains our results refuting random constraint satisfaction problems. In Chapter 4 we give our degree-4 lower bound for the planted clique problem. Finally, Chapter 5 describes our methodology for obtaining fast spectral algorithms from SoS analyses. Some additional technical underpinnings are contained in Appendix A.

Chapter 2

Preliminaries

2.1 Notation and Conventions

Indexing. We will use the shorthand $[n] = \{1, \ldots, n\}$. We will also use the notation $[n]^k$ to denote the set of all multisets of k elements of $[n]$, and the notation $[n]^{\leq k}$ to denote the set of all multisets of at most k elements of $[n]$. By $\binom{[n]}{k}$ (or $\binom{[n]}{\leq k}$), we denote the set of all sets of (up to) k elements of $[n]$, without repetitions.

Polynomials. We will often be concerned with maximizing or minimizing real multivariate polynomials; in these cases, we will most often denote with x the vector of n variables. We will sometime represent monomials with the notation $x^S = \prod_{i \in S} x_i$.

Norms and Inner Products. For two vectors $u, v \in \mathbb{R}^n$, we let $\langle u, v \rangle = \sum_{i \in [n]} u_i v_i$. Similarly for two matrices $A, B \in \mathbb{R}^{n \times m}$, we let $\langle A, B \rangle = \mathrm{Tr}(AB^\top)$. The ℓ_p-norm of a vector $v \in \mathbb{R}^n$ is denoted by $\|v\|_p$. For matrices, the default norm will be the operator norm, $\|A\| := \|A\|_{op}$. We will also use the Frobenius norm, $\|A\|_F := Tr(AA^\top)^{1/2}$.

Linear Algebra. We use $\succ$ ($\succeq$) to denote positive (semi)definiteness, so that for an $n \times n$ symmetric real matrix A and for a vector $x \in \mathbb{R}^n$, $A \succ 0$ ($A \succeq 0$) if and only if $x^\top Ax > 0$ ($x^\top Ax \geq 0$) for all $x \in \mathbb{R}^n$. We also use $\succeq$ to denote the PSD or Loewner ordering on matrices, so that $A \succeq B$ if and only if $A - B \succeq 0$.

A vector of indeterminates may be denoted $x = (x_1, \ldots, x_n)$, although we may sometimes switch to parenthetical notation for indexing, i.e. $x = (x(1), \ldots, x(n))$ when subscripts are already in use. We denote by $[n]$ the set of all valid indices for a vector in $\mathbb{R}^n$. Let e_i be the ith canonical basis vector so that $e_i(i) = 1$ and $e_i(j) = 0$ for $j \neq i$.

For a vectors space V, we may denote by $\mathcal{L}(V)$ the space of linear operators from V to V. The space orthogonal to a vector v is denoted $v^\perp$.

For a matrix M, we use M^{-1} to denote its inverse or its Moore-Penrose pseudoinverse; which one it is will be clear from context. For M PSD, we write $M^{-1/2}$ for the unique PSD matrix with $(M^{-1/2})^2 = M^{-1}$.

Tensors. We represent tensors by boldface letters such as $\mathbf{T}$. We refer to the map from a tensor $\mathbf{T}$ with entries indexed by $[n]^{\otimes 2k}$ to a matrix T indexed by $[n]^k \times [n]^k$ as the "natural flattening" of $\mathbf{T}$. For a matrix or vector $M \in \mathbb{R}^{n \times m}$, the notation $M^{\otimes d}$ refers both to the $n^d \times m^d$ d-wise Kronecker power of M, or to the $n \times \cdots \times m$ tensor given by the d-wise tensor product of M with itself.

For an order-3 tensor in $(\mathbb{R}^n)^{\otimes 3}$, we denote by $\mathbf{T}(x, y, z)$ the multilinear function in $x, y, z \in \mathbb{R}^n$ such that $\mathbf{T}(x, y, z) = \sum_{i,j,k \in [n]} T_{i,j,k} x_i y_j z_k$, applying x, y, and z to the first, second, and third modes of the tensor $\mathbf{T}$ respectively. If the arguments are matrices P, Q, and R instead, this lifts $\mathbf{T}(P, Q, R)$ to the unique multilinear tensor-valued function such that $[\mathbf{T}(P, Q, R)](x, y, z) = \mathbf{T}(Px, Qy, Rz)$ for all vectors x, y, z.

Probability. We will often refer to collections of independent and identically distributed (or *iid*) random variables. The Gaussian distribution with mean μ and variance σ^2 is denoted $\mathcal{N}(\mu, \sigma^2)$. We will also use $\mathcal{N}(M, \Sigma)$ to denote the multivariate Gaussian distribution with mean M and covariance matrix Σ. Sometimes we state that an event happens with overwhelming probability. This means that its probability is at least $1 - n^{-\omega(1)}$.

Asymptotic Bounds. We will use standard O-notation; we will use $\widetilde{O}(\cdot)$ to suppress polylogarithmic factors. We will sometimes abuse notation by taking $\widetilde{O}(1)$ to denote $O(\operatorname{polylog} n)$, and we hope this will be clear from context.

2.2 Optimization and Semidefinite Programming

Suppose we have some set of points S, and we wish to optimize some objective function f^{obj} over S. *Optimization problems* of this form will be our primary subject of study.

Definition 2.2.1. An *optimization* problem consists of an *objective function* $f^{\text{obj}} : \mathbb{R}^n \to \mathbb{R}$ and a closed feasible set $S \subseteq \mathbb{R}^n$. In the *decision version*, we are given some target $c \in \mathbb{R}$, and we are asked to determine whether

$$\min_{x \in S} f^{\text{obj}}(x) \leq c.$$

In the *search version*, we are asked to find the element of S minimizing f^{obj},

$$\operatorname*{argmin}_{x \in S} f^{\text{obj}}(x).$$

Remark 2.2.2. If one can solve the decision version, then one can also determine the value $\min_{x \in S} f^{\text{obj}}(x)$ by applying binary search to the target c.

Examples of optimization problems that we will consider in this manuscript include maximizing constraint satisfaction problems such as 3-XOR and finding maximum cliques.

Example 2.2.3 (MAX 3-XOR)**.** Given a 3-XOR instance on n variables with equations $\{x_{i_1}x_{i_2}x_{i_3} = b_i\}_{i\in[m]}$ for index tuples $(i_1, i_2, i_3) \in [n]^3$ and signs $b_i \in \{\pm 1\}$, the MAX 3-XOR optimization problem asks us to find the assignment $x \in \{\pm 1\}^n$ that maximizes the number of satisfied equations,

$$\max_{x\in\{\pm\}^n} \sum_{i\in[m]} \frac{1}{2}\left(1 + b_i \cdot x_{i_1}x_{i_2}x_{i_3}\right) .$$

Here we are minimizing $f^{\text{obj}} = -\sum_{i\in[m]} \frac{1}{2}\left(1 + b_i \cdot x_{i_1}x_{i_2}x_{i_3}\right)$, over the set $S = \{\pm 1\}^n \subset \mathbb{R}^n$. It is not difficult to verify that for $x \in \{\pm 1\}^n$ and $b \in \{\pm 1\}^m$, each term in the summation is 1 if $x_{i_1}x_{i_2}x_{i_3} = b_i$, and 0 otherwise.

Example 2.2.4 (MAX CLIQUE)**.** Given a graph $G = (V, E)$ on n vertices indexed by $[n]$, the MAX CLIQUE optimization problem asks us to find the largest subset of vertices $K \subseteq V$ such that K is a clique in G. Letting $x_K \in \mathbb{R}^n$ be the 0/1 indicator of $K \subseteq [n]$, here $f^{\text{obj}}(x) = -\|x_K\|_1$, and $S \subset \mathbb{R}^n$ is the set of all 0/1 indicators of the cliques of G.

Convex Relaxations

In the case where f^{obj} is a convex function and $S \subseteq \mathbb{R}^n$ is also convex, the *Ellipsoid Algorithm* can solve the *search version* of an optimization problem efficiently.

Theorem 2.2.5 (Efficient Optimization over Convex Domains (see e.g. [PS82]).)**.** *Suppose $S \subset \mathbb{R}^n$ is a closed convex set and f^{obj} is a convex function, and suppose we have an upper bound of M on $\min_{x\in S} f^{\text{obj}}(x) \le M$. Let V be the volume of the convex set $K \stackrel{\text{def}}{=} S \cup \{x \in \mathbb{R}^n \mid f^{\text{obj}}(x) \le M\}$, and suppose that K is either empty or contains the ball $\mathcal{B}(c, r)$ for some $r > 0$ and some $c \in \mathbb{R}^n$. Let T be the runtime of a separation oracle for K.*

Then given an error parameter ε and access to a separation oracle for K, if K is non-empty then the Ellipsoid Algorithm *finds $x^* \in S$ such that*

$$\left| f^{\text{obj}}(x^*) - \min_{x\in S} f^{\text{obj}}(x) \right| \le \varepsilon$$

in time poly $\left(\log\left(\frac{1}{\varepsilon}\right), \log(V), M, T\right)$.

This is a classical theorem in convex optimization, and its proof may be found in many optimization textbooks (for example [PS82]). The subtler requirements of the theorem (such as K containing a ball of radius r) can often be made to hold for problems of interest (with some gentle massaging).

However, often we are interested in solving optimization problems over *non-convex* domains S, as we have seen above in the examples MAX 3-XOR ($S = \{\pm 1\}^n$) and MAX CLIQUE ($S = \{0, 1\}^n$). Many of these problems are **NP**-Hard, and so we cannot hope for efficient algorithms which solve them exactly. This motivates the concept of *convex relaxations*.

Definition 2.2.6 (Convex Relaxation)**.** Suppose we have the optimization problem $\mathcal{P} = (f^{\text{obj}}, S)$, in which we wish to minimize the function $f^{\text{obj}}(x)$ over the domain $S \subseteq \mathbb{R}^n$. A *convex relaxation* for $\mathcal{P}$ is an optimization problem $\mathcal{Q} = (f^{\text{rel}}, S^{\text{rel}})$ with objective function f^{rel} over domain $S^{\text{rel}} \subseteq \mathbb{R}^N$, so that there exists a map $R : \mathbb{R}^n \to \mathbb{R}^N$ such that

- The set S^{rel} contains the image of S under R. That is,

$$\forall x \in S,\ R(x) \in S^{\text{rel}}.$$

- The function f^{rel} is identical to the function f^{obj} over $R(S)$. That is,

$$\forall x \in S,\ f^{\text{rel}}(R(x)) = f^{\text{obj}}(x).$$

Fact 2.2.7. *For a minimization problem $\mathcal{P} = (f^{\text{obj}}, S)$ and its convex relaxation $\mathcal{Q} = (f^{rel}, S^{rel})$, we have that*

$$\text{opt}(\mathcal{Q}) \leq \text{opt}(\mathcal{P}).$$

Proof. This is because, if $x^* \in S$ is the minimizer of f^{obj}, then by definition of a relaxation,

$$\text{opt}(\mathcal{P}) = f^{\text{obj}}(x^*) = f^{\text{rel}}(R(x^*)) \geq \min_{y \in S^{\text{rel}}} f^{\text{rel}}(y) = \text{opt}(\mathcal{Q}).$$

□

If we are interested in knowing $\text{opt}(\mathcal{P})$ for a non-convex **NP**-hard optimization problem $\mathcal{P}$, we have a compelling algorithmic alternative: we can instead design a *convex*, easy-to-optimize relaxation $\mathcal{Q}$, and solve $\mathcal{Q}$ to obtain a lower bound on $\text{opt}(\mathcal{P})$. Of course, we may have $\text{opt}(\mathcal{Q}) \ll \text{opt}(\mathcal{P})$, in which case solving $\mathcal{Q}$ does not give us very much information. We measure the quality of a convex relaxation by the distance between $\text{opt}(\mathcal{P})$ and $\text{opt}(\mathcal{Q})$. This distance is called the *integrality gap*.

Definition 2.2.8 (Integrality gap)**.** Suppose we have a minimization problem $\mathcal{P}$ and a convex relaxation $\mathcal{Q}$ of $\mathcal{P}$. The *integrality gap* of $\mathcal{Q}$ is defined to be the smallest real number that bounds the ratio of the optima of $\mathcal{Q}$ and $\mathcal{P}$:

$$\inf\{\alpha \geq 1 \mid \text{opt}(\mathcal{P}) \leq \alpha \cdot \text{opt}(\mathcal{Q})\}.$$

For a family of optimization problems $\{\mathcal{P}_I\}$ and their convex relaxations $\{\mathcal{Q}_I\}$, sometimes we also refer to an *α-integrality gap instance* or a *gap instance*—this is a particular instance $\mathcal{P}_I$ of the optimization problem for which the optimum of the corresponding relaxation $\mathcal{Q}_I$ is smaller by a factor of $1/\alpha$.

Semidefinite Programs

Of particular interest to us is a category of convex programs known as *semidefinite programs*, or SDPs. Semidefinite programs allow us to optimize a linear objective over a domain defined by linear constraints, as well as matrix eigenvalue constraints. Defined more precisely:

Definition 2.2.9 (Semidefinite Program (primal))**.** A *semidefinite program* is a convex optimization program over symmetric matrices $X \in \mathbb{R}^{n\times n}$, of the form

$$\begin{aligned} \max_{X\in\mathbb{R}^{n\times n}} \quad & \langle C, X\rangle \\ s.t. \quad & \langle A_i, X\rangle = b_i \quad \forall i \in [m] \\ & X \succeq 0, \end{aligned}$$

where C and $\{A_i\}_{i\in[m]}$ are $n \times n$ real matrices, and $b \in \mathbb{R}^m$.

One can also incorporate linear inequality constraints; see [BV04] for a reference. The Lagrangian dual of a SDP is also an SDP:

Proposition 2.2.10 (Semidefinite Program (dual))**.** *The dual of the SDP in Definition 2.2.9 is a convex optimization program over scalars $y \in \mathbb{R}^m$, of the form*

$$\begin{aligned} \min_{y\in\mathbb{R}^m} \quad & \langle b, y\rangle \\ s.t. \quad & C \preceq \sum_{i=1}^{m} y_i A_i. \end{aligned}$$

See [BV04] for a proof.

2.3 The Sum-of-Squares SDP Hierarchy: Polynomial Optimization

Suppose we wish to maximize an n-variate polynomial $f^{\text{obj}}(x)$ over $x \in S \subset \mathbb{R}^n$, where S is some subset of $\mathbb{R}^n$ defined by polynomial constraints. This problem is clearly **NP**-hard in general, as it captures **NP**-hard problems such as MAX CLIQUE and 3-SAT.

The Sum-of-Squares (SoS) hierarchy is a sequence of increasingly powerful (and computationally intensive) SDP relaxations for polynomial optimization problems. The idea is as follows: for each even positive number d, define the degree-d SoS relaxation for a polynomial optimization problem $\mathcal{P}$ by introducing a variable for each monomial of degree at most d. Then, replace every polynomial constraint in $\mathcal{P}$ with a linear constraint in the variables, and include a positive-semidefiniteness constraint (which mimics the positive-semidefiniteness of true evaluations of the monomials). Formally,

Definition 2.3.1 (Degree-d SoS relaxation)**.** Suppose we are given a polynomial minimization problem $\mathcal{P} = (f^{\text{obj}}, S)$ over $x \in \mathbb{R}^n$, with a feasible region S defined by the constraints $\{g_i(x) = 0\}_{i\in[m]}$, and let k be an upper bound on the degrees of f^{obj} and g_i for all $i \in [m]$.

For any even $d \geq k$, we define the degree-d SoS relaxation to be the following SDP over matrices $X \in \mathbb{R}^{N\times N}$, where $N = (n+1)^{d/2}$. We think of the rows and columns of X as being indexed by multisets $S \in [n]^{\leq d/2}$, and we identify the matrix entry $X_{S,T}$ with the monomial $x^S \cdot x^T$ (recalling that we have defined $x^S = \prod_{i\in S} x_i$). Our constraints are as follows:

- Scaling/Normalization: we think of our SDP variable $X_\emptyset$ as the monomial 1, because in general we want $X_{S\cup T}$ to represent $x^S \cdot x^T$, which implies that we want $X_{\emptyset\cup S}$ to represent $x^\emptyset \cdot x^S$. Thus we fix the scale appropriately.

$$X_\emptyset = 1 \tag{2.3.1}$$

- Commutativity/Symmetry: for any list of indices S and any permutation $\pi(S)$, we have that $x^S = x^{\pi(S)}$. We add a corresponding constraint to our variables,

$$X_{A,B} = X_{C,D} \qquad \forall A, B, C, D \in [n]^{\leq d/2} \quad \text{s.t.} \quad (A, B) = (C, D) \text{ as multisets}. \tag{2.3.2}$$

- Positive-semidefiniteness: the matrix X is a relaxation of the rank-1 positive-semidefinite matrix $\left([1\ x]^{\otimes d/2}\right)^\top ([1\ x]^{\otimes d/2}) \succeq 0$. We thus enforce the constraint that

$$X \succeq 0. \tag{2.3.3}$$

- Constraints from $\mathcal{P}$: in $\mathcal{P}$, we have the polynomial constraints that $g_i(x) = 0$, which in turn imply the polynomial constraint that $x^S \cdot g_i(x) = 0$ for any monomial x^S. The polynomial $x^S \cdot g_i$ is a linear function of the monomials, and so we enforce the constraint that

$$x^S \cdot g_i(X) = 0 \qquad \forall S \subseteq [n] \quad \text{s.t.} \quad \deg(x^S \cdot g_i) \leq d, \tag{2.3.4}$$

  where we have abused notation and used $x^S \cdot g_i(X)$ to denote the linear function $\sum_{T\subseteq[n]^{\leq d}} g_i^T \cdot X_{S\cup T}$, where g_i^T is the coefficient of x^T in g_i.

- Objective function: We take the linear objective function

$$\min f^{\text{obj}}(X), \tag{2.3.5}$$

  where again we define $f^{\text{obj}}(X) = \sum_{T\subseteq[n]^{\leq d}} (f^{obj})^T \cdot X_T$.

Remark 2.3.2. As a consequence of the Ellipsoid Algorithm, we may conclude that (ignoring issues of bit complexity, which sometimes do arise [O'D16] but are irrelevant for most problems of interest [RW17]) the degree-d SoS relaxation can be solved in $n^{O(d)}$ time.

In the above definition, we have pointed out that the degree-d SoS program is indeed a relaxation, as when we apply the map $R(x) = [1\ x]^\top [1\ x]$ to any $x \in \mathbb{R}^n$, $R(x)$ is a feasible point for the relaxation with the same objective value as $f^{\text{obj}}(x)$.

Fact 2.3.3. *A useful alternate characterization of (2.3.3) is that for any polynomial p of degree at most $d/2$, we have that $p^2(X) \geq 0$, where $p^2(X)$ is the function given by evaluating the coefficients of p^2 at the stand-in monomials given by the variables of the program.*

Proof. This can be seen by taking the vector $\hat{p}$ whose Sth entry is given by the coefficient of x^S in p, and noticing that $0 \leq \hat{p}^\top X \hat{p} = p^2(X)$. □

Definition 2.3.4 (Pseudoexpectation operator)**.** Given a particular degree-d SoS program, we define the corresponding *pseudoexpectation* operator to be the linear operator $\tilde{\mathbb{E}} : \mathbb{R}[x]^{\leq 2d} \to \mathbb{R}$, which maps any monomial of degree at most $2d$ to the SoS variable identified with it.

It is sometimes instructive to think of the variable X_S as a *pseudoexpectation* or a *pseudomoment* of the monomial $\prod_{i \in S} x_i$ over feasible solutions which minimize the objective function:

$$X_S = \underset{x \text{ minimizing } f_{obj}}{\tilde{\mathbb{E}}} \left[\prod_{i \in S} x_i \right].$$

In some sense, the constraints of the SDP force the solution to behave like the moments of a probability distribution over feasible maximizing solutions, although they needn't correspond to the moments of a true distribution, hence the term *pseudomoment*. See e.g. [Bar14] for more discussion.

2.4 The Sum-of-Squares SDP Hierarchy: Sum-of-Squares Proofs

Given a polynomial minimization problem $\mathcal{P} = (f^{\text{obj}}, S)$, if we solve the degree-d SoS relaxation for $\mathcal{P}$, we have a lower bound c on the value of f^{obj} over S. If we take the dual of the degree-d SoS program, we get an interesting object: a Sum-of-Squares proof that $f^{\text{obj}} \geq c$ over S.

Definition 2.4.1 (Sum-of-Squares proof)**.** Given a polynomial $f(x)$ in $x \in \mathbb{R}^n$, a set S defined by polynomial equalities $\{g_i(x) = 0\}_{i \in [m]}$, and a $c \in \mathbb{R}$ such that $f(x) \geq c$ for all $x \in \mathbb{R}^n$, a *Sum-of-Squares* proof is a polynomial identity of the form

$$f(x) - c = \sum_{j \in [N]} p_j^2(x) + \sum_{i \in [m]} q_i(x) \cdot g_i(x),$$

where $\{p_j\}_{j \in [N]}$ and $\{q_i\}_{i \in [m]}$ are polynomials in $x \in \mathbb{R}^n$.

A *degree-d Sum-of-Squares proof* is one in which all of the polynomials $q_i \cdot g_i$ and p_j^2 have degree at most d.

Proposition 2.4.2. *For a polynomial optimization problem $\mathcal{P}$, the dual of the degree-d SoS SDP,* $\text{dual}(\mathcal{Q})$*, gives a degree-d SoS proof that* $\text{opt}(\mathcal{P}) \geq \text{opt}(\text{dual}(\mathcal{Q}))$.

Proof. The degree-d SoS SDP is an SDP of the form

$$\begin{aligned}
\max \quad & \langle -F, X \rangle \quad (\equiv \min \quad \langle F, X \rangle) \\
s.t. \quad & \langle \mathbb{I}_{\emptyset,\emptyset}, X \rangle = 1 \\
& \langle \mathbb{I}_{S,T} - \mathbb{I}_{U,V}, X \rangle = 0 \quad \forall S, T, U, V \in [n]^{\leq d/2} \text{ s.t. } (S,T) = (U,V) \text{ as multisets} \\
& \langle G_i^T, X \rangle = 0 \quad \forall i \in [m],\ T \in [n]^{\leq d - \deg(g_i)}
\end{aligned}$$

$$X \succeq 0$$

where F is a matrix representation of f^{obj} and G_i^T is a matrix representation of $x^T g_i$ and $\mathbb{I}_{S,T}$ is the $N \times N$ matrix with a single entry in the S,Tth coordinate.

By Proposition 2.2.10, the dual SDP of Definition 2.3.1 has the form

$$\begin{aligned} \min_{c,q,y} \quad & -c \quad (\equiv \quad \max c) \\ s.t. \quad & -F \preceq -c \cdot \mathbb{I}_{\emptyset,\emptyset} + \sum_i \sum_{T \in [n]^{\leq d - \deg(g_i)}} q_{i,T} \cdot G_i^T + \sum_{\substack{S,T,U,V \in [n]^{\leq d/2} \\ (S,T)=(U,V)}} y_{S,T,U,V} \cdot (\mathbb{I}_{S,T} - \mathbb{I}_{U,V}). \end{aligned}$$

We can re-express the above inequality as an equality by introducing a PSD slack matrix $S \succeq 0$.

$$S - F = -c \cdot \mathbb{I}_{\emptyset,\emptyset} + \sum_i \sum_{T \in [n]^{\leq d - \deg(g_i)}} q_{i,T} \cdot G_i^T + \sum_{\substack{S,T,U,V \in [n]^{\leq d/2} \\ (S,T)=(U,V)}} y_{S,T,U,V} \cdot (\mathbb{I}_{S,T} - \mathbb{I}_{U,V})$$

Now, we transform the matrix identity into a polynomial identity by taking the inner product of the left and right and sides with the variable matrix $\hat{x}^{\otimes d}$ (where $\hat{x} = [1 \; x]^\top$):

$$\begin{aligned} &\langle S, \hat{x}^{\otimes d} \rangle - \langle F, \hat{x}^{\otimes d} \rangle \\ &\quad = -c \cdot \langle \mathbb{I}_{\emptyset,\emptyset}, \hat{x}^{\otimes d} \rangle + \sum_{\substack{i,T \\ i \in [m] \\ T \in [n]^{\leq d - \deg(g_i)}}} \langle q_{i,T} \cdot G_i^T, \hat{x}^{\otimes d} \rangle + \sum_{\substack{S,T,U,V \in [n]^{\leq d/2} \\ (S,T)=(U,V)}} y_{S,T,U,V} \cdot \langle \mathbb{I}_{S,T} - \mathbb{I}_{U,V}, \hat{x}^{\otimes d} \rangle \end{aligned}$$

By definition of the SoS relaxation, we have that $\langle F, \hat{x}^{\otimes d} \rangle = f^{\text{obj}}(x)$, and that the same holds for the G_i^T. We also use that $\langle \hat{x}^{\otimes d}, \mathbb{I}_{\emptyset,\emptyset} \rangle = 1$ and that $\langle \hat{x}^{\otimes d}, \mathbb{I}_{S,T} - \mathbb{I}_{U,V} \rangle = 0$ when $(S,T) = (U,V)$ as multisets. Simplifying, we have

$$\langle S, \hat{x}^{\otimes d} \rangle - f^{\text{obj}}(x) = -c + \sum_{\substack{i,T \\ i \in [m] \\ T \in [n]^{\leq d - \deg(g_i)}}} q_{i,T} \cdot x^T \cdot g_i(x)$$

Now, for each $i \in [m]$ we can define the degree-$(d - \deg(g_i))$ polynomial $q_i(x) = \sum_{T \in [n]^{d - \deg(g_i)}} q_i^T x^T$, so that we can further simplify,

$$\langle S, \hat{x}^{\otimes d} \rangle - f^{\text{obj}}(x) = -c + \sum_{i \in [m]} q_i(x) \cdot g_i(x)$$

Finally, we use that for $S \succeq 0$, we can always write $S = PP^\top$ for some matrix P. If we take the polynomial $p_j(x)$ to be the degree-$d/2$ polynomial defined by the form $\langle p_j, \hat{x}^{d/2} \rangle$ for p_j the jth column of S, we have that

$$\sum_j p_j(x)^2 - f^{\text{obj}}(x) = -c + \sum_{i \in [m]} q_i(x) \cdot g_i(x) \,.$$

Rearranging,

$$f^{\text{obj}}(x) - c = \sum_j p_j(x)^2 - \sum_{i \in [m]} q_i(x) \cdot g_i(x) \,.$$

which is a SoS proof that $f^{\text{obj}}(x) \geq c$ for x in the feasible region of $\mathcal{P}$. □

Often, when we formulate an SoS relaxation for a polynomial optimization problem, we will use the dual formulation to bound its integrality gap. That is, if we wish to minimize the polynomial $f^{\text{obj}}(x)$, to lower bound the value of the degree-d SoS relaxation for f^{obj}, we can use any degree-d SoS proof as a dual certificate.

Corollary 2.4.3. *If we have a degree-d Sum-of-Squares proof that polynomial minimization problem $\mathcal{P}$ has $\text{opt}(\mathcal{P}) \geq c$, then the degree-$d$ SoS relaxation $\mathcal{Q}$ for $\mathcal{P}$ also has $\text{opt}(\mathcal{Q}) \geq c$.*

Proof. Since any degree-d SoS proof is a feasible dual solution for the dual of $\mathcal{Q}$, we must have $\text{opt}(\mathcal{Q}) \geq \text{opt}(\text{dual}(\mathcal{Q})) = c^* \geq c$. □

In the next section, we will introduce some useful low-degree some-of-squares proofs that will be valuable to us in bounding SoS relaxations.

2.5 Common Sum-of-Squares Proofs and Feasible Dual Points

In this section, we list some useful low-degree Sum-of-Squares identities and provide their proofs. Statements of additional useful lemmas, and their proofs, can be found in the appendix of [BKS14].

The following lemma is the basis of the connection between Sum-of-Squares relaxations and spectral algorithms.

Lemma 2.5.1 (SoS matrix inner product)**.** *Let M be an $[n]^d \times [n]^d$ matrix. Then there is a degree-$2d$ Sum-of-Squares proof of the fact that*

$$\langle x^{\otimes 2d}, M \rangle \leq \|x\|^{2d} \cdot \lambda_{\max}(M).$$

Proof. For convenience let $\lambda \stackrel{\text{def}}{=} \lambda_{\max}(M)$. By definition $\lambda \cdot I - M \succeq 0$, and therefore the expression $I - M$ can be written as a Sum-of-Squares proof of degree at most $2d$. We thus have a degree-$2d$ Sum-of-Squares proof that

$$\langle x^{\otimes 2d}, M \rangle \;\leq\; \langle x^{\otimes 2d}, \lambda \cdot I - M \rangle + \langle x^{\otimes 2d}, M \rangle \;=\; \lambda \cdot \|x\|^{2d},$$

as desired. □

Many standard inequalities, such as the Cauchy-Schwarz inequality and Hölder's inequality, have low-degree Sum-of-Squares proofs, or are true for low-degree pseudodistributions.

Lemma 2.5.2 (SoS Cauchy-Schwarz inequality)**.** *Let p, q be polynomials of degree at most d in x, and let $\tilde{\mathbb{E}}$ be a degree-$2d$ pseudodistibution over $x \in \mathbb{R}^n$. Then*

$$\tilde{\mathbb{E}}[p(x)q(x)] \leq \sqrt{\tilde{\mathbb{E}}[p^2(x)]\,\tilde{\mathbb{E}}[q^2(x)]}$$

Proof. If $\tilde{\mathbb{E}}[p(x)^2] = 0$ or $\tilde{\mathbb{E}}[q(x)^2] = 0$, the left-hand side is also zero—to see this, we think of the PSD matrix of pseudomoments of $\tilde{\mathbb{E}}$, and realize that this implies that the vector of coefficients of p (respectively q) is in its null space.

Assuming that both numbers are strictly positive, we take $\widehat{p}(x) = \frac{p(x)}{\sqrt{\tilde{\mathbb{E}}[p(x)^2]}}$, and $\widehat{q}(x) = \frac{q(x)}{\sqrt{\tilde{\mathbb{E}}[q(x)^2]}}$—this does not increase the degree of the proof, as the quantities $\tilde{\mathbb{E}}[p(x)^2]$ and $\tilde{\mathbb{E}}[q(x)^2]$ are fixed positive scalars. Since $(\hat{p}(x) - \hat{q}(x))^2$ is a square of a degree-d polynomial, we have a degree-$2d$ SoS proof that

$$\widehat{p}(x)\widehat{q}(x) \leq \frac{1}{2}\left(\widehat{p}^2(x) + \widehat{q}^2(x)\right),$$

which implies that

$$\frac{1}{\sqrt{\tilde{\mathbb{E}}[p^2(x)]\,\tilde{\mathbb{E}}[q^2(x)]}} \cdot \tilde{\mathbb{E}}[p(x)q(x)] \;=\; \tilde{\mathbb{E}}[\widehat{p}(x)\widehat{q}(x)] \leq \frac{1}{2}\left(\tilde{\mathbb{E}}[\widehat{p}^2(x)] + \tilde{\mathbb{E}}[\widehat{q}^2(x)]\right) \;=\; 1,$$

since $\tilde{\mathbb{E}}$ is a linear operator which is non-negative for squares of polynomials of degree at most $2d$. Multiplying both sides by the scalar $\sqrt{\tilde{\mathbb{E}}[p^2(x)]\,\tilde{\mathbb{E}}[q^2(x)]}$ gives the desired conclusion. □

The following fact will be useful in relating the objective value of SoS programs to the higher-degree pseudomoments.

Fact 2.5.3 (SoS-Convexity)**.** *For any k which is an integer power of 2, if $p(x)$ is a polynomial of degree at most d, and $\tilde{\mathbb{E}}$ is a degree-$2kd$ pseudoexpectation operator over $x \in \mathbb{R}^n$, then*

$$\tilde{\mathbb{E}}[p(x)]^{2k} \leq \tilde{\mathbb{E}}[p(x)^{2k}].$$

Proof. We prove this by induction on k. When $k = 1$, this is equivalent to the SoS Cauchy-Schwarz (Lemma 2.5.2). Now, assuming this is true for k, we prove it for $2k$. Since $\tilde{\mathbb{E}}$ is a degree-$4dk$ pseudoexpectation by assumption, we have that

$$0 \leq \tilde{\mathbb{E}}\left[\left(p(x)^{2k} - \tilde{\mathbb{E}}[p(x)^{2k}]\right)^2\right]$$
$$\tilde{\mathbb{E}}_m\, u\left[p(x)^{2k}\right]^2 \leq \tilde{\mathbb{E}}\left[p(x)^{4k}\right].$$

Applying the induction hypothesis, the conclusion follows. □

Chapter 3

Strong Refutation of Constraint Satisfaction Problems

3.1 Introduction

Random instances of constraint satisfaction problems (CSPs) have been a subject of intense study in computer science, mathematics and statistical physics. Even if we restrict our attention to random k-SAT, there is already a vast body of work across various communities– see [Ach09] for a survey. In this chapter, our focus is on *refuting* random CSPs: the task of algorithmically proving that a random instance of a CSP is unsatisfiable. Refutation is a well-studied problem with connections to myriad areas of theoretical computer science including proof complexity [BB02], inapproximability [Fei02], SAT solvers, cryptography [ABW10a], learning theory [DLS14a], statistical physics [CLP02] and complexity theory [BKS13].

For the sake of concreteness, we will for a moment restrict our attention to k-SAT, the most well-studied random CSP. In the random k-SAT model, we choose a k-uniform CNF formula Φ over n variables by drawing m clauses independently and uniformly at random. The density of Φ is given by the ratio $\alpha = m/n$. It is conjectured that for each k, there is a critical value α_k such that Φ is satisfiable with high probability if $\alpha < \alpha_k$, and unsatisfiable with high probability for $\alpha > \alpha_k$. Such phase transition phenomena are conjectured to occur for all nontrivial random CSPs; for the specific case case of k-SAT, it was only recently rigorously established for all sufficiently large k [DSS15].

In the unsatisfiable regime, when $\alpha > \alpha_k$, the natural algorithmic problem we associate with random k-SAT formulas is the problem of *refutation*. We define the notion of a refutation algorithm formally:

Definition 3.1.1. (Refutation Algorithm) An algorithm $\mathcal{A}$ is a *refutation* algorithm for random k-SAT at density α, if given a random instance Φ of k-SAT with density α, the algorithm $\mathcal{A}$:

- Outputs UNSAT with probability at least $\frac{1}{2}$ over the choice of Φ.[1]
- Outputs SAT if Φ is satisfiable.

[1]The choice of the fraction $\frac{1}{2}$ here is arbitrary, and one could potentially consider any fixed constant.

Note that if the algorithm $\mathcal{A}$ outputs UNSAT on an instance Φ, it certifies that the instance Φ is unsatisfiable.

Refuting random k-SAT is a seemingly intractable problem in that the best polynomial-time algorithms require density $\alpha > \tilde{O}(n^{k/2-1}) \gg \widetilde{O}(1)$. We survey the prior work on refuting CSPs in Section 3.1.

At densities far exceeding the unsatisfiability threshold, i.e., $\alpha \gg \alpha_k$, a simple union bound argument can be used to show that a random instance Φ has no assignment satisfying more than a $1 - \frac{1}{2^k} + \delta(\alpha)$ fraction of constraints, where $\delta(\alpha) \to 0$ as $\alpha \to \infty$. In this regime, a natural algorithmic task is *strong refutation*:

Definition 3.1.2. (Strong Refutation) An algorithm $\mathcal{A}$ is a *strong refutation* algorithm for random k-SAT at density α, if for a fixed constant $\delta > 0$, given a random instance Φ of k-SAT with density α, the algorithm $\mathcal{A}$:

- Outputs UNSAT with probability at least $\frac{1}{2}$ over the choice of Φ.
- Outputs SAT if Φ has an assignment satisfying at least a $(1-\delta)$-fraction of clauses.

An important conjecture in complexity theory is Feige's "R3SAT hypothesis," which states that for any $\delta > 0$, there exists some constant c such that there is no polynomial-time algorithm that can certify that a random 3-SAT instance has value at most $1 - \delta$ (that is, strongly refute 3-SAT) at clause density $m/n = c$. Feige exhibited hardness of approximation results based on the hypothesis for a class of otherwise elusive problems such as densest-k subgraph and min-bisection [Fei02]. This hypothesis has subsequently been used as the starting point in a variety of reductions (see e.g. [AAM+11, BKS13, DLS13]).

The problem of strong refutation is non-trivial even for polynomial-time solvable CSPs such as k-XOR.[2] A random k-XOR instance Φ on n variables $x_1, \ldots, x_n \in \{\pm 1\}$ consists of m equations of the form $x_{i_1} \cdot x_{i_2} \cdots x_{i_k} = \pm 1$ By a simple union bound, one can show that at all super-linear densities $m/n = \omega(1)$, with high probability, no assignment satisfies more than $\frac{1}{2} + o(1)$-fraction of the equations.[3] The problem of strong refutation for random k-XOR amounts to certifying that no assignment satisfies more than $1 - \delta$ fraction of equations for some constant $\delta > 0$. A natural spectral algorithm can efficiently strongly refute k-XOR at densities $m/n \geq n^{k/2-1}$ [CGL07, AOW15, BM16]. However, strong refutation at any lower density is widely believed to be an intractable problem [ABW10a, BM16, DLS14a, Dan16]. We refer the reader to [Dan16] for a survey of the evidence pointing to the intractability of the problem.

To expose the stark difficulty of strongly refuting random k-XOR, consider the easier task of distinguishing random k-XOR instances from those generated from the following distribution: first, sample a satisfiable instance of k-XOR uniformly at random, by sampling a planted solution $z \in \{\pm 1\}^n$ and randomly choosing m equations, each on k variables, satisfied by z. Then, corrupt each of the m equations (so that z does not satisfy it) with probability δ. Equivalently, this problem can be described as *learning parity with noise*, wherein $z \in \{\pm 1\}^n$ defines the unknown parity and each equation C_i is an *example* to the learning algorithm. An algorithm to learn parity from noisy examples can be used to distinguish the planted

[2] The *weak* refutation problem for k-XOR can be easily solved using Gaussian elimination.

[3] Random k-XOR can also be equivalently defined in terms of equations of the form $x_{i_1} \oplus \cdots x_{i_k} = 0/1$. The equivalence follows by mapping $0 \to 1$, $1 \to -1$, and $\oplus \to \cdot$.

instances sampled as described above from uniformly random instances of k-XOR. There is no known distinguishing algorithm at any density $m/n < n^{k/2-1}$, and the computational intractability of this problem has recently been used to obtain lower bounds for improper learning [Dan16].

Sum-of-Squares Refutations. A natural proof system for strong refutation is the sum-of-squares (SoS) proof system. Unfortunately, the lower bounds of Grigoriev [Gri01b] and Schoenebeck [Sch08] rule out efficient strong SoS refutations for random k-XOR and random k-SAT at densities significantly smaller than $m/n < n^{k/2-1}$. Specifically, Schoenebeck's result implies that with high probability over k-XOR instances Φ with clause density $m/n < O(n^{(k/2-1)(1-\delta)})$, the SoS hierarchy cannot refute Φ at degree $O(n^\delta)$.

Note that this leaves open the possibility that random k-XOR and random k-SAT admit subexponential-sized strong refutations well-below the $n^{k/2-1}$ threshold. This sets the stage for our main result.

Theorem 3.1.3. *For all $\delta \in [0,1)$ given a random k-XOR instance Φ on n variables, with high probability over Φ, the degree $O(n^\delta)$ sum-of-squares hierarchy can strongly refute Φ, certifying that*

$$\mathrm{opt}(\Phi) \leq \frac{1}{2} + \varepsilon\,,$$

for any constant $\varepsilon > 0$ as long as Φ has clause density $m/n \geq \widetilde{O}(n^{(k/2-1)(1-\delta)})$, where the $\widetilde{O}$ notation hides logarithmic factors and a dependence on ε and k. Further, there is a spectral algorithm achieving the same guarantees by computing the eigenvalue of an $2^{\widetilde{O}(n^\delta)} \times 2^{\widetilde{O}(n^\delta)}$ matrix.

Remark 3.1.4. The algorithm from Theorem 3.1.3 yields *tight* refutations–certifying a tight upper bound of $\mathrm{opt}(\Phi) + \varepsilon$ for any constant $\varepsilon > 0$.

Notice that the result establishes a smooth trade-off between the clause density of Φ and the running time of the refutation algorithm. Specifically for all $\delta \in [0,1)$, the algorithm strongly refutes at density $m/n = \widetilde{O}(n^{(k/2-1)(1-\delta)})$ in time $\exp(\widetilde{O}(n^\delta))$, so that when $\delta = 0$ the result matches the performance of the best known polynomial-time algorithms, and at $\delta = 1$, the algorithm refutes instances just above the threshold of satisfiability in exponential time. Moreover, the degree of the sum-of-squares refutations matches the degree lower bounds of [Gri01b, Sch08] up to polylogarithmic factors.

Feige [Fei02] introduced a connection between the refutation of random XOR instances and the refutation of other CSPs, and this connection was later used in several other works (e.g. [FKO06, AOW15, BM16]). Using the machinery developed by Allen et al. [AOW15], we apply our algorithm for k-XOR to refute other random CSPs involving arbitrary Boolean predicates P; for example to k-SAT.

Theorem 3.1.5. *Let $P : \{\pm 1\}^k \to \{0,1\}$ be a predicate with expected value $\mathbb{E}[P]$ over a random assignment in $\{\pm 1\}^k$. For all $\delta \in (0,1]$, given an instance Φ of a random k-CSP with predicate P on n variables, the degree $O(n^\delta)$ SoS hierarchy strongly refutes Φ with high probability, certifying that*

$$\mathrm{opt}(\Phi) \leq \mathbb{E}[P] + \varepsilon\,,$$

for any constant $\varepsilon > 0$ so long as Φ has density at least $m/n \geq \widetilde{O}(n^{(k/2-1)(1-\delta)})$, where the $\widetilde{O}$ hides a dependence on a polylog factor, k and ε. Further, there is a spectral algorithm achieving the same guarantees.

We can extend Theorem 3.1.5 so that the density/runtime trade-off depends on the *independence parameter* of the predicate P as defined by [AOW15], giving a better runtime/density tradeoff for some CSPs. We defer the details to Section 3.5.

Injective Tensor Norm

The proof techniques we develop are applicable beyond strongly refuting random k-XOR, to the problem of certifying upper bounds on the injective tensor norm of random tensors.

The injective tensor norm generalizes the matrix operator norm, in the following sense. For an order-k symmetric tensor with all dimensions equal to n, the injective tensor norm is defined as

$$\|\mathbf{T}\|_{inj} \stackrel{\text{def}}{=} \max_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} \left|\left\langle \mathbf{T}, x^{\otimes k} \right\rangle\right|,$$

where by $x^{\otimes k}$ we mean the symmetric rank-1 tensor of order k given by tensoring x with itself, and by the inner product we mean the entry-wise sum of the products of the entries of $\mathbf{T}$ and $x^{\otimes k}$, as is standard.

When $k = 2$, computing $\|\mathbf{T}\|_{inj}$ is equivalent to computing the matrix operator norm. Yet when $k \geq 3$, the injective tensor norm is hard to compute. The hardness of approximating the injective tensor norm is not fully understood, but we do know that, assuming the exponential-time hypothesis, the injective tensor norm requires quasipolynomial time to approximate, even within super-constant factors [BBH+12]. There are also reductions to the problem from a variety of problems such as Planted Clique [BV09] and Small-Set Expansion [BBH+12].

The problem is nontrivial even when the tensor has i.i.d. random entries. It is well-known that the norm of a tensor with i.i.d. symmetric subgaussian entries is of the same order as the norm of a random matrix:

Theorem 3.1.6 ([TS14])**.** *If $k \in \mathbb{N}$ is constant and $\mathbf{T}$ is a symmetric order-k tensor of dimension n with i.i.d. symmetrically distributed subgaussian entries, then with probability at least $1 - o(1)$, $\|\mathbf{T}\|_{inj} \leq \widetilde{O}(\sqrt{n})$.*

So the question arises naturally: is it easy to certify tensor norm bounds under distributional assumptions on the entries? The current known polynomial-time algorithms fall short of the bound $\widetilde{O}(\sqrt{n})$, and can only certify bounds of $\|\mathbf{T}\|_{inj} \leq \widetilde{O}(n^{k/4})$ for tensors of order k [RM14, HSS15, HSSS16]. The algorithm of Hopkins et al. [HSS15] is based on the degree-k SoS relaxation for the tensor norm problem. They also give a lower bound for the SoS relaxation for the order-3 tensor at degree 4, proving that the relaxation has value $\widetilde{\Omega}(n^{3/4})$, which implies that their analysis is tight for the SoS hierarchy at degree 4.

By applying our techniques for random k-XOR refutations to the problem of certifying bounds on tensor norms, we have the following result:

Theorem 3.1.7. *For any $\delta \in [0, 1/120)$, given a symmetric order-k tensor $\mathbf{T}$ with i.i.d. standard Gaussian entries, with high probability over the choice of $\mathbf{T}$, the degree $O(n^\delta)$ SoS hierarchy relaxation certifies that*

$$\|\mathbf{T}\|_{inj} \leq \widetilde{O}(n^{1/2+(k-2)(1-\delta)/4+3k^2\delta^2}),$$

where the $\widetilde{O}$ notation hides a polylogarithmic factor and a dependence on k. Furthermore, there is a spectral algorithm that computes the eigenvalues of a $2^{\widetilde{O}(n^\delta)} \times 2^{\widetilde{O}(n^\delta)}$ matrix that certifies the same bound.

We remark that the above theorem also holds, up to constants, for symmetric tensors with i.i.d. entries from any symmetric distribution $\mathcal{D}$ over $\mathbb{R}$ with subgaussian tails. Strong refutation for k-XOR instances can be thought of as a special case of the problem of bounding the norm of a random tensor–we elaborate on the connection at the start of Section 3.2. However, the underlying distribution for random k-XOR yields tensors which are extremely *sparse*, which poses several additional technical challenges.

In an independent work, Bhattiprolu et al. [BGL16] have obtained a result similar to Theorem 3.1.7 with bounds that are tighter as a function of δ. They also obtain a tight lower bound on the integrality gap of degree-k SoS relaxations for k-tensor norms.

Related work

We briefly survey the prior work on refuting random CSPs–we refer the reader to [AOW15] for a thorough survey on the topic. Work on refuting random CSPs began with Chvátal and Szemerédi [CS88], who showed that a random k-SAT instance with clause density $\alpha > c$ (for c constant) with high probability requires Resolution refutations of exponential size. This lower bound was later complemented by the works of [Fu98, BKPS98], which show that at clause density $\alpha \geq O(n^{k-1})$, polynomial-sized resolution proofs exist and can be found efficiently. At the turn of the century, Goerdt and Krivelevich [GK01] pioneered the spectral approach to refuting CSPs, showing that a natural spectral algorithm gives refutations for k-SAT in polynomial time when $\alpha = m/n \geq n^{\lceil k/2 \rceil - 1}$. A series of improvements followed, first achieving bounds for $\alpha \geq \widetilde{O}(n^{1/2})$ for the special case of 3-SAT [FG01, FGK05, CGL07], then achieving strong refutation at densities $\alpha \geq \widetilde{O}(n^{\lceil k/2 \rceil - 1})$ [CCF10]. Finally, the works of Allen et al. and Barak and Moitra gave spectral algorithms for strongly refuting k-XOR and k-SAT for any $\alpha \geq \widetilde{O}(n^{k/2-1})$ [AOW15, BM16], and Allen et al. also give a reduction from any CSP which is far from supporting a t-wise independent distribution to t-XOR. These spectral algorithms are the algorithmic frontier for efficient refutations of random CSPs.

Though not algorithmic, the work of [FKO06] is worth mentioning as well. Feige et al. show that, at clause density $\alpha = m/n \geq \widetilde{O}(n^{0.4})$, there exists a polynomial-sized (weak) refutation for random 3-SAT given by a subset of $O(n^{0.2})$ unsatisfiable clauses. Understanding whether polynomial-sized weak refutations exist for smaller α is an intriguing open problem.

Organization

In Section 3.2, we illustrate the technical core of our ideas via a detailed exposition of our proof for certifying bounds on the norm of order-4 tensors, and explain how these techniques

can be built upon to strongly refute CSPs. Section 3.3 contains the full proof our tensor norm results. In Section 3.4, we give our results for refuting k-XOR instances. In Section 3.5, we combine our k-XOR refutation algorithms with the framework of [AOW15] to refute other CSPs. Finally, in Section 3.6 we argue that our spectral algorithms give SoS proofs.

3.2 Main Ideas: Proof for Random 4-Tensors

In this section, we will explain the technical core of our result by proving Theorem 3.1.7 (our tensor norm certification algorithm) for the case of random 4-tensors. This specific case yields the simplest proof, while encapsulating the core ideas of our techniques for both injective tensor norm and k-XOR. We formally state the injective tensor norm problem here.

Problem 3.2.1 (Certifying injective tensor norm)**.** Given an order-k tensor $\mathbf{T}$ with dimension n, certify that for all $x \in \mathbb{R}^n$ with $\|x\| = 1$, $|\langle \mathbf{T}, x^{\otimes k}\rangle| \leq \|\mathbf{T}\|_{inj} \leq \tau$ for some upper bound τ.

From k-XOR to Tensor Norms. First, we briefly outline the connection between k-XOR refutation and certifying bounds on tensor norms. Let Φ be a random k-XOR formula on $x \in \{\pm 1\}^n$ with $m \approx pn^k$ clauses, sampled as follows: for each $S \subset [n]^k$ independently with probability p, add the constraint that $\prod_{i \in S} x_i = \eta_S$ where η_S is a uniform bit ± 1, and with probability $1 - p$, add no constraint. We can form an order-k tensor $\mathbf{T}$ so that for each $S \in [n]^k$, $\mathbf{T}_S = 0$ if there is no constraint, and otherwise $\mathbf{T}_S = \eta_S$.

For any assignment $x \in \{\pm 1\}^n$, the inner product $\langle \mathbf{T}, x^{\otimes k}\rangle$ is equal to the difference in the number of Φ's constraints that x does and does not satisfy. Since Φ has m constraints in all, certifying that $\max_{x \in \{\pm 1\}^n} |\langle \mathbf{T}, x^{\otimes k}\rangle| \leq o(m)$ is equivalent to certifying that $\mathrm{opt}(\Phi) \leq \frac{1}{2} + o(1)$. On the other hand, certifying the injective tensor norm amounts to exhibiting an upper bound on $\max_{\|y\| \leq 1} |\langle \mathbf{T}, y^{\otimes k}\rangle|$ where the maximization is over all unit vectors y. Every Boolean vector $x \in \{\pm 1\}^n$ is of length $\|x\| = \sqrt{n}$, which implies that $\max_{x \in \{\pm 1\}^n} |\langle \mathbf{T}, x^{\otimes k}\rangle| \leq n^{k/2} \cdot \|\mathbf{T}\|_{inj}$.

While the above reduction from certifying that $\mathrm{opt}(\Phi) \leq \frac{1}{2} + o(1)$ to certifying a bound on $\|\mathbf{T}\|_{inj}$ exposes the connection between the two problems, it is too lossy to be useful. In fact, for $p < 1/n^{k/2}$, the sparsity of the tensor $\mathbf{T}$ implies that the form $\langle y^{\otimes k}, \mathbf{T}\rangle$ is maximized by sparse real-valued vectors y that are completely unlike Boolean vectors. In other words, almost surely there exists a sparse $y \in \mathbb{R}^n$, $\|y\| = \sqrt{n}$ with $|\langle \mathbf{T}, y^{\otimes k}\rangle| \gg \max_{x \in \{\pm 1\}^n} |\langle \mathbf{T}, x^{\otimes k}\rangle|$. As a result, our refutation algorithm for k-XOR is more involved than the certification algorithm for injective tensor norm in two ways. First, it crucially uses the *non-sparseness* of Boolean vectors and second, the sparsity of the tensor $\mathbf{T}$ calls for more nuanced concentration arguments. We give a short overview of these differences in Section 3.2 after presenting the broad strokes of the proof, via our algorithm for certifying tensor norms, and full details in Section 3.4.

Certifying injective tensor norm. In what follows, we will give a *spectral algorithm* for Problem 3.2.1 for random 4-tensors with i.i.d. subgaussian entries. We will show that this

spectral algorithm is subsumed by a SoS relaxation of appropriate degree in Section 3.6. The rest of this section is organized as follows.

1. We first **d**escribe the matrix whose maximum eigenvalue provides the upper bound on the injective tensor norm. Rather than writing down the matrix immediately, we will build up our intuition by first considering a simple spectral approach, and then seeing how we can improve.
2. We then **o**btain bounds on the eigenvalues of the matrix, which will hold with high probability for tensors with i.i.d. subgaussian entries–this is the step in which we analyze the performance of our algorithm. Because our matrix is somewhat complicated and not amenable to the application of black-box matrix concentration inequalities, we will apply the *trace power method*. This amounts to bounding the expected trace of a large power of our matrix, a goal which we split in to two steps.
   a) First, we *reduce computing the expected trace to a hypergraph counting problem.*
   b) Then, we simplify the counting by *analyzing a particular hypergraph sampling process.*

Improving on the Natural Spectral Algorithm with Higher-Order Symmetries

A natural spectral algorithm for Problem 3.2.1 is to flatten the tensor to a matrix, and then compute the operator norm of the matrix. This is a valid relaxation because, given an order-4 tensor $\mathbf{A}$ with symmetric i.i.d. standard normal entries, if we take A to be the natural $n^2 \times n^2$ matrix flattening of $\mathbf{A}$,

$$\|\mathbf{A}\|_{inj} = \max_{x\in\mathbb{R}^n:\|x\|=1} \left|(x\otimes x)^\top A(x\otimes x)\right| \leq \max_{y\in\mathbb{R}^{n^2}:\|y\|=1} \left|y^\top Ay\right| = \|A\|_{op}\,. \tag{3.2.1}$$

So $\|A\|_{op}$ gives a valid upper bound for $\|\mathbf{A}\|_{inj}$. This is great–on the left, we have a program that we cannot efficiently optimize, and on the right we have a relaxation which we can compute in polynomial time.

On the other hand this bound is quite loose–classical results from random matrix theory assert that with high probability, $\|A\|_{op} = \widetilde{\Theta}(n)$ whereas with high probability $\|\mathbf{A}\|_{inj} \leq O(\sqrt{n})$. The issue is that the relaxation in (3.2.1) is too lenient–the large eigenvalues of A correspond to eigenvectors $y \in \mathbb{R}^{n^2}$, that are far from vectors of the form $x\otimes x : x \in \mathbb{R}^n$. We want to decrease the spectrum of A along these asymmetric *non-tensor product* directions.

A tensored vector of the form $x\otimes x$ satisfies the symmetry that $(x\otimes x)_{ij} = (x\otimes x)_{ji} = x_i x_j$. Therefore, a natural approach to decrease the spectrum of A along the *non-tensor product* directions is to average the matrix A, along these symmetries. Specifically, for each (i,j), we would average the ij^{th} and ji^{th} rows, and then repeat the same operation on columns. Formally, the averaged matrix A' is given by,

$$A' = \mathop{\mathbb{E}}_{\Sigma,\Pi\in\hat{\mathcal{S}}_2} [\Sigma A \Pi]$$

where $\hat{\mathcal{S}}_2$ is the set of matrices which perform the permutations corresponding to the symmetric group on 2 elements on the rows and columns of matrices indexed by $[n]^2$. Unfortunately,

for a symmetric 4-tensor $\mathbf{A}$, the matrix A is also symmetric with respect to these operations, so that $A' = A$.

To better exploit the symmetries of tensored vectors $x \otimes x$, we will work with higher powers of the injective tensor norm. For any $d \in \mathbb{N}$, we can write the d^{th}-power of $\|\mathbf{A}\|_{inj}$ as

$$\|\mathbf{A}\|_{inj}^d = \max_{x \in \mathbb{R}^n, \|x\|=1} \left|\langle x^{\otimes 4}, \mathbf{A}\rangle^d\right| = \max_{x \in \mathbb{R}^n, \|x\|=1} \left|(x^{\otimes 2d})^\top A^{\otimes d} x^{\otimes 2d}\right|,$$

where $A^{\otimes d}$ is the natural $n^{2d} \times n^{2d}$ matrix flattening of $\mathbf{A}^{\otimes d}$. The symmetric vector $x^{\otimes 2d}$ is fixed by averaging over any permutation of the indices, so averaging over such permutations does not change the maximum:

$$= \max_{x \in \mathbb{R}^n, \|x\|=1} \left| \mathop{\mathbb{E}}_{\Pi, \Sigma \in \hat{\mathcal{S}}_{2d}} \left[(\Pi x^{\otimes 2d})^\top A^{\otimes d} (\Sigma x^{\otimes 2d})\right] \right|,$$

and by linearity of expectation,

$$= \max_{x \in \mathbb{R}^n, \|x\|=1} \left| (x^{\otimes 2d})^\top \left(\mathop{\mathbb{E}}_{\Pi, \Sigma \in \hat{\mathcal{S}}_{2d}} \left[\Pi^\top A^{\otimes d} \Sigma\right] \right) x^{\otimes 2d} \right| \quad \leq \quad \left\| \mathop{\mathbb{E}}_{\Pi, \Sigma \in \hat{\mathcal{S}}_{2d}} \left[\Pi^\top A^{\otimes d} \Sigma\right] \right\|_{op}. \tag{3.2.2}$$

The operator norm of the above described matrix will certify our upper bounds:

Proposition 3.2.2. *Let $k \in \mathbb{N}$ be even. Let $\hat{\mathcal{S}}_{kd/2}$ be the set of matrices performing the permutations of $\mathcal{S}_{kd/2}$ on matrices with rows and columns indexed by $[n]^{kd/2}$. For any order-k tensor $\mathbf{A}$ with matrix flattening A,*

$$\|\mathbf{A}\|_{inj} \leq \left(\left\| \mathop{\mathbb{E}}_{\Pi, \Sigma \in \hat{\mathcal{S}}_{kd/2}} \left[\Pi A^{\otimes d} \Sigma\right] \right\|_{op} \right)^{1/d}.$$

Proof. The sequence of calculations culminating in (3.2.2) gives the proof. □

Now, how can this give an improved upper bound over $\|A^{\otimes d}\|_{op} = \|A\|_{op}^d$? The reason is that although $\mathbf{A}$ had 4-wise symmetry, the tensor $\mathbf{A}^{\otimes d}$ *does not have* $4d$-wise symmetry. For $I, J \in [n]^{2d}$, $I = (i_1, i_1'), \ldots, (i_d, i_d')$ and $J = (j_1, j_1'), \ldots, (j_d, j_d')$ and for permutations π, σ on $2d$ elements,

$$(A^{\otimes d})_{I,J} \;=\; \prod_{\substack{\ell=1 \\ (i_\ell, i_\ell') \in I \\ (j_\ell, j_\ell') \in J}}^{d} A_{i_\ell, i_\ell', j_\ell, j_\ell'} \;\neq\; \prod_{\substack{\ell=1 \\ (a_\ell, a_\ell') \in \pi(I) \\ (b_\ell, b_\ell') \in \sigma(J)}}^{d} A_{a_\ell, a_\ell', b_\ell, b_\ell'} \;=\; (A^{\otimes d})_{\pi(I), \sigma(J)},$$

because the identity of the base variables in the expression may change under the permutation of the indices I and J. Thus, the typical entry of $\mathbb{E}_{\Pi, \Sigma \in \hat{\mathcal{S}}_{2d}} \left[\Pi (A^{\otimes d}) \Sigma\right]$ is an average of $(d/2!)^2$ random variables, which are not independent, but also not identical. Since the

entries of A are distributed symmetrically about zero, we expect the magnitude of the typical entry to drop after this averaging. If we indulge the heuristic assumption that the entries of $\mathbb{E}_{\Pi,\Sigma\in\hat{\mathcal{S}}_{2d}}\left[\Pi(A^{\otimes d})\Sigma\right]$ are averages of $d^{\Omega(d)}$ *independent* random symmetric variables of constant variance, then the magnitude of the typical entry should be $\approx \frac{1}{d^{\Omega(d)}}$. So heuristically, we have that

$$\left\|\mathop{\mathbb{E}}_{\Pi,\Sigma\in\hat{\mathcal{S}}_{2d}}\left[\Pi(A^{\otimes d})\Sigma\right]\right\|_F \leq \frac{1}{d^{\Omega(d)}}\cdot\|A^{\otimes d}\|_F.$$

By Wigner's semicircle law, matrices with independent entries have eigenvalues that are all roughly of the same magnitude. Because our matrix has roughly independent entries, we may hope that the semicircle law holds for us, so that from the above heuristic calculations and from (3.2.2),

$$\|\mathbf{A}\|_{inj} \leq \left(\left\|\mathop{\mathbb{E}}_{\Pi,\Sigma\in\hat{\mathcal{S}}_{2d}}\left[\Pi(A^{\otimes d})\Sigma\right]\right\|_{op}\right)^{1/d} \leq \left(\frac{1}{d^{\Omega(d)}}\cdot\|A^{\otimes d}\|_{op}\right)^{1/d} \leq \frac{n}{d^{\Omega(1)}}\,.$$

Thus, we expect that as we increase d, and therefore increase the symmetry of the *tensored vectors* $x\otimes x$ relative to the "noisy" non-tensor product eigenvectors of A, we can certify a tighter upper bound on $\|\mathbf{A}\|_{inj}$. Of course, since our certificate is the eigenvalue of a $n^{2d}\times n^{2d}$ matrix, the running time the refutation algorithm grows exponentially in the choice of d.

Matrix concentration for the certificate

Our algorithm is now clear: we form our matrix certificate by averaging over rows and columns corresponding to permutations of row and column indices in $A^{\otimes d}$, then use the certificate matrix's eigenvalues to upper bound $\|\mathbf{A}\|_{inj}^d$ (by Proposition 3.2.2).

Theorem 3.2.3. *Let $n, d\in\mathbb{N}$. Let $\mathbf{A}$ be an order-4 tensor with independent entries, distributed according to subgaussian distribution symmetric about 0. Then if $d\log n \ll n$, with high probability over $\mathbf{A}$,*

$$\left\|\mathop{\mathbb{E}}_{\Pi,\Sigma}\left[\Pi A^{\otimes d}\Sigma\right]\right\|^{1/d} \leq \widetilde{O}\left(\frac{n}{d^{1/2}}d^{12\frac{\log d}{\log n}}\right)\,.$$

As a corollary of Theorem 3.2.3 and Proposition 3.2.2, we get Theorem 3.1.7 for the case of order-4 tensors.

At the end of the previous subsection we gave a heuristic argument that a statement along the lines of Theorem 3.2.3 should be true. While the heuristic argument is plausible, it is very far from a formal proof; we need to prove that the eigenvalues of $\mathbb{E}[\Pi A^{\otimes d}\Sigma]$ are bounded by $\approx \widetilde{O}(n/\sqrt{d})^d$ with high probability. But the matrix $\mathbb{E}[\Pi A^{\otimes d}\Sigma]$ is not a sum of independent random matrices, and it does not have independent entries, so sophisticated matrix concentration tools (like the semicircle law or matrix Chernoff bounds) do not apply. For tasks of this sort, the trace power method, or the method of moments, is the tool of choice:

Proposition 3.2.4 (Trace power method)**.** *Let $n, \ell \in \mathbb{N}$, let $c \in \mathbb{R}$, and let M be an $n \times n$ random matrix. Then*

$$\mathop{\mathbb{E}}_{M}[\mathrm{Tr}((MM^\top)^\ell)] \leq \beta \implies \mathbb{P}\left(\|M\| \geq c \cdot \beta^{1/2\ell}\right) \geq 1 - c^{-2\ell}\,.$$

The proof is essentially an application of Markov's inequality; we give it in Appendix A.3.

From bounding the expected trace to a hypergraph counting problem

A classic way to apply the trace power method is to reduce to a graph counting problem. For example, let M be a symmetric $n \times n$ random matrix with independent Rademacher entries. We can view the row/column index set $[n]$ as a set of "vertices," and the entry $M_{i,j}$ as an "edge" variable between vertices i and j. The trace $\mathrm{Tr}(M^\ell)$ is the sum over products of edge variables along closed walks of length ℓ in the graph defined by M. When we take $\mathbb{E}_M[\mathrm{Tr}(M^\ell)]$, any closed walk in which an edge appears with odd multiplicity does not contribute to the sum, since $\mathbb{E}[M_{i,j}^m] = 0$ for odd m. Therefore, $\mathbb{E}[\mathrm{Tr}(M^\ell)]$ is equal to the number of closed walks of length ℓ in which every edge appears with even multiplicity, within the complete graph K_n, and bounding $\mathbb{E}[\mathrm{Tr}(M^\ell)]$ becomes a counting problem.

We make a similar reduction for our matrix $C \stackrel{\text{def}}{=} \mathbb{E}_{\Sigma,\Pi}[\Pi A^{\otimes d}\Sigma]$. The rows and columns of A are indexed by pairs $[n]^2$, we interpret each variable $A_{ij,k\ell}$ as a (multi)hyperedge between the vertices (i,j) and (k,ℓ) corresponding to the row and column indices respectively. In the Kronecker power $A^{\otimes d}$, the rows and columns are indexed by vertex multisets $I, J \in [n]^{2d}$, $I = (i_1, i_1', \ldots, i_d, i_d')$, $J = (j_1, j_1', \ldots, j_d, j_d')$, and the entry $(A^{\otimes d})_{I,J}$ is the product of the hyperedges $\prod_{k=1}^d A_{i_k i_k', j_k j_k'}$. We view this as a hyperedge matching between I, J, in which the vertices (i_k, i_k') are matched with the vertices (j_k, j_k') for each $k \in [d]$ (see Figure 3.1).

Now to obtain our matrix C, we average over row and column symmetries, so that $C_{I,J} = \mathbb{E}_{\pi,\sigma \in \mathcal{S}_{2d}}[(A^{\otimes d})_{\pi(I),\sigma(J)}]$. In each entry of C, we average over the permutations of the left and right vertex sets, which is the same as averaging over all perfect hypergraph matchings from I to J (again see Figure 3.1).

Just as in the case of the simple random matrix M, we can interpret $\mathrm{Tr}((CC^\top)^\ell)$ as the sum over all closed walks of length 2ℓ on the complete graph (with self-loops) on the vertex set $[n]^{2d}$, where the edge variable between I, J is the *average* over all possible hyperedge matchings between I and J. When we take the expectation over $\mathbf{A}$, $\mathbb{E}_{\mathbf{A}}[\mathrm{Tr}((CC^\top)^\ell)]$, any *hyperedge* appearing with odd multiplicity will cause the contribution of the closed walk to be 0, since the entries of $\mathbf{A}$ are distributed symmetrically about 0.

Our reduction is now complete. Because we will be dealing with subgaussian random variables, the entries of $\mathbf{A}$ will concentrate well enough for us to reduce to the Rademacher case.

Lemma 3.2.5. *Let $\mathbf{A}$ be an order-4 tensor with i.i.d. Rademacher entries, and let A be its matrix flattening. Let $C_d \stackrel{\text{def}}{=} \mathbb{E}_{\Sigma,\Pi \in \hat{\mathcal{S}}_{2d}}[\Pi A^{\otimes d}\Sigma]$. For the 2ℓ multisets of vertices $I_1, \ldots, I_{2\ell} \in [n]^{2d}$, let $\mathcal{H}$ be the set of all sequences of perfect hyperedge matchings between each I_j and $I_{j+1 \mod 2\ell}$, so that each hyperedge has 2 vertices from I_j and 2 vertices from I_{j+1}. For a fixed sequence of hyperedge matchings $H \in \mathcal{H}$, let $\mathcal{E}_{I_1,\ldots,I_{2\ell}}(H \text{ even})$ be the event that every*

hyperedge appears with even multiplicity. Then

$$\mathop{\mathbb{E}}_{\mathbf{A}}\left[\mathrm{Tr}\left((C_d C_d^\top)^\ell\right)\right] = \sum_{I_1,\dots,I_{2\ell}\in[n]^{2d}} \mathop{\mathbb{P}}_{H\sim\mathcal{H}}[\mathcal{E}_{I_1,\dots,I_{2\ell}}(H \text{ even})]$$

Proof. Any product of Rademacher random variables has expectation 0 if some variable appears with odd multiplicity, and 1 otherwise. This, along with the observations preceding the lemma statement, implies that each $I_1,\dots,I_{2\ell}$ contributes exactly the probability that hyperedges chosen for it all have even multiplicity (where we get a probability since each entry $C_{I,J}$ is the average over hyperedge matchings from I to J). □

Bounding the probability of an even hypergraph

From Lemma 3.2.5 and Proposition 3.2.4, in order to prove Theorem 3.2.3 it suffices for us to bound

$$\sum_{I_1,\dots,I_{2\ell}\in[n]^{2d}} \mathop{\mathbb{P}}_{H\sim\mathcal{H}}[\mathcal{E}_{I_1,\dots,I_{2\ell}}(H \text{ even})] \le \widetilde{O}(n/d^{1/2})^{2d\ell}, \tag{3.2.3}$$

for $\ell = \Omega(\log n)$. Since each probability is bounded by 1 and there are $n^{4d\ell}$ terms in the sum, (3.2.3) easily gives us an upper bound of $n^{4d\ell}$. We need to improve upon this naive bound twofold: first, we need the dependence on n to be $n^{2d\ell}$. This would give a bound of $\|\mathbb{E}[\Sigma A^{\otimes d}\Pi]\| \le \widetilde{O}(n^d)$ w.h.p., but we can get this bound trivially by ignoring the symmetrization, as $\|A^{\otimes d}\| \le \widetilde{O}(n^d)$ w.h.p. To fully reap the rewards of symmetrization, we must improve by a factor of $\approx (\sqrt{d})^{-2d\ell}$.

At first, bounding (3.2.3) seems daunting–it is unclear how to count the number of such hypergraphs with even multiplicity, while simultaneously getting the correct dependence on n and d. It will be helpful to use the following two-step process for sampling hypergraphs: for a fixed vertex configuration $I_1,\dots,I_{2\ell}\in[n]^{2d}$,

1. First, sample perfect simple edge matchings between I_j, I_{j+1} for each $j\in[2\ell]$.
2. Next, pair up the edges between I_j, I_{j+1} and merge each pair to form a hyperedge.

We will use step 1 to bound the dependence on n, and step 2 to bound the dependence on d. In particular, our arguments from Lemma 3.2.5 give us the following lemma almost immediately:

Lemma 3.2.6. *Let $\mathcal{M}$ be the set of all possible choices of edge sets sampled in step 1. Let $\mathcal{E}_{I_1,\dots,I_{2\ell}}(E \text{ even})$ be the event that the graph given by the edges $E\in\mathcal{M}$ on $I_1,\dots,I_{2\ell}$ has every edge appearing with even multiplicity. Then*

$$\sum_{I_1,\dots,I_{2\ell}\in[n]^{2d}} \mathop{\mathbb{P}}_{E\sim\mathcal{M}}[\mathcal{E}_{I_1,\dots,I_{2\ell}}(E \text{ even})] = \mathop{\mathbb{E}}_{M}\left[\mathrm{Tr}\left((BB^\top)^\ell\right)\right], \tag{3.2.4}$$

where M is an $n\times n$ matrix with i.i.d. Rademacher entries, and $B \stackrel{\text{def}}{=} \mathbb{E}_{\Pi,\Sigma\in\hat{\mathcal{S}}_{2d}}[\Pi M^{\otimes 2d}\Sigma]$.

Lemma 3.2.6 lets us relate the probability that we sample a perfect matching in which every edge appears with even multiplicity in step 1 to the norm of a matrix M with i.i.d.

Rademacher entries, which is an object we understand well: with very high probability, $\|M\| \le O(\sqrt{n})$, and because of the connection between the expected trace and the norm of a matrix, we can then bound (3.2.4) by the desired $\widetilde{O}(n^{1/2})^{4d\ell}$.

To use (3.2.4), we need to relate the probability that the edges sampled in step 1 have even multiplicity to the probability that the hyperedges sampled in step 2 have even multiplicity.

Lemma 3.2.7 (somewhat informal statement)**.** *Let $I_1, \dots, I_{2\ell} \in [n]^{2d}$, and suppose we have sampled hyperedges $H \in \mathcal{H}$ by first sampling simple edges $E \in \mathcal{M}$ as in step 1 and then grouping them as in step 2. Then*

$$\mathbb{P}(\mathcal{E}_{I_1,\dots,I_{2\ell}}(E \text{ even}) \mid \mathcal{E}_{I_1,\dots,I_{2\ell}}(H \text{ even})) \ge \left(\frac{1}{2}\right)^{2d\ell}.$$

Proof (sketch). For any given hyperedge $(i,j,k,\ell) \in H$, with $i,j \in I_a$ and $k,\ell \in I_{a+1}$, there are only two ways it could have been sampled as pairs of edges, either as a merge of $(i,k),(j,\ell) \in E$ or of $(i,\ell),(j,k) \in E$. If all copies of a hyperedge of even multiplicity m are sampled the same way, then the corresponding edges also have even multiplicity.[4] For a hyperedge of multiplicity m, every copy of the hyperedge is sampled in the same way with probability at least $(1/2)^m$, which becomes $(1/2)^{2d\ell}$ for the $2d\ell$ hyperedges in the graph. □

Now, using the shorthand $\mathcal{E}(\cdot) \stackrel{\text{def}}{=} \mathcal{E}_{I_1,\dots,I_{2\ell}}(\cdot \text{ even})$, we already have that

$$\mathop{\mathbb{P}}_{H\sim\mathcal{H}}[\mathcal{E}(H)] = \frac{\mathbb{P}[\,\mathcal{E}(H),\mathcal{E}(E)\,]}{\mathbb{P}[\mathcal{E}(E) \mid \mathcal{E}(H))} \le 2^{2d\ell} \cdot \mathbb{P}[\mathcal{E}(H) \mid \mathcal{E}(E)] \cdot \mathbb{P}[\mathcal{E}(E)].$$

Further, we have our bound from Lemma 3.2.6, so if we could bound $\max_{I_1,\dots,I_{2\ell}} \mathbb{P}[\mathcal{E}(H \text{ even}) \mid \mathcal{E}(E \text{ even})] \le d^{-2k\ell}$, we would be done. But this conditional probability is not always small–for example, there is the case when $I_1 = \cdots = I_{2\ell}$ are all multisets containing the same vertex $i \in [n]$ with multiplicity $2d$. In this case, the probability that we sample an even hypergraph is 1.

Still, so long as there are sufficiently many different vertices in $I_1, \dots, I_{2\ell}$, we can prove that this conditional probability is small enough:

Lemma 3.2.8. *Let $E_1, \dots, E_{2\ell} \in [n \times n]^{2d}$ be multisets of edges such that every edge is present in the union at least twice, and the number of distinct edges in the union is at least $(1-\beta)2d\ell$, i.e., $|\cup_{i=1}^{2\ell} E_i| \ge (1-\beta)2d\ell$.*

Let P_i denote a uniformly random pairing of elements within E_i sampled independently for each $i \in [2\ell]$. Then there exists a constant c_β depending only on β such that

$$\mathbb{P}[\cup_i P_i \text{ has every pair with even multiplicity}] \le \left(\frac{c_\beta}{d}\right)^{(1-10\beta)d\ell}.$$

[4] In the formal proof, we'll have to take care to start with an asymmetric tensor, with $\mathbf{A}_{ijk\ell} \ne \mathbf{A}_{\pi(ijk\ell)}$ for permutations π, so that no hyperedge can appear with even multiplicity by being grouped from the edges $(i,k),(j,\ell)$ and also $(i,j),(k,\ell)$.

Proof (sketch, details in proof of Lemma 3.3.8). Suppose we make our pairing decisions one multiset at a time. We must pair the last copy of each edge correctly, so that all its pairs have even multiplicity. There are $2d$ edges per matching, so the probability that we make this last decision correctly is $\approx \Omega(d)^{-1}$. We make d pairing decisions per matching, and we make the "last" decision about half of the time since every edge appears close to twice on average–this gives the probability to be roughly $\Omega(d)^{-d\ell}$. □

Now, as there are only $\approx n^{(1-\alpha)\cdot 2d\ell}$ choices of sets $I_1, \ldots, I_{2\ell}$ which could have at most $(1-\alpha)\cdot 2d\ell$ different edges, these sets contribute negligibly to the sum, and we have that

$$\begin{aligned}\sum_{I_1,\ldots,I_{2\ell}} \mathop{\mathbb{P}}_{H\sim\mathcal{H}}\left[\mathcal{E}_{I_1,\ldots,I_{2\ell}}(H)\right] &\le 2^{2d\ell}\cdot\left(n^{(1-\alpha)\cdot 2d\ell} + \mathbb{P}[\mathcal{E}(H)\mid\mathcal{E}(E)]\sum_{I_1,\ldots,I_{2\ell}}\mathbb{P}\left[\mathcal{E}_{I_1,\ldots,I_{2\ell}}(E)\right]\right)\\ &\le 2^{2d\ell}\cdot\left(n^{(1-\alpha)\cdot 2d\ell} + \left(\frac{c_\alpha}{\sqrt{d}}\right)^{(1-10\alpha)2d\ell}\cdot n^{2d\ell}\right)\end{aligned}$$

Balancing the terms concludes the proof; we will fill in the few remaining details in Section 3.3.

From Tensor Norms to Odd-Order Tensors and k-XOR

The proof of Theorem 3.2.3 generalizes to tensors of all even orders k almost immediately. For odd k we need an extra idea or two, since all natural flattenings of the tensor to a matrix result in a non-square matrix. We give the details for even and odd k in Section 3.3 and Section 3.3 respectively.

As hinted earlier, to apply these ideas to strongly refute k-XOR we need to overcome two main hurdles. First, as the number of clauses is small, $m \approx p\cdot n^k < n^{k/2}$, the tensor corresponding to the instance is sparse enough that the injective tensor norm $\max_{y\in\mathbb{R}^n}|\langle \mathbf{T}, y^{\otimes k}\rangle|$ is maximized by *sparse* vectors y. Sparse vectors $y\in\mathbb{R}^n$ are too far from the solutions of interest, namely Boolean vectors $x\in\{\pm1\}^n$, which are in a sense maximally dense.

To address this issue, we will consider a sub-matrix of the tensored matrix $A^{\otimes d}$. Again, let us consider the case of $k=4$. Recall that, $\max_{x\in\{\pm1\}^n}\langle \mathbf{A}, x^{\otimes 4}\rangle^d = \max_{x\in\{\pm1\}^n}\left|x^{\otimes 2d}A^{\otimes d}x^{\otimes 2d}\right|$. The rows and columns of $A^{\otimes d}$ are indexed by $I, J\in[n]^{2d}$. We refer to a tuple $I\in[n]^{2d}$ as *high multiplicity* if there is some $i\in[n]$ which has multiplicity greater than $100\log n$ in I (since we are interested in the case when $d = n^\delta \gg \log n$). The rows and columns of $A^{\otimes d}$ corresponding to such tuples will be referred to as high-multiplicity rows and columns. Let Γ denote the projection on to the low-multiplicity indices, $(\Gamma x)_I = x_I\cdot\mathbb{I}[I \text{ not high-multiplicity}]$.

The key idea is that for a Boolean vector $x\in\{\pm1\}^n$, almost all of the ℓ_2-norm of $x^{\otimes 2d}$ is concentrated within the low-multiplicity indices, i.e., $\|\Gamma x^{\otimes 2d}\| \approx \|x^{\otimes 2d}\|$. However, for a sparse vector $y\in\mathbb{R}^n$, $\|\Gamma y^{\otimes 2d}\| \ll \|y^{\otimes 2d}\|$. Therefore, we eliminate the sparse maxima of the polynomial, by restricting the matrix to the low-multiplicity rows and columns, and then apply the averaging over row and column permutations. Specifically, the spectral upper bound used by the refutation algorithm is,

$$\max_{x\in\{\pm1\}^n}\left|\langle \mathbf{A}, x^{\otimes 4}\rangle\right|^d = \max_{x\in\{\pm1\}^n}\left|(x^{\otimes 2d})^\top A^{\otimes d}x^{\otimes 2d}\right| \approx \max_{x\in\{\pm1\}^n}\left|(x^{\otimes 2d})^\top\left(\Gamma A^{\otimes d}\Gamma^T\right)x^{\otimes 2d}\right|$$

$$\leq n^{2d} \cdot \left\| \mathop{\mathbb{E}}_{\Pi,\Sigma\in\hat{\mathcal{S}}_{2d}} \left[\Pi \left(\Gamma A^{\otimes d}\Gamma^T\right)\Sigma\right] \right\| .$$

The second challenge is that, in the sparse regime where $p \leq 1/n^{k/2}$, the entries of the random matrix A are ill-behaved. Specifically, the entries of A have distributions with unusually large higher moments, completely unlike Gaussian or Rademacher random variables. For example, the $2r^{th}$ moment of an entry $\mathbb{E}[A_{ijk\ell}^{2r}] = p \gg (\mathbb{E}[A_{ijk\ell}^2])^r = p^r$. In the trace calculation we outlined earlier, each term of the sum was either 0 if any variable had odd multiplicity, and otherwise 1. In the sparse regime, different terms in the trace contribute vastly different amounts, depending on the multiplicities involved. So we must count our hypergraphs precisely, taking into account the multiplicity of each hyperedge, rather than the just the parity. We use the *encoding technique* to count the number of hypergraph structures accurately, in a way reminiscent of similar arguments in random matrix theory (e.g. [FK81]). Although the counting argument involved is more subtle than the case of random 4-tensors (see Section 3.4), we are still able to use the same 2-step hyperedge sampling process to simplify the counting.

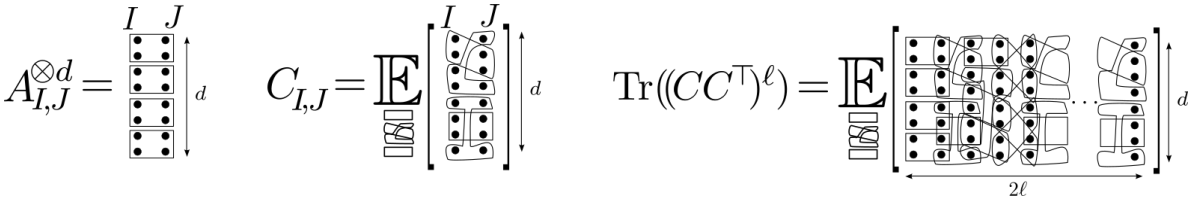


Figure 3.1: Hypergraph interpretations of the entries of $A^{\otimes d}$, C_d, and $\mathrm{Tr}((CC^\top)^\ell)$.

3.3 Injective Tensor Norm for Subgaussian Random Tensors

In this section, we show how to certify bounds on the norm of a random tensor, building on our proof of the order-4 case in Section 3.2. We handle the even-order and odd-order cases separately, as the odd-order case contains some additional intricacies.

Section 3.3 contains the proof for even tensors. Section 3.3 contains the proof for odd tensors. In Section 3.3, we prove a combinatorial lemma that we rely upon in both proofs.

Even-Order Tensors

The case of order-k tensors when k is even is almost completely outlined in Section 3.2, in the proof overview of Theorem 3.2.3. Some of the statements from the overview need additional proof, and some need generalization for $k > 4$. We briefly fill in the gaps.

Recall that in our setting, we are given a symmetric order-k tensor $\mathbf{A}$ with i.i.d. standard Gaussian entries, where k is even. Our algorithm consists of computing the operator norm of a certificate matrix; though we described this certificate ion Section 3.2, we will require one small twist to make our proofs easier:

Algorithm 3.3.1 (Certifying even k-tensor norms)**.**
Input: An order-k dimension-n tensor $\mathbf{A}$, for even k.

1. Form the asymmetric tensor $\mathbf{A}'$ from $\mathbf{A}$ as follows. For each $S \in [n]^k$,
   a) if S is lexicographically first among all permutations of S, set $\mathbf{A}'_S = \sum_{\pi \in \mathcal{S}_k} \mathbf{A}_{\pi(S)}$.
   b) otherwise, set $\mathbf{A}'_S = 0$.
2. Take the natural $n^{k/2} \times n^{k/2}$ matrix flattening A of $\mathbf{A}'$, and form $A^{\otimes d}$.
3. Letting $\hat{\mathcal{S}}_{dk/2}$ be the set of all permutation matrices that perform the index permutations corresponding to $\mathcal{S}_{dk/2}$ on the rows and columns of $A^{\otimes d}$, form

$$C_d \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{\Pi,\Sigma \in \hat{\mathcal{S}}_{dk/2}} \left[\Pi A^{\otimes d} \Sigma\right] .$$

Output: $\|C_d\|^{1/d}$ as a bound on the objective value.

First, we verify the completeness of the certificate:

Lemma 3.3.2. *Let $\mathbf{A}$ be a symmetric order-k tensor for even k, and let A be the natural matrix flattening of $\mathbf{A}'$ the asymmetrization of $\mathbf{A}$ described in Algorithm 3.3.1. Let $\mathcal{S}_{dk/2}$ be the symmetric group on $dk/2$ elements, and further let $\hat{\mathcal{S}}_{dk/2}$ be the set of $n^{dk/2} \times n^{dk/2}$ matrices that apply the permutations of $\mathcal{S}_{dk}$ to matrices whose rows and columns are identified with multisets in $[n]^{dk}$. Then*

$$\left\| \mathop{\mathbb{E}}_{\Pi,\Sigma \in \hat{\mathcal{S}}_{dk}} \left[\Pi (A^{\otimes d}) \Sigma\right] \right\|^{1/d} \geq \|\mathbf{A}\|_{inj} .$$

Proof. The proof is identical to that of Proposition 3.2.2, up to noticing that $\langle \mathbf{A}, x^{\otimes k} \rangle = \langle \mathbf{A}', x^{\otimes k} \rangle$. □

Now, we will prove that in the case that $\mathbf{A}$ is a random tensor with i.i.d. subgaussian entries, our certification algorithm improves smoothly upon the simple spectral algorithm as we invest more computational resources.

Theorem 3.3.3. *Let $n, k, d \in \mathbb{N}$, with even k. Let $\mathbf{A}$ be a symmetric order-k tensor with independent entries distributed symmetrically about 0. Let A be the matrix flattening of $\mathbf{A}'$, the asymmetrization of $\mathbf{A}$ described in Algorithm 3.3.1. Then if $d \ll n^{1/3k^2}$, there exists a constant c such that with high probability over $\mathbf{A}$,*

$$\left\| \mathop{\mathbb{E}}_{\Pi,\Sigma} \left[\Pi A^{\otimes d} \Sigma\right] \right\|^{1/d} \leq (c \log^2 n)^k \cdot d^{\frac{k^2 \log d}{4 \log n}} \cdot \frac{n^{k/4}}{d^{(k-2)/4}} .$$

The proof is nearly identical to the $k = 4$ case from Section 3.2, so we will be brief.

Proof. We will assume that each entry of $\mathbf{A}'$ is bounded in absolute value by $\gamma = O(\sqrt{d \log n})$, as by the subgaussian assumption this is true with high probability, even after symmetrization. This assumption preserves the symmetry of the distribution.

As in the proof of the $k = 4$ case from Section 3.2, we will use the trace power method. For shorthand, let $C \stackrel{\text{def}}{=} \mathbb{E}_{\Pi,\Sigma\in\mathcal{S}_{dk}}\left[\Pi A^{\otimes k}\Sigma\right]$. Let $\mathcal{H}$ be the set of all hyperedge configurations possible (the set of all possible length-2ℓ sequences of hypergraph matchings on two sets of $dk/2$ vertices). Let $\mathcal{V}$ be the set of all vertex configurations possible (the set of all possible length-2ℓ sequences of vertex multisets $I_1, \ldots, I_{2\ell} \in [n]^{dk/2}$). We note now that there are not many vertex configurations which use few vertices in $[n]$:

Fact 3.3.4. *Let $\mathcal{V}_\alpha$ be the set of vertex configurations on $dk\ell$ vertices containing fewer than $\alpha dk\ell/2$ distinct vertices from $[n]$. Then*

$$|V_\alpha| \le (\alpha dk\ell/2)^{(1-\alpha/2)dk\ell} \cdot n^{\alpha dk\ell/2}.$$

Proof. There are only $n^{\alpha dk\ell/2}$ choices for vertex labels, and then $(\alpha dk\ell/2)^{(1-\alpha/2)dk\ell}$ choices for the rest. □

For $H \in \mathcal{H}$ and $V \in \mathcal{V}$, we let $w_{\mathbf{A}}(V, H)$ denote the product of all hyperedge weights in the hyperedge cycle (V, H) when the weights are given by entries of the tensor $\mathbf{A}$. Because the entries are distributed symmetrically about 0, we have that

$$\begin{aligned}\mathop{\mathbb{E}}_{\mathbf{A}}\left[\mathrm{Tr}((CC^\top)^\ell)\right] &= \sum_{V\in\mathcal{V}} \mathop{\mathbb{E}}_{H\in\mathcal{H}}\left[\mathop{\mathbb{E}}_{\mathbf{A}}[w_{\mathbf{A}}(V,H)]\right] \\ &\le \sum_{V\in\mathcal{V}} \mathop{\mathbb{E}}_{H\in\mathcal{H}}\left[\gamma^{2d\ell}\cdot\mathbb{I}[(V,H)\text{ even}, \neq 0\,]\right] \;=\; \gamma^{2d\ell}\cdot\sum_{V\in\mathcal{V}}\mathop{\mathbb{P}}_{H\in\mathcal{H}}[(V,H)\text{ even}, \neq 0\,],\end{aligned}$$

where $\mathbb{I}[\cdot]$ is the $0-1$ indicator for an event. Notice that now, evenness is not enough to ensure that we have nonzero contribution–because we asymmetrized $\mathbf{A}$, every hyperedge also has to be lexicographically first, meaning it appears either as $\mathbf{A}'_{S,T}$ or $\mathbf{A}'_{T,S}$ depending on whether it comes from a C or $C^\top$ term. Using Fact 3.3.4 to argue that the number of vertex configurations with fewer than $(1-\beta)dk\ell/2$ distinct vertices (the number of $V \in \mathcal{V}_{(1-\beta)}$) cannot be too large,

$$\le \gamma^{2d\ell}\left(\left(\left(\frac{1+\beta}{2}dk\ell\right)^{(1+\beta)}n^{(1-\beta)}\right)^{dk\ell/2} + \sum_{V\notin\mathcal{V}_{(1-\beta)}}\mathop{\mathbb{P}}_{H\in\mathcal{H}}[(V,H)\text{ even}, \neq 0\,]\right). \tag{3.3.1}$$

So for a fixed $V \in \mathcal{V}$, we will bound $\mathbb{P}_H[(V, H)$ even, $\neq 0]$.

To do this, we repeat our argument from Section 3.2. Fixing a vertex configuration $V = I_1, \ldots, I_{2\ell}$, we sample $H \sim \mathcal{H}$ uniformly in two steps:

1. Sample a random perfect matching (of edges, not hyperedges) between every two consecutive vertex sets I_i, I_{i+1}, letting the configuration of edges we chose be E from the set of all such possible configurations $\mathcal{M}$.
2. Group the edges between I_i and I_{i+1} into groups of size $k/2$, and merge every group into a hyperedge (of order k).

Let (V, E) be the intermediate graph in this process that produces the hypergraph (V, H). Notice that now, We restate, then prove, a more precise version of Lemma 3.2.7

Lemma 3.3.5 (formal version of Lemma 3.2.7)**.** *Let $V = I_1, \dots, I_{2\ell} \in [n]^{dk/2}$, and suppose we have sampled hyperedges $H \in \mathcal{H}$ by first sampling simple edges $E \in \mathcal{M}$ as in step 1 and then grouping them into groups of $k/2$ as in step 2. Then*

$$\mathbb{P}((V, E) \text{ even} \mid (V, H) \text{ even}, \neq 0) \geq \left(\frac{1}{\frac{k}{2}!}\right)^{2d\ell}.$$

Proof. Suppose every hyperedge in H is lexicographically first and has even multiplicity. Each hyperedge h in H, h was sampled from one of the $(k/2)!$ matchings of its left-hand vertices to its right-hand vertices with equal probability. Let $h_1, \dots, h_m$ be the distinct labeled hyperedges of our hypergraph. Since all our hyperedges are lexicographically first, the same bipartition of vertices is common to every appearance of h_i for all $i \in [m]$. Thus, if we choose a uniformly random perfect matching of simple edges in each hyperedge of the hypergraph, we choose the same simple matching for all copies of h_i with probability at least $(\frac{k}{2}!)^{-\#h_i}$. It follows that if all hyperedges appear in (V, H) with even multiplicity, then with probability at least $(\frac{k}{2}!)^{-2d\ell}$ all simple edges in (V, E) appear with even multiplicities. □

Applying Lemma 3.3.5,

$$\begin{aligned}\mathbb{P}((V, H) \text{ even}, \neq 0) &= \frac{\mathbb{P}((V, H) \text{ even}, \neq 0 \ \&\ \ (V, E) \text{ even})}{\mathbb{P}((V, E) \text{ even} \mid (V, H) \text{ even}, \neq 0)} \\ &\leq \left(\frac{k}{2}!\right)^{2d\ell} \cdot \mathbb{P}((V, E) \text{ even}) \cdot \mathbb{P}((V, H) \text{ even} \mid (V, E) \text{ even}). \qquad (3.3.2)\end{aligned}$$

We now relate the quantity $\sum_{V \in \mathcal{V}} \mathbb{P}_E[(V, E) \text{ even}]$ to a matrix quantity we can control well. Letting B be an $n \times n$ matrix with symmetric i.i.d. entries uniform from $\{\pm 1\}$, and letting $C' = \mathbb{E}[\Pi B^{\otimes dk/2} \Sigma]$,

$$\mathbb{E}\left[\mathrm{Tr}((C'C'^{\top})^{\ell})\right] = \sum_{V \in \mathcal{V}} \mathop{\mathbb{P}}_{E}[(V, E) \text{ even}].$$

We now prove and apply the following proposition, which is a restatement of Lemma 3.2.6 for arbitrary k:

Proposition 3.3.6. *Let $n, d, k, \ell \in \mathbb{N}$ so that $dk\ell \log n \ll n$. Let $C' = \mathbb{E}_{\Pi, \Sigma \in \mathcal{S}_{dk/2}}\left[\Pi B^{\otimes dk/2} \Sigma\right]$, for an $n \times n$ matrix B with i.i.d. Rademacher entries. Then*

$$\mathbb{E}\left[\mathrm{Tr}((C'C'^{\top})^{\ell})\right] \leq 2^{4dk\ell+1} n^{dk\ell/2+dk/2}.$$

Proof. Let B be an $n \times n$ matrix with i.i.d. Rademacher entries, and let $d, \ell \in \mathbb{N}$. We have that for any $N \times N$ PSD matrix P, $\mathrm{Tr}\left(P^{\ell}\right) \leq N \cdot \left\|P^{\ell}\right\|$, and because $C'C'^{\top}$ is PSD it follows that

$$\mathrm{Tr}\left((C'C'^{\top})^{\ell}\right) \leq n^{dk/2} \cdot \left\|(C'C'^{\top})^{\ell}\right\|. \qquad (3.3.3)$$

We will get a bound on $\mathbb{E}\,\|(C'C'^{\top})^{\ell}\|$. Because C' is symmetric, $C'C'^{\top} = (C')^2$. Thus, a bound on $\mathbb{E}\,\|C'^{2\ell}\|$ will suffice. We apply the triangle inequality and the submultiplicativity of the norm to deduce that for any B,

$$\|C'^{2\ell}\| \;=\; \left\| \left(\mathop{\mathbb{E}}_{\Pi,\Sigma\in\hat{\mathcal{S}}_{2d}} \left[\Pi(B^{\otimes dk/2})\Sigma\right] \right)^{2\ell} \right\| \leq \left(\mathop{\mathbb{E}}_{\Pi,\Sigma\in\hat{\mathcal{S}}_{2d}} \left[\|\Pi\| \cdot \|(B^{\otimes dk/2})\| \cdot \|\Sigma\|\right] \right)^{2\ell} \;\leq\; \|B\|^{dk\ell},$$

and now, we can use standard arguments from random matrix theory to get tail bounds on $\|B\|$. From Theorem A.3.2, we have that $\mathbb{P}[\|B\| - 12n^{1/2} \geq s] \leq \exp(-s^2/16)$, and we also have that $\|B\| \leq \|B\|_F \leq n$, and thus it follows that

$$\begin{aligned}\mathbb{E}\left[\|C'^{2\ell}\|\right] \;\leq\; \mathbb{E}\left[\|B\|^{dk\ell}\right] &\leq \mathbb{P}[\|B\| \leq 16\sqrt{n}] \cdot (16\sqrt{n})^{dk\ell} + \mathbb{P}[\|B\| > 16\sqrt{n}] \cdot n^{dk\ell} \\ &\leq (1 - \exp(-n)) \cdot (16\sqrt{n})^{dk\ell} + \exp(dk\ell \log n - n) \;\leq\; 2(16\sqrt{n})^{dk\ell},\end{aligned}$$

and the conclusion follows from combining the above with (3.3.3). □

We thus have

$$\sum_{V\in\mathcal{V}_{(1-\beta)}} \mathop{\mathbb{P}}_{E}[(V,E) \text{ even}] \leq \sum_{V\in\mathcal{V}} \mathop{\mathbb{P}}_{E}[(V,E) \text{ even}] \leq \mathbb{E}\left[\mathrm{Tr}((C'C'^{\top})^{\ell})\right] \leq 2^{4dk\ell+1} n^{dk\ell/2+dk/2}. \tag{3.3.4}$$

Now, from (3.3.2) we are left to bound $\mathbb{P}[(V,H) \text{ even} \mid (V,E) \text{ even},\ V \in \mathcal{V}_{(1-\beta)}]$. We apply the following lemma:

Lemma 3.3.7. *Let $m, n, c \in \mathbb{N}$, and let G be a graph which is a union of at most c disjoint cycles. Suppose furthermore that each vertex receives labels from the set $[n]$, that every labeled edge appears with even multiplicity, and that there are exactly m distinct labeled edges. Then letting L be the number of distinct vertex labels, we have*

$$L \leq m + c.$$

The proof of Lemma 3.3.7 proceeds by a cute inductive argument, which we will reserve for Section 3.3.

Lemma 3.3.7 implies that if (V,E) has at least $(1-\beta)dk\ell/2$ distinct vertices, then it must have at least $(1-\beta)dk\ell/2 - dk/2$ distinct edges. Let $E_1, \ldots, E_{2\ell}$ be the matchings in E so that E_i gives the edges between I_i, I_{i+1}. We invoke and prove a generalization of Lemma 3.2.8:

Lemma 3.3.8. *Fix $M, r, \ell, N \in \mathbb{N}$ and $\beta \in (0,1)$. Let $E_1, \ldots, E_M \in [N]^{r\cdot c}$ be multisets of elements such that the number of distinct elements in the union $\cup_{i\in[M]} E_i$ is at least $(1-\beta)M \cdot r \cdot c/2$. Let G_i denote a uniformly random r-grouping of elements within E_i, sampled independently for each $i \in [M]$. Let $\bigoplus_i G_i$ denote the set of r-groups $(a_1, \ldots, a_r) \in [N]^r$ that appear an odd number of times within $\cup_i G_i$. Then for any $0 < \delta < 3.5\beta$,*

$$\mathbb{P}[|\oplus_i G_i| \leq \delta Mc] \leq \left(\frac{112}{\beta c}\right)^{(1-(4r+1)\beta-2\delta)(r-1)Mc/2}$$

We will prove Lemma 3.3.8 in Section 3.3.

We apply Lemma 3.3.8 to the multisets E_i with parameters $M \leftarrow 2\ell$, $c \leftarrow d$, $r \leftarrow k/2$, to conclude that if the S_i are each grouped into matchings of hyperedges with d edges each, then

$$\begin{aligned}\mathbb{P}\left[(V,H) \text{ even} \mid (V,E) \text{ has} \geq (1-\beta)dk\ell/2 \text{ edges }\right] &\leq \left(\frac{112}{\beta d}\right)^{(1-(2k+1)\beta)(k/2-1)d\ell} \\ &\leq c_\beta^{dk\ell/2}\left(\frac{1}{d}\right)^{(1-3k\beta)(k/2-1)d\ell}.\end{aligned}$$

for some constant c_β depending only on β. Putting this together with (4.4.2),(3.3.2), and (3.3.4),

$$\begin{aligned}&\mathop{\mathbb{E}}_{\mathbf{A}}\left[\mathrm{Tr}((CC^\top)^\ell)\right] \\ &\leq \gamma^{2d\ell}\left(\left(((1+\beta)dk\ell/2)^{(1+\beta)}\cdot n^{(1-\beta)}\right)^{\frac{dk\ell}{2}} + \left(\frac{k}{2}!\right)^{2d\ell}(2^8 n)^{\frac{dk\ell+dk}{2}} c_\beta^{\frac{dk\ell}{2}}\left(\frac{1}{d}\right)^{(1-3k\beta)(k/2-1)d\ell}\right) \\ &\leq (c'_\beta\cdot k^k\gamma^2\ell^k)^{d\ell}n^{dk/2}\left((d^{1+\beta}n^{1-\beta})^{dk\ell/2} + \left(\frac{1}{d}\right)^{(1-3k\beta)(k/2-1)d\ell} n^{dk\ell/2}\right)\end{aligned}$$

for some constant c'_β. Choosing $\beta = \frac{2(k-1)\log d}{k(3k-7)\log d+\log n}$ balances the terms, so for smaller β we have

$$\mathop{\mathbb{E}}_{\mathbf{A}}\left[\mathrm{Tr}((CC^\top)^\ell)\right] \leq 2(c'_\beta\cdot k^k\gamma^2\ell^k)^{d\ell}n^{dk/2}\cdot\left(\frac{n^{k/2}}{d^{k/2-1}}\right)^{d\ell}\cdot d^{\beta(k/2-1)d\ell}.$$

Now, requiring that $d \leq n^{1/3k^2}$ and choosing $\beta \leftarrow (k-1)\frac{\log d}{\log n}$, we have that

$$\mathbb{E}\left[\mathrm{Tr}((CC^\top)^\ell)\right]^{1/2\ell} \leq 2(c'_\beta\cdot k^k\gamma^2\ell^k)^{d/2}n^{dk/4\ell}\cdot\left(\frac{n^{k/2}}{d^{k/2-1}}\right)^{d/2}\cdot d^{\frac{k^2\log d}{2\log n}\cdot d/2}$$

Taking $\ell = O(\log n)$ and applying Proposition 3.2.4, the conclusion of Theorem 3.3.3 follows. □

Odd-Order Tensors

In this section, we give our algorithm for certifying bounds on the injective tensor norm of random odd-order tensors. Because there is no canonical way to flatten an odd-order tensor to a square matrix, the algorithm includes an additional step, similar to the one we employ for k-XOR instances when k is odd (Section 3.4). We remark that this additional step is not new, and has appeared before (as early as e.g. [FG01])—however it does introduce some new challenges in our analysis.

We begin with a brief high-level overview of our algorithm. To begin with, let $\mathbf{A} \in \mathbb{R}^{[n]^k}$ be an order-k symmetric tensor of dimension n. For convenience, we define an integer κ such

that $k = 2\kappa + 1$. For the rest of this section, we will use A_i to denote the $[n]^\kappa \times [n]^\kappa$ matrix obtained by flattening the i^{th} slice of $\mathbf{A}$, i.e.,

$$A_i(I, J) \stackrel{\text{def}}{=} \mathbf{A}_{(i,I,J)} \qquad \forall I, J \in [n]^\kappa .$$

Using the Cauchy-Schwarz inequality, we can bound the injective norm in terms of the matrices A_i,

$$\begin{aligned}\langle x^{\otimes 2\kappa+1}, \mathbf{A}\rangle &= \sum_i x_i \cdot \langle x^{\otimes\kappa}, A_i x^{\otimes\kappa}\rangle \\ &\leq \left(\sum_i x_i^2\right)^{1/2} \cdot \left(\sum_i \langle x^{\otimes\kappa}, A_i x^{\otimes\kappa}\rangle^2\right)^{1/2} = \left(\langle x^{\otimes 2\kappa}, \left(\sum_i A_i \otimes A_i\right) x^{\otimes 2\kappa}\rangle\right)^{1/2} .\end{aligned} \tag{3.3.5}$$

Therefore, in order to bound $\|\mathbf{A}\|_{inj}$, it is sufficient to bound the following quantity.

$$\max_{\|x\|\leq 1} \left\langle x^{\otimes 2\kappa}, \left(\sum_i A_i \otimes A_i\right) x^{\otimes 2\kappa} \right\rangle \tag{3.3.6}$$

For a tensor $\mathbf{A}$ whose entries are i.i.d. subgaussian variables, we bound the value of the maximization problem (3.3.6).

The matrix $\sum_i A_i \otimes A_i$ has large diagonal entries. However, our tensoring and symmetrizing algorithm requires a matrix with eigenvalues roughly symmetric about 0 (see the heuristic explanation in Section 3.2). Thus, we will work with a diagonal-free version of the matrix. Define the matrix $N \in \mathbb{R}^{[n]^{2\kappa}\times[n]^{2\kappa}}$ as follows:

$$N_i((a, b), (c, d)) = A_i(a, c) \cdot A_i(b, d) \cdot \mathbb{I}[(a, c) \neq (b, d)] \qquad \forall a, b, c, d \in [n]^\kappa$$

We can rewrite the polynomial in (3.3.6) as,

$$\left\langle x^{\otimes 2\kappa}, \left(\sum_i A_i \otimes A_i\right) x^{\otimes 2\kappa} \right\rangle = \left\langle x^{\otimes 2\kappa}, \left(\sum_i N_i\right) x^{\otimes 2\kappa} \right\rangle + \sum_{i\in[n],a,b\in[n]^\kappa} x_a^2 x_b^2 \mathbf{A}_i^2(a, b)$$

And we can upper bound the latter term by

$$\sum_{i\in[n],a,b\in[n]^\kappa} x_a^2 x_b^2 \mathbf{A}_i^2(a, b) \;\leq\; \sum_{a,b\in[n]^\kappa} x_a^2 x_b^2 \left(\sum_i \mathbf{A}_i^2(a, b)\right) \;\leq\; \max_{a,b} \left(\sum_i \mathbf{A}_i^2(a, b)\right) \tag{3.3.7}$$

where we have used the fact that $\|x\|^2 = 1$. Bounding the norm of tensor $\mathbf{A}$ thus reduces to upper bounding $\langle x^{\otimes 2\kappa}, (\sum_i N_i)\, x^{\otimes 2\kappa}\rangle$. Now our strategy is as before–we take a dth tensor power of our matrix, then average over the symmetries of $x^{\otimes 2\kappa d}$.

Having discussed the differences between the even and odd cases, we are ready to give our algorithm.

Algorithm 3.3.9 (Odd-order Injective tensor norm)**.**
Input: A random tensor $\mathbf{A}$ of dimension n and odd order $k = 2\kappa + 1$, and a parameter d.

1. Form the asymmetric tensor $\mathbf{A}'$ as described in Algorithm 3.3.1, so that $\langle x^{\otimes k}, \mathbf{A}\rangle = \langle x^{\otimes k}, \mathbf{A}'\rangle$ but only lexicographically first entries are nonzero.
2. Let A_i be the $n^\kappa \times n^\kappa$ matrix flattening of the ith slice of $\mathbf{A}'$, and form the matrix
$$M := \sum_{i\in[n]} A_i \otimes A_i$$
3. Zero out all entries of the matrix corresponding to $(I_1, I_2), (J_1, J_2) \in [n]^{2\kappa}$ such that $(I_1, J_1) = (I_2, J_2)$, forming a new matrix N:
$$N_{(I_1,I_2),(J_1,J_2)} := M_{(I_1,I_2),(J_1,J_2)} \cdot \mathbb{I}((I_1, J_1) \neq (I_2, J_2)) \,.$$
4. Take the dth tensor power of N,
$$N \to N^{\otimes d} \,.$$
5. Symmetrize the rows and columns of $N^{\otimes d}$ according to the symmetries of $\mathcal{S}_{2d\kappa}$ to obtain the matrix C,
$$C_d \overset{\text{def}}{=} \mathop{\mathbb{E}}_{\Pi,\Sigma\in\hat{\mathcal{S}}_{2d\kappa}} \left[\Pi(N^{\otimes d})\Sigma\right] \,.$$

Output: The quantity $\left(\|C_d\|^{1/d} + \max_{a,b\in[n]^\kappa} \sum_{i\in[n]} A_i(a,b)^2\right)^{1/2}$ as an upper bound on $\|\mathbf{A}\|_{inj}$.

Proposition 3.3.10. *For any symmetric tensor* $\mathbf{A}$*, Algorithm 3.3.9 outputs a valid upper bound on* $\|\mathbf{A}\|_{inj}$*.*

Proof. Our asymmetrization in step 1 ensures that $\langle x^{\otimes k}, \mathbf{A}\rangle = \langle x^{\otimes k}, \mathbf{A}'\rangle$. The proof then follows from the calculations above, beginning at (3.3.5) and ending at (3.3.7), and then using that the symmetrization step fixes vectors of the form $x^{\otimes 2d\kappa}$. □

We prove that when $\mathbf{A}$ has subgaussian, centered, independent entries, Algorithm 3.3.9 improves over the basic spectral algorithm.

Theorem 3.3.11. *For any symmetric tensor* $\mathbf{A}$ *with independent subgaussian centered entries, with high probability over the choice of* $\mathbf{A}$*, Algorithm 3.3.9 certifies that*
$$\|\mathbf{A}\|_{inj} \leq \tilde{O}\left(\frac{n^{k/4}}{k^{(k-2)/4}} \cdot d^{\frac{k^2 \log d}{2 \log n}}\right) \,.$$
so long as $d \log n \ll n^{1/120}$*.*

First, the very straightforward observation that subtracting the maximum element cannot have too strong of a negative effect:

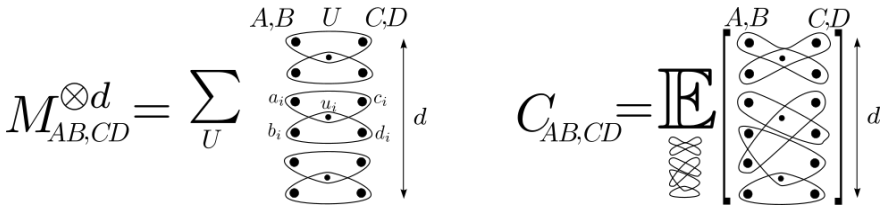


Figure 3.2: Hypergraphs corresponding to odd certificate entries.

Lemma 3.3.12. *If* $\mathbf{A}$ *is an order-*D *tensor with i.i.d. symmetric subgaussian entries, then*

$$\max_{a,b\in[n]^d} \sum_{i\in[n]} \mathbf{A}(i,a,b)^2 \leq O(n\log n),$$

with high probability.

Proof. The lemma follows from the fact that the variables are subgaussian, and by applying first a Chernoff bound and then a union bound over the indices. □

Now, we bound the norm of the matrix $\|C_d\|$.

Theorem 3.3.13. *So long as* $kd\ell < 4n^{\beta/4}$*, there exists some absolute constant* c_β *depending on* β *such that with high probability over the choice of* $\mathbf{A}$*,*

$$\|C_d\| \leq \left(c_\beta^d \log n \cdot \frac{n^{k/4}}{d^{(k-2)/4-6k\beta}} \cdot n^{1/2\ell}\right)^d .$$

Proof. Because the entries of $\mathbf{A}$ are subgaussian, with high probability all entries of the tensor are bounded in magnitude by $\gamma = O(\sqrt{\kappa \log n})$. We will assume this to be the case in the remainder of the proof.

We bound the expected trace $\mathbb{E}[\mathrm{Tr}\left((CC^\top)^\ell\right)]$ over the choice of $\mathbf{A}$, in order to apply the tensor power method. Let $M := (\sum_i A_i \otimes A_i)^{\otimes d}$ for convenience. The $(A,B),(C,D)$th entry of M (for $A,B,C,D \in [n]^{d\kappa}$ with $A = a_1,\ldots,a_d$ with $a_i \in [n]^\kappa$, and with similar decompositions defined for B,C,D) has value

$$M_{(A,B),(C,D)} = \prod_{i\in[d]} \left(\sum_{u\in[n]} \mathbf{A}_{a_i,c_i,u} \cdot \mathbf{A}_{b_i,d_i,u}\right) = \sum_{U\in[n]^d} \prod_{i\in[d]} \left(\mathbf{A}_{a_i,c_i,u_i} \cdot \mathbf{A}_{b_i,d_i,u_i}\right) .$$

Interpreting the variables $\mathbf{A}_{a_i,c_i,u_i}$ as $k = (2\kappa+1)$-uniform hyperedges, we have that each entry is a sum over hypergraphs indexed by $U \in [n]^d$. For each $U \in [n]^d$, we have a hypergraph on the following vertex configuration: on the left, we have the vertices from the multiset A,B. On the right, we have the vertices from the multiset C,D. In the center, we have the vertices from U. On this vertex set, we have $2d$ hyperedges. Of these hyperedges, d form a tripartite matching on the vertices in A,C,U, with κ vertices from each of A,C and one vertex in U. The other d form a similar tripartite matching on the vertices in B,D,U. Every hyperedge on A,C,U shares exactly one vertex in U with exactly one hyperedge from B,D,U. See Figure 3.2 for an illustration.

The subtraction of the square terms squares$(A_u \otimes A_u)$ forces us to never have two hyperedges sharing a vertex in U if they contain vertices of the same type in $[n]$: that is, we can never have $(a_i, c_i) = (b_i, d_i)$ as ordered multisets. Then, the averaging operation $\mathbb{E}_{\Pi,\Sigma\in\hat{\mathcal{S}}_{2d\kappa}}$ takes each such entry to an average over all allowed hyperedge configurations on the vertex set $(A, B), (C, D), U$.

When we take $\mathrm{Tr}(C_{(d)}C_{(d)}^\top)^\ell$, we are taking a sum over all "cycles" of length 2ℓ in such hypergraphs, where the vertices in the cycle are given by the (A, B) multisets, and the edges are given by the average hyperedge configuration between (A, B) and the next (C, D), with the U vertices in between.

To this end, we describe an equivalent definition of the matrix $C_{(d)}$. Specifically, given $a, b \in [n]^{2\kappa d}$ the entry $C_{(d)}(a, b)$ can be evaluated as follows:

1. Sample a random matching $\mathcal{E} = \{e_1, \ldots, e_{2\kappa d}\}$ between the multisets a and b.
2. Group the edges of $\mathcal{E}$ in to $2d$ groups of size κ, to obtain $2d$ *blocks* $\mathcal{F} = \{f_1, \ldots, f_{2d}\}$.
3. Pick a random matching $\mathcal{M}$ between the blocks in $\mathcal{F}$. Let $\mathcal{M}$ be given by d pairs $\{(h_i, h_i')\}_{i\in[d]}$.
4. For each choice of *"pivot vertices"* $\sigma \in [n]^d$, we get a $(2\kappa + 1)$-uniform hypergraph $\mathcal{H}_\sigma$ with $2d$ hyperedges given by
$$\{(\sigma_i, h_i), (\sigma_i, h_i') | i \in [d]\}.$$
5. Output the value $\sum_{\sigma\in[n]^d} \prod_{i\in[d]} A_{(\sigma_i,h_i)} \cdot A_{(\sigma_i,h_i')} \cdot \mathbb{I}[h_i \neq h_i']$.

The entries of the matrix C are given by,

$$C(a, b) = \mathop{\mathbb{E}}_{\mathcal{E}} \mathop{\mathbb{E}}_{\mathcal{F}} \mathop{\mathbb{E}}_{\mathcal{M}} \sum_{\sigma\in[n]^d} \left[\prod_{i\in[d]} T_{(\sigma_i,f_i)} \cdot T_{(\sigma_i,g_i)} \cdot \mathbb{I}[h_i \neq h_i'] \right]$$

Returning to the quantity $\mathrm{Tr}((CC^\top)^\ell)$, we can understand this as a sum over cycles in the entries of C, which gives us a sum over products of random variables corresponding to the edges in cyclic hypergraphs. Since we have assumed the entries of $\mathbf{A}$ are distributed symmetrically about 0, each term in the sum $\mathbb{E}\,\mathrm{Tr}((CC^\top)^\ell)$ is non-zero only if every hyperedge appears with even multiplicity.

We can organize the terms in $\mathrm{Tr}\left((CC^\top)^\ell\right)$ as follows:

- For each vertex configuration $V = \{a_1, b_1, a_2, \ldots, b_\ell, a_1\} \in \mathcal{V} \subset [n]^{2\kappa d}$
  1. Sample matchings $\mathcal{E} = \{\mathcal{E}_1, \ldots, \mathcal{E}_{2\ell}\}$
  2. Group the edges in to *blocks* $\mathcal{F} = \{\mathcal{F}_1, \ldots, \mathcal{F}_{2\ell}\}$
  3. Pick random matchings $\mathcal{M} = \{\mathcal{M}_1, \ldots, \mathcal{M}_{2\ell}\}$ between the blocks.
  4. For each choice of *"pivots"* $\sigma = \{\sigma_1, \ldots, \sigma_{2\ell}\} \subset [n]^d$ we get a $(2\kappa + 1)$-uniform hypergraph $\mathcal{H}_\sigma$ with $2d\ell$ hyperedges.

We will call the hypergraph $\mathcal{H}_\sigma$ *diagonal-free* (or *d-free*) if there are no pairs of identical blocks matched with each other in $\mathcal{M}$. We will use the notation $\|\cdot\|_\oplus$ to denote the number of elements of odd multiplicity in a multiset, and similarly the notation $\|\cdot\|_0$ to denote the number of distinct elements in a multiset. We will say $\mathcal{H}_\sigma$ is *even* if the number of

occurrences of each hyperedge is even. Now, dividing by our upper bound on the absolute value of the maximum entry,

$$\begin{aligned}
&\gamma^{-2\kappa d\ell} \mathop{\mathbb{E}}_{T} \left[\mathrm{Tr}\left((CC^\top)^\ell\right)\right] \\
&\quad\le \sum_{V\in\mathcal{V}} \mathop{\mathbb{E}}_{\mathcal{E}} \mathop{\mathbb{E}}_{\mathcal{F}} \mathop{\mathbb{E}}_{\mathcal{M}} \left[\sum_{\sigma} \mathbb{I}[\mathcal{H}_\sigma \text{ even \& d-free}]\right] \\
&\quad\le (\kappa!)^{4d\ell} \cdot \sum_{V\in\mathcal{V}} \mathop{\mathbb{E}}_{\mathcal{E}} \mathop{\mathbb{E}}_{\mathcal{F}} \mathop{\mathbb{E}}_{\mathcal{M}} \left[\sum_{\sigma} \mathbb{I}[\mathcal{H}_\sigma \text{ even \& d-free}] \cdot \mathbb{I}[\|\mathcal{E}\|_\oplus = 0]\right] \qquad \text{(by Lemma 3.2.7)} \\
&\quad= (\kappa!)^{4d\ell} \cdot \sum_{V\in\mathcal{V}} \mathop{\mathbb{E}}_{\mathcal{E}} \mathbb{I}\left[\|\mathcal{E}\|_\oplus = 0\right] \mathop{\mathbb{E}}_{\mathcal{F}} \mathop{\mathbb{E}}_{\mathcal{M}} \left[\sum_{\sigma} \mathbb{I}\left[\mathcal{H}_\sigma \text{ even \& d-free}\right]\right] \\
&\quad\le (\kappa!)^{4d\ell} \cdot \sum_{V\in\mathcal{V}} \mathop{\mathbb{E}}_{\mathcal{E}} \mathbb{I}\left[\|\mathcal{E}\|_\oplus = 0 \wedge \|\mathcal{E}\|_0 \ge 2d\kappa\ell(1-\beta)\right] \mathop{\mathbb{E}}_{\mathcal{F}} \mathop{\mathbb{E}}_{\mathcal{M}} \left[\sum_{\sigma} \mathbb{I}[\mathcal{H}_\sigma \text{ even \& d-free}]\right] \qquad (3.3.8) \\
&\quad+ (\kappa!)^{4d\ell} \cdot \sum_{V\in\mathcal{V}} \mathop{\mathbb{E}}_{\mathcal{E}} \mathbb{I}\left[\|\mathcal{E}\|_\oplus = 0 \wedge \|\mathcal{E}\|_0 \le 2d\kappa\ell(1-\beta)\right] \mathop{\mathbb{E}}_{\mathcal{F}} \mathop{\mathbb{E}}_{\mathcal{M}} \left[\sum_{\sigma} \mathbb{I}[\mathcal{H}_\sigma \text{ even \& d-free}]\right] \qquad (3.3.9)
\end{aligned}$$

First we will bound the value of term in (3.3.9). Recall that by Lemma 3.3.7, if $\mathcal{E}$ is even then the number of distinct labels in $V \in \mathcal{V}$ is less than $\|\mathcal{E}\|_0$. Therefore,

$$|\{\mathcal{E} \mid \|\mathcal{E}\|_\oplus = 0 \wedge \|\mathcal{E}\|_0 \le 2d\kappa\ell(1-\beta)\}| < n^{2d\kappa\ell(1-\beta)}(4d\kappa\ell)!$$

Now, we will use the following claim:

Claim 3.3.14. For every choice of $V \in \mathcal{V}, \mathcal{E}, \mathcal{F}, \mathcal{M}$,

$$\sum_{\sigma} \mathbb{I}[\|\mathcal{H}_\sigma\|_\oplus = 0 \wedge \mathcal{H}_\sigma \text{ is diagonal-free}] \le (2d\ell)! \cdot n^{d\ell}$$

Proof. If $\mathcal{H}_\sigma$ is diagonal-free and even, then we claim that each *pivot* value appears twice. Suppose not, if σ_i is such that $\sigma_i \neq \sigma_j$ for all $j \neq i$. Since $\mathcal{H}_\sigma$ is diagonal free, the two hyperedges involving σ_i are distinct. Since this is the unique occurrence of these two hyperedges in $\mathcal{H}_\sigma$, $\mathcal{H}_\sigma$ cannot be *even*–a contradiction. With each *pivot* appearing at least twice, the number of distinct choices of σ is at most $(2d\ell)!n^{d\ell}$. □

By Claim 3.3.14, for each $\mathcal{E}$ the corresponding term is at most $(2d\ell)!n^{d\ell}$. In all, this shows that (3.3.9) can be bounded as

$$(3.3.9) \le (\kappa!)^{4d\ell} \cdot n^{2d\kappa\ell(1-\beta)}(4d\kappa\ell)! \cdot \left((2d\ell)! \cdot n^{d\ell}\right) \le \left(\frac{(d\ell\kappa)^{5\kappa} \cdot n^{2\kappa+1}}{n^{2\kappa\beta}}\right)^{d\ell} \le \left(\frac{n^{2\kappa+1}}{d^{2\kappa-1}}\right)^{d\ell} \qquad (3.3.10)$$

where the final simplification uses $d\kappa\ell < n^\beta 4$.

Now we bound (3.3.8). Using Proposition 3.3.6 and reasoning similar to that in the proof of Theorem 3.3.3, by making an analogy between the set of configurations with even $\mathcal{E}$ and the norm of a random matrix under our tensoring and averaging operations, we know that

$$\sum_{V \in \mathcal{V}} \mathop{\mathbb{E}}_{\mathcal{E}} \mathbb{I}[\|\mathcal{E}\|_\oplus = 0] \leq c^{2\kappa d\ell} n^{2\kappa d\ell + \kappa d}$$

for an absolute constant $c > 0$. Moreover, conditioned on $\|\mathcal{E}\|_0 \geq 2d\kappa\ell(1-\beta)$, we will show the following bound

$$\mathop{\mathbb{E}}_{\mathcal{F}} \mathop{\mathbb{E}}_{\mathcal{M}} \sum_{\sigma \in [n]^{d\ell}} \mathbb{I}[\|\mathcal{H}_\sigma\|_\oplus = 0] \leq \left(\ell c_{\kappa\beta} \cdot \frac{n}{d^{(2\kappa-1)-8\kappa^2\beta}}\right)^{d\ell}$$

for a constant $c_{\kappa\beta}$ depending only on κ, β in Lemma 3.3.15. By the preceding pair of inequalities, we get that

$$(3.3.8) \leq \left(\ell c'_{\kappa\beta} \cdot \frac{n^{2\kappa+1}}{d^{2\kappa-1-8\kappa^2\beta}} \cdot n^{\kappa/\ell}\right)^{d\ell} \tag{3.3.11}$$

From (3.3.10) & (3.3.11), we conclude that

$$\left(\mathop{\mathbb{E}}_{T}\left[\mathrm{Tr}\left((CC^T)^\ell\right)\right]\right)^{1/2\ell} \leq \left(\ell c_{\kappa\beta} \cdot \frac{n^{\kappa+1/2}}{d^{\kappa-1/2-4\kappa^2\beta}} \cdot n^{\kappa/2\ell}\right)^{d}$$

By Proposition 3.2.4, taking $\ell = O(\log n)$, we conclude that

$$\mathbb{P}\left[\|C\| \leq \left(c'_{\kappa\beta} \log n \cdot \frac{n^{\kappa+1/2}}{d^{\kappa-1/2-4\kappa^2\beta}}\right)^{d}\right] \geq 1 - n^{-100},$$

□

We can now put together the easy bound on the maximum diagonal entry with the bound on $\|C\|$ to prove Theorem 3.3.11.

Proof of Theorem 3.3.11. Algorithm 3.3.9 returns the upper bound

$$\left(\|C\|^{1/d} + \max_{I,J}\left(\sum_i \mathbf{A}^2_{i,I,J}\right)\right)^{1/2}.$$

We combine Lemma 3.3.12 with Theorem 3.3.13, and we have that with high probability, for constants c_β and c_2,

$$\left(\|C\|^{1/d} + \max_{I,J}\left(\sum_i \mathbf{A}^2_{i,I,J}\right)\right)^{1/2} \leq \left(\log n \cdot c_{\kappa\beta} \cdot \frac{n^{\kappa+1/2}}{d^{\kappa-1/2-4\kappa^2\beta}} + c_2 n \log n\right)^{1/2} \tag{3.3.12}$$

By picking the best possible β under the constraint $\beta < 1/30$ and $d\kappa\ell < n^{\beta/4}$, we have that the former term always dominates, and we get the bound:

$$\|\mathbf{A}\|_{inj} \leq \tilde{O}\left(\frac{n^{(2\kappa+1)/4}}{d^{(2\kappa-1)/4}} \cdot d^{2\kappa^2 \frac{\log d}{\log n}}\right).$$

This concludes the proof. □

Now, we prove some of the lemmas we have relied upon in the proof of Theorem 3.3.13. We begin with a lemma bounding the probability that the hyperedges we sample all have even multiplicity.

Lemma 3.3.15. *Suppose $k < n^{\beta/4}$ and $\beta < 1/30$. Then conditioned on an $\mathcal{E}$ such that $\|\mathcal{E}\|_0 \geq 2kd\ell(1-\beta)$,*

$$\mathop{\mathbb{E}}_{\mathcal{F}} \mathop{\mathbb{E}}_{\mathcal{M}} \sum_{\sigma \in [n]^{k\ell}} \mathbb{I}[\|\mathcal{H}_\sigma\|_\oplus = 0] \leq \left(\ell c_{d\beta} \cdot \frac{n}{k^{(2d-1)-8d^2\beta}}\right)^{k\ell}$$

where $c_{d\beta}$ is a constant depending on β and d.

Proof. Note that $\|\mathcal{H}_\sigma\|_\oplus = 0$ implies that $\|\mathcal{F}\|_\oplus = 0$. By applying Lemma 3.3.8 with $r \leftarrow d$, $c \leftarrow 2k$, $M \leftarrow 2\ell$, and $E_i \leftarrow \mathcal{E}_i$, we obtain the following bound over the choice of $\mathcal{F}$.

$$\mathop{\mathbb{P}}_{\mathcal{F}}[\|\mathcal{F}\|_\oplus = 0 | \|\mathcal{E}\|_0 \geq 2kd\ell(1-\beta)] \leq \left(\frac{112}{2\beta k}\right)^{2(d-1)k\ell(1-(4d+1)\beta)} \tag{3.3.13}$$

Furthermore, if $\|\mathcal{E}\|_0 \geq 2kd\ell(1-\beta)$ then clearly $\|\mathcal{F}\|_0 \geq 2k\ell(1-\beta)$. By Lemma 3.3.16, for every $\mathcal{F}$ with $\|\mathcal{F}\|_0 \geq 2k\ell(1-\beta)$ we have,

$$\begin{aligned}\mathop{\mathbb{E}}_{\mathcal{M}} \sum_{\sigma \in [n]^{k\ell}} \mathbb{I}[\|\mathcal{H}_\sigma\|_\oplus = 0] &\leq (4\beta k\ell)! \cdot n^{k\ell} \cdot \left(\left(\frac{112}{\beta k}\right)^{k\ell(1-10\beta)} + n^{-\beta k\ell/3}\right) \\ &\leq \left(k^{4\beta}\ell^{4\beta} \cdot n \cdot \left(\left(\frac{112}{\beta k}\right)^{(1-10\beta)} + n^{-\beta/3}\right)\right)^{k\ell} \end{aligned} \tag{3.3.14}$$

Using (3.3.13) and (3.3.14) we conclude that,

$$\mathop{\mathbb{E}}_{\mathcal{F}} \mathop{\mathbb{E}}_{\mathcal{M}} \sum_{\sigma \in [n]^{k\ell}} \mathbb{I}[\|\mathcal{H}_\sigma\|_\oplus = 0] \leq \left(\left(\frac{112}{2\beta k}\right)^{2(d-1)(1-(4d+1)\beta)} \cdot k^{4\beta}\ell^{4\beta} \cdot n \cdot \left(\left(\frac{112}{\beta k}\right)^{(1-10\beta)} + n^{-\beta/3}\right)\right)^{k\ell}$$

Since $k < n^{\beta/4}$ and $\beta < 1/30$, we have that $k^{1-14\beta} << n^{\beta/3}$, and so the first term in the latter parenthesis dominates. This implies that,

$$\mathop{\mathbb{E}}_{\mathcal{F}} \mathop{\mathbb{E}}_{\mathcal{M}} \sum_{\sigma \in [n]^{k\ell}} \mathbb{I}[\|\mathcal{H}_\sigma\|_\oplus = 0] \leq \left(c_{d\beta}\ell \cdot \frac{n}{k^{(2d-1)-8d^2\beta}}\right)^{k\ell}$$

where $c_{d\beta}$ is a constant depending on d and β, and where we have used the fact that $8d^2 \geq 8d^2 - 6d + 4$ for all $d \geq 1$.

□

The following lemma we employ in bounding the probability that our blocks from $\mathcal{F}$ are matched in a way that gives hyperedges with even multiplicity. We do this via reducing the problem to counting the number of multigraphs with labeled edges in which every subgraph induced by a given label is Eulerian.

Lemma 3.3.16. *For every $\mathcal{F}$ with $\|\mathcal{F}\|_0 \geq 2k\ell(1-\beta)$,*

$$\mathop{\mathbb{E}}_{\mathcal{M}} \sum_{\sigma \in [n]^{k\ell}} \mathbb{I}[\|\mathcal{H}_\sigma\|_\oplus = 0] \leq (4\beta k\ell)! \cdot n^{k\ell} \cdot \left(\left(\frac{112}{\beta k} \right)^{k\ell(1-10\beta)} + n^{-\beta k\ell/3} \right)$$

Proof. Define a multigraph $\mathcal{G}$ as follows. In the multigraph $\mathcal{G}$, there is a vertex v_f for each *distinct* block $f \in \mathcal{F}$. There is an edge in $\mathcal{G}$ for each edge in the matchings $\mathcal{M}$ between the blocks. Every choice of *pivot* vertices $\sigma \in [n]^k$ corresponds to a labeling of the edges $\sigma : E(\mathcal{G}) \to [n]$. For each edge $e \in E(\mathcal{G})$ incident at a vertex $v_f \in V(\mathcal{G})$, there is a hyperedge in $\mathcal{H}_\sigma$ corresponding to $(\sigma(e), v_f)$. The hypergraph $\mathcal{H}_\sigma$ is *even* if and only if for each pivot vertex $i \in [n]$, and each vertex $v_f \in V(\mathcal{G})$, the number of edges labeled i incident at v_f is even. This implies that $\sigma^{-1}(i)$ form an Eulerian subgraph for each $i \in [n]$. By Lemma 3.3.17, the number of such labelings $\sigma : E(\mathcal{G}) \to [n]$ is at most $(2|E(\mathcal{G})| - 2|V(\mathcal{G})|)! \cdot n^{E(\mathcal{G})/2 - E_\oplus(\mathcal{G})/6}$.

By definition of the graph $\mathcal{G}$, $|V(\mathcal{G})| = \|\mathcal{F}\|_0 \geq 2k\ell(1-\beta)$ and $E(\mathcal{G}) = 2k\ell$. Moreover, by applying Lemma 3.3.8 with $r \leftarrow 2$, $c \leftarrow k$, $M \leftarrow 2\ell$, $\delta \leftarrow 1$, and $E_i \leftarrow \mathcal{F}_i$, we conclude that the graph $\mathcal{G}$ has many odd multiedges with high probability over the choice of $\mathcal{M}$. Formally,

$$\mathbb{P}[|E_\oplus(\mathcal{G})| \leq 2\beta k\ell] \leq \left(\frac{112}{\beta c} \right)^{k\ell(1-10\beta)}$$

Now we are ready to wrap up the proof of the lemma.

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathcal{M}} \sum_{\sigma \in [n]^{k\ell}} \mathbb{I}[\|\mathcal{H}_\sigma\|_\oplus = 0] &\leq \mathop{\mathbb{E}}_{\mathcal{M}} (2E(\mathcal{G}) - 2V(\mathcal{G}))! \cdot n^{E(\mathcal{G})/2 - E_\oplus(\mathcal{G})/6} \\
&\leq \mathop{\mathbb{E}}_{\mathcal{M}} (4\beta k\ell)! \cdot n^{k\ell} \cdot n^{-|E_\oplus(\mathcal{G})|/6} \\
&= (4\beta k\ell)! \cdot n^{k\ell} \cdot \mathop{\mathbb{E}}_{\mathcal{M}} n^{-|E_\oplus(\mathcal{G})|/6} \\
&\leq (4\beta k\ell)! \cdot n^{k\ell} \cdot \left(\mathbb{P}[|E_\oplus(\mathcal{G})| \leq 2\beta k\ell] + n^{-2\beta k\ell/6} \right) \\
&\leq (4\beta k\ell)! \cdot n^{k\ell} \cdot \left(\left(\frac{112}{\beta k} \right)^{k\ell(1-10\beta)} + n^{-2\beta k\ell/6} \right) \\
&\leq \left(k^{4\beta} \ell^{4\beta} \cdot n \cdot \left(\left(\frac{112}{\beta k} \right)^{(1-10\beta)} + n^{-\beta/3} \right) \right)^{k\ell}
\end{aligned}$$

□

Our final lemma of this section is a bound on the number of labelings of a multigraph such that the subgraphs induced by all edge labels are Eulerian, given a bound on the number of multi-edges appearing with odd multiplicity.

Lemma 3.3.17. *Given a multigraph $\mathcal{G}$, a labeling of its edges $\sigma : E(\mathcal{G}) \to [n]$ is said to be* even*, if the preimage of every label i forms an Eulerian subgraph (not necessarily connected) of $\mathcal{G}$. Specifically, the set of edges $\sigma^{-1}(i) \subseteq E(\mathcal{G})$ induce a subgraph where the degree of every vertex is even.*

$$|\{\sigma : E(\mathcal{G}) \to [n] | \sigma \text{ is even } \}| \leq (2|E(\mathcal{G})| - 2|V(\mathcal{G}|))! \cdot n^{|E(\mathcal{G})|/2 - |E_{\oplus}(\mathcal{G})|/6}$$

where $|E_{\oplus}(\mathcal{G})|$ is the number of multi-edges with odd multiplicity within $\mathcal{G}$.

Proof. We will count the number of *even* labelings σ as follows:

- Pick a unordered partition of the edges of the graph in to Eulerian subgraphs. By Claim 3.3.18, there are at most $(2|E(\mathcal{G})| - 2|V(\mathcal{G})|)!$ of them.
- Assign a label from $[n]$ to each Eulerian subgraph in the partition. The number of labelings is clearly at most n^t where t is the number of subgraphs in the partition. By Claim 3.3.19, there are at most $\frac{|E(\mathcal{G})|}{2} - \frac{|E_{\oplus}(\mathcal{G})|}{6}$ subgraphs in any partition. Hence, there are at most $n^{\frac{|E(\mathcal{G})|}{2} - \frac{|E_{\oplus}(\mathcal{G})|}{6}}$ labelings for each partition of $\mathcal{G}$ in to Eulerian subgraphs.

The lemma follows immediately from the Claim 3.3.18 and Claim 3.3.19 which we will show now.

Claim 3.3.18. The number of unordered partitions of the edges of the graph in to Eulerian subgraphs is at most $(2|E(\mathcal{G})| - 2|V(\mathcal{G})|)!$.

Proof. Let d_v denote the degree of vertex $v \in V(\mathcal{G})$. We can specify a partition of the edges of $\mathcal{G}$ in to Eulerian subgraphs, by specifying a sequence of Eulerian traversals whose union covers all the edges in the graph exactly once.

Consider a vertex v. Any sequence of traversals induces a matching M_v between the edges incident at v – where e, e' are matched if one of the traversals goes along $e \to v \to e'$. Furthermore, given a set of matchings $\{M_v | v \in V(\mathcal{G})\}$, it uniquely identifies a set of traversals.

Therefore the number of partitions of $E(\mathcal{G})$ in to Eulerian subgraphs is at most

$$\begin{aligned}
\prod_{v \in V(\mathcal{G})} |\#\text{ matchings of edges incident at} v| &\leq \prod_{v \in V(\mathcal{G})} \left(\frac{d_v!}{(d_v/2)!} \cdot \frac{1}{2^{d_v}} \right) \\
&\leq \prod_{v \in V(\mathcal{G})} (d_v - 2)! \\
&\leq (\sum_{v \in V(\mathcal{G})} (d_v - 2))! \leq (2E(\mathcal{G}) - 2V(\mathcal{G}))!
\end{aligned}$$

□

Claim 3.3.19. In any partition of $\mathcal{G}$ in to Eulerian subgraphs, the number of partitions is at most $\frac{|\mathcal{E}(G)|}{2} - \frac{|\mathcal{E}_{\oplus}(G)|}{6}$

Proof. Suppose $E(\mathcal{G}) = \cup_{i=1}^{t} E_i(\mathcal{G})$ denote a partition of $E(\mathcal{G})$ into Eulerian subgraphs. For each edge $e \in E_i(\mathcal{G})$ assign a weight $w_e = \frac{1}{|E_i(\mathcal{G})|}$. By definition of the weights, we have

$$\sum_{e \in E(\mathcal{G})} w_e = t\,.$$

Note that $w_e \leq \frac{1}{2}$ for all $e \in E(\mathcal{G})$, since each subset $E_i(\mathcal{G})$ contain at least two edges by virtue of being Eulerian. Moreover, $w_e = \frac{1}{2}$ if the edge e belongs to an Eulerian subgraph $E_i(\mathcal{G})$ with exactly two edges. In particular, $E_i(\mathcal{G}) = \{e, e'\}$ where e and e' form a 2-cycle. For every multiedge (a, b) with odd multiplicity, at least one of its edges has $w_e \leq \frac{1}{3} = \frac{1}{2} - \frac{1}{6}$.

Therefore we conclude that

$$t = \sum_{e \in E(\mathcal{G})} w_e \leq \sum_{e \in E(\mathcal{G})} \frac{1}{2} - \sum_{(a,b) \in E_{\oplus}(\mathcal{G})} \frac{1}{6} = \frac{|E(\mathcal{G})|}{2} - \frac{|E_{\oplus}(\mathcal{G})|}{6}\,.$$

□

These claims together finish the proof. □

Useful Combinatorial Lemmas

Define an *r-grouping* to be a partition of a set of size $c \cdot r$ into c subsets of size r. The following lemma bounds the probability that, given a multiset with many distinct elements, an r-grouping of the elements results in few r-sets with odd multiplicity. We rely on this lemma in our injective tensor norm upper bounds, to bound the probability that a hypergraph sampled from a simple graph has the evenness property.

Lemma (Restatement of Lemma 3.3.8)**.** *Fix $M, r, \ell, N \in \mathbb{N}$ and $\beta \in (0, 1)$. Let $E_1, \ldots, E_M \in [N]^{r \cdot c}$ be multisets of elements such that the number of distinct elements in the union $\cup_{i \in [M]} E_i$ is at least $(1 - \beta)M \cdot r \cdot c/2$. Let G_i denote a uniformly random r-grouping of elements within E_i, sampled independently for each $i \in [M]$. Let $\bigoplus_i G_i$ denote the set of r-groups $(a_1, \ldots, a_r) \in [N]^r$ that appear an odd number of times within $\cup_i G_i$. Then for any $0 < \delta < 3.5\beta$,*

$$\mathbb{P}[|\oplus_i G_i| \leq \delta M c] \leq \left(\frac{112}{\beta c}\right)^{(1-(4r+1)\beta-2\delta)(r-1)Mc/2}$$

Proof. We will refer to each $s \in [N]$ as a "type". Call a type $s \in [N]$ *infrequent* if the number of occurrences of s within $\cup_i E_i$ is nonzero but at most 8.

Suppose a type $s \in [N]$ appears exactly once in the sets $E_1, \ldots, E_M$,then irrespective of the choice of the grouping, the group involving s appears exactly once. If there are more than $r\delta Mc$ types that appear exactly once then,

$$\mathbb{P}[|\oplus_i G_i| \leq \delta M c] = 0,$$

and the lemma holds. Henceforth, we assume that all but $r\delta Mc$ types appear at least twice.

Call a type to be *frequent* if it occurs more than 8 times within $\cup_i E_i$. Out of the rMc elements, at most an 8β fraction are occurrences of *frequent* types. Otherwise, the number of distinct types would be less than $rMc\,((1-8\beta)/2 + 8\beta/8 + \delta) < \frac{rMc}{2}(1-\beta)$.

Moreover, this implies that the number of distinct *frequent* types is at most $8\beta rMc/8 \leq \beta rMc$. Finally, the number of distinct infrequent types is at least $\frac{rcM}{2}(1-\beta) - \beta rMc \geq \frac{rMc}{2} \cdot (1-3\beta)$.

Let us sample uniform random r-groupings $\{G_i\}_{i\in[M]}$ one group at a time. Specifically, we will sample groups $g_1, \ldots, g_{cM}$ where $G_i = \{g_{(i-1)c+1}, \ldots, g_{ic}\}$, one group at a time. We sample the i^{th} grouping G_i as follows:

- For $j = 1$ to c
- Pick the element s with the smallest number of ungrouped occurrences left within $\cup_{j=i}^{M} E_j$ (breaking ties lexicographically).
- Sample the group $g_{(i-1)c+j}$ by picking the remaining $r-1$ elements uniformly at random from ungrouped elements in E_i

It is clear that the above sampling procedure picks a uniformly random grouping $\{G_i\}_{i\in[M]}$.

We will refer to the groups picked at any stage to be *configuration*. So, the configuration at the end of i^{th} stage is $\mathcal{E}_i \stackrel{\text{def}}{=} \{g_1, \ldots, g_i\}$. Given a current configuration $\mathcal{E}_i$, there is a unique element $s(\mathcal{E}_i)$ that will be grouped in the next step. A configuration $\mathcal{E}_i$ is said to be *critical* if

1. $s(\mathcal{E}_i)$ is its final ungrouped occurrence of an *infrequent* type.
2. All previous occurrences of $s(\mathcal{E}_i)$ has been grouped with *infrequent* types.
3. There are at least βrc ungrouped elements within the current multiset E_j that is being grouped.

Claim 3.3.20. For every sequence of random choices, the sampling procedure encounters at least $\frac{cM}{2} \cdot (1-(4r+1)\beta)$ *critical* configurations.

Proof. There are at most $8\beta rcM$ occurrences of *frequent* types. This implies that among the $\frac{rMc}{2}(1-3\beta)$ infrequent labels, at least $\frac{rMc}{2}(1-3\beta) - (8\beta rMc)(r-1) \geq \frac{rMc}{2}(1-(4r-1)\beta)$ are grouped only with infrequent types.

For each of these $\frac{rMc}{2}(1-(4r-1)\beta)$ types there is one final ungrouped occurrence. Even assuming we match all these final occurrences among themselves, both conditions (1) & (2) are met at least $\frac{Mc}{2}(1-(4r-1)\beta)$ times during the sampling procedure.

Finally, there are at most βc groups that are picked among the final βrc elements within the sets E_i. Therefore, for at least $\frac{Mc}{2}(1-(4r-1)\beta) - \beta Mc \geq \frac{Mc}{2}(1-(4r+1)\beta)$ steps, $\mathcal{E}_i$ is a critical configuration. □

Define random variables $\{Z_i\}_{i\in[m]}$ as follows:

$$Z_i \stackrel{\text{def}}{=} \mathbb{I}[g_i \text{ is final occurrence of an odd group in } \cup_j G_j]\,.$$

By definition, we have

$$|\oplus_j G_j| = \sum_{i\in[cM]} Z_i$$

Set $\alpha = \left(\frac{56}{\beta c}\right)^{r-1}$. In order to obtain concentration bounds on $\sum_{i\in[cM]} Z_i$ we will bound $\mathbb{E}[\alpha^{\sum_i Z_i}]$.

Claim 3.3.21. For all $\alpha \leq \left(\frac{56}{\beta c}\right)^{c-1}$, for all $t \in [cM]$ and all critical configurations $\mathcal{E}_t$,

$$\mathbb{E}[\alpha^{\sum_{i=t}^{cM} Z_{i+1}} | \mathcal{E}_t] \leq 2\alpha \cdot \max_{\mathcal{E}_{t+1}|\mathcal{E}_t} \mathbb{E}[\alpha^{\sum_{i=t+1}^{cM} Z_{i+1}} | \mathcal{E}_{t+1}]$$

where the maximum is taken over all feasible configurations $\mathcal{E}_{t+1}$ from $\mathcal{E}_t$.

Proof. At a critical configuration $\mathcal{E}_t$, the next group is the last occurrence of $s(\mathcal{E}_t)$. Recall that $s(\mathcal{E}_t)$ is infrequent in that it has at most 7 previous occurrences. Moreover, each of its previous occurrences is grouped to an infrequent type (appearing less than 8 times).

There are at least βrc ungrouped elements from which the remaining $r-1$ elements of the group are chosen. For all but at most $(56)^{r-1}$ group choices, the group contains a type s' such that this is the first occurrence of s with s' in a group.

Therefore, for all but at most $(56)^{r-1}$ choices, the group sampled is its first and only occurrence. In particular, this implies that for a critical configuration $\mathcal{E}_t$,

$$\mathbb{P}[Z_{t+1} = 0 | \mathcal{E}_t] \leq \frac{(56)^{r-1}}{\binom{\beta rc}{r-1}} \leq \left(\frac{56}{\beta c}\right)^{r-1}$$

Finally, we have

$$\begin{aligned}
\mathbb{E}[\alpha^{\sum_{i=t+1}^{cM}}] &= \mathbb{P}[Z_{t+1} = 1 | \mathcal{E}_t] \cdot \alpha \cdot \mathbb{E}[\alpha^{\sum_{i=t+2}^{cM} Z_i} | \mathcal{E}_t, Z_{t+1} = 1] \\
&\quad + \mathbb{P}[Z_{t+1} = 0 | \mathcal{E}_t] \cdot \mathbb{E}[\alpha^{\sum_{i=t+2}^{cM} Z_i} | \mathcal{E}_t, Z_{t+1} = 0] \\
&\leq \alpha \cdot \mathbb{E}[\alpha^{\sum_{i=t+2}^{cM} Z_i} | \mathcal{E}_t, Z_{t+1} = 1] + \alpha \cdot \mathbb{E}[\alpha^{\sum_{i=t+2}^{cM} Z_i} | \mathcal{E}_t, Z_{t+1} = 0] \\
&\leq 2\alpha \cdot \max_{\mathcal{E}_{t+1}|\mathcal{E}_t} \mathbb{E}[\alpha^{\sum_{i=t+1}^{cM} Z_{i+1}} | \mathcal{E}_{t+1}]
\end{aligned}$$

□

Combining Claim 3.3.21 and Claim 3.3.20, we have that

$$\mathbb{E}[\alpha^{\sum_{i\in[cM]} z_i}] \leq (2\alpha)^{Mc(1-(4r+1)\beta)/2},$$

which yields the following concentration bound for all $\delta > 0$,

$$\mathbb{P}[|\oplus_i G_i| \leq \delta cM] < (2\alpha)^{(1-(4r+1)\beta-2\delta)cM/2} \leq \left(\frac{112}{\beta k}\right)^{(1-(4r+1)\beta-2\delta)(r-1)Mc/2}$$

□

The lemma below shows that in a simple graph formed by matchings with the evenness property, there cannot be too many more distinct vertices than distinct edges.

Lemma (Restatement of Lemma 3.3.7)**.** *Let $m, n, c \in \mathbb{N}$, and let G be a graph which is a union of at most c disjoint cycles. Suppose furthermore that each vertex receives labels from the set $[n]$, that every labeled edge appears with even multiplicity, and that there are exactly m distinct labeled edges. Then letting L be the number of distinct vertex labels, we have*

$$L \leq m + c.$$

Proof. We first prove the following claim:

Claim. If each labeled edge appears with multiplicity exactly 2, then $L \leq m + c$.

Proof. In this case, there are exactly $2m$ edges and exactly $2m$ vertices. We proceed by induction on c and m. In the base case, we have $c = 1$ component with 2 vertices, in which case we have at most 2 distinct labels on the vertices, confirming the claim.

Assuming the claim for $c \geq 1$ components and $2m \geq 2$ vertices, consider an instance on $2m + 2$ vertices. If all labels appear ≥ 2 times, we are done, since there are $2m + 2$ vertices and thus at most $L \leq m + 1$ labels. Otherwise, locate a vertex v whose label has multiplicity 1.

If v is in a cycle of length 2, remove v and its neighbor from the graph, obtaining a smaller instance with L' labels, c' components, and m' distinct edge types, with $L' + 2 \geq L$, $c' = c - 1$, and $m' = m$. By the induction hypothesis, $L' \leq m' + c' = m + c - 1$, and therefore $L \leq m + 1 + c$, as desired.

If v's cycle has length > 2, both v's vertex neighbors must have the same label in order for the edges incident on v to appear twice. We remove v and identify its neighbors, obtaining an instance with $L' + 1 = L$, $m' = m$, $c' = c$. Appealing to the induction hypothesis, we have $L' \leq m' + c$, from which we conclude that $L \leq m + 1 + c$, as desired. □

Now, we reduce our lemma to the above case. Say an edge appears with even multiplicity $\mu > 2$, and that the labels of the edge are $(a, b) \in [n]^2$. We will remove the occurrences of this edge, and put the graph segments back together. When we remove all occurrences of the edge (a, b), we get 3 kinds of graph segments: paths from a–b, paths from a–a, and paths from b–b. Since a, b each have to appear μ times, we can form a matching between segments of type a–b, gluing them together at the a endpoint to get a b–b segment. Now, we make one cycle by gluing together a–a segments, and a separate cycle by gluing together b–b segments. Our number of distinct edges has decreased by 1, and our number of cycles has increased by at most 1, since we broke up at least one cycle to remove the edge (a, b). We recursively apply this process to our instance, until we reach an instance in which there are only edges of multiplicity 2, never increasing the quantity $m + c$. In conjunction with our above claim, the conclusion follows. □

3.4 Refuting Random k-XOR Instances

In this section, we give our algorithm for refuting random k-XOR instances. In Section 3.4, we describe the algorithm for even k; in Section 3.4, we describe the algorithm for odd k. We first recall the problem:

Definition 3.4.1 (Random k-XOR with density $\alpha = pn^{(k-1)/2}$)**.** A *random instance of k-XOR with density $\alpha = pn^{(k-1)/2}$* is a formula Φ on n variables $x \in \{\pm 1\}^n$, sampled so that for each $S \in [n]^k$:

- Independently with probability p, add constraint $C_S : \prod_{i \in S} x_i = \eta_S$, for η_S a uniformly random Rademacher variable.
- Otherwise, with probability $1 - p$, add no constraint.

We let $m \approx pn^k$ be the number of constraints, and for any assignment $x \in \{\pm 1\}^n$, $P_\Phi(x)$ is the fraction of constraints satisfied by x.

Problem 3.4.2 (Strongly refuting random k-XOR)**.** Given a random k-XOR instance Φ, certify with high probability over the choice of Φ that for all assignments $x \in \{\pm 1\}^n$,

$$P_\Phi(x) \leq \frac{1}{2} + \delta + o(1),$$

for some constant $\delta \in [0, 1/2)$, where $P_\Phi(x)$ is the fraction of Φ's constraints satisfied by x.

As described in Section 3.2, there is a natural random order-k tensor that we can identify with any k-XOR instance Φ. Given a k-XOR instance Φ with constraints $C_1, \ldots, C_m$, form the tensor $\mathbf{T}_\Phi$ as follows: for each constraint $C_i : \prod_{j \in S_i} x_i = \eta_{S_i}$, set the entry $\mathbf{T}_{S_i} = \eta_{S_i}$; in all other entries place a 0. We then have that for any assignment $x \in \{\pm 1\}^n$,

$$\langle \mathbf{T}_\Phi, x^{\otimes k} \rangle = \sum_{i \in [m]} \eta_{S_i} \cdot \prod_{j \in S_i} x_j = m \cdot \left(P_\Phi(x) - \frac{1}{2} \right)$$

That is, the inner product $\langle \mathbf{T}_\Phi, x^{\otimes k} \rangle$ gives the difference between the number of constraints x satisfies and the number of constraints x violates. Our strong refutation algorithm will be based on showing that

$$\left| \left\langle \mathbf{T}_\Phi, x^{\otimes k} \right\rangle \right| \leq (\delta + o(1)) \cdot m \quad \forall x \in \{\pm 1\}^n, \tag{3.4.1}$$

for a constant δ arbitrarily close to 0.

From (3.4.1), it is clear that a good bound on $\|\mathbf{T}_\Phi\|_{inj}$ would give a refutation algorithm, and so we could hope that our algorithms for bounding tensor norms would suffice. However, when the probability of sampling a constraint $p \leq n^{-k/2}$, the tensor $\mathbf{T}_\Phi$ becomes sparse enough that its norm is maximized by sparse vectors, so that $\|\mathbf{T}_\Phi\|_{inj} \approx 1$. We are only interested in balanced vector $x \in \{\pm 1\}^n$, and so this is a poor upper bound–it will only let us certify that $P_\Phi(x) \leq \frac{1}{2} + \frac{n^{k/2}}{m} \geq 1$.

So our algorithm for the case of k-XOR is almost identical to our algorithm for bounding tensor norms, but with an additional twist to get rid of the sparse vectors. We form our certificate as we did in the tensor norm algorithm: we flatten $\mathbf{T}_\Phi$ to a matrix T, then take the dth Kronecker power of T, and we average over rows and columns corresponding to permutations of the same index set. But now, there is one additional step: we delete any row or column indexed by a multiset $S \in [n]^{kd/2}$ which contains an element i with multiplicity greater than $O(\log n)$.

It is not difficult to see why this should help: supposing we started with a sparse vector, say the standard basis vector $e_1 \in \mathbb{R}^n$, this will ensure that $e_1^{\otimes kd/2}$ has 0 projection onto our matrix. On the other hand, the choice of $O(\log n)$ as our upper bound on the multiplicity makes sense, since we are eliminating an $o(1)$-fraction of the Frobenius norm of the certificate in this way, even when $d \geq n^{1/2}$–if we were to delete all rows and columns in which an element appears with multiplicity ≥ 2, then once $d = n^{1/2}$ we would be deleting a constant fraction of the rows and columns, by the birthday paradox.

This introduces some technicalities in the analysis–in particular, once we delete these rows and columns, it is no longer obvious that we are working with a valid relaxation of $\langle \mathbf{T}_\Phi, x^{\otimes k} \rangle$ over $x \in \{\pm 1\}^n$. But as before, the main theorems of this section will have to do with bounding the norm of our matrix certificate–arguing that the matrix certificate is valid will be straightforward.

We begin by detailing our algorithm for even k, then give the somewhat more involved analysis for odd k (the additional complication introduced by the lack of a natural matrix flattening for odd-order tensors).

Even k-XOR

We begin by describing our matrix certificate for this case, and establishing an upper bound on its norm–as mentioned above, this is the main result of this section. Later, in Section 3.4, we will show how to use the matrix to get a valid certificate.

Algorithm 3.4.3 (Even k-XOR Certificate at level d)**.**
Input: A k-XOR instance Φ for even k on n variables and m clauses. Parameters $d \in \mathbb{N}$.

1. Form the tensor $\mathbf{T}_\Phi$ from Φ as described above (see (3.4.1)).
2. Take the natural $n^{k/2} \times n^{k/2}$ matrix flattening T of $\mathbf{T}_{inj}$, and take the Kronecker power $T^{\otimes d}$
3. Letting $\hat{\mathcal{S}}_{dk/2}$ be the set of all permutation matrices that perform the permutations corresponding to $\mathcal{S}_{dk/2}$ on the rows and columns of T, form
$$C_{(d)} \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{\Pi,\Sigma \in \hat{\mathcal{S}}_{dk/2}} \left[\Pi T^{\otimes d} \Sigma\right] .$$
4. Zero out any row or column of $C_{(d)}$ indexed by a multiset in $[n]^{kd/2}$ containing more than $10 \log n$ copies of any $i \in [n]$.

Output: The value $\|C_{(d)}\|$.

The following theorem gives a bound on the value output by Algorithm 3.4.3.

Theorem 3.4.4. *Let $k, n, d, k \in \mathbb{N}$, so that $d \log n \ll n$, k is even. Let Φ be a random instance of k-XOR on n variables with $\Theta(pn^k)$ clauses (so each constraint is sampled uniformly and independently with probability p). Let $C_{(d)}$ be the matrix formed from the instance Φ as described in Algorithm 3.4.3. Then if $p \cdot d^{(k/2-1)} n^{k/2} > 1$, there is a constant c_k depending on k such that with high probability,*
$$\|C_{(d)}\|^{1/d} \leq \left(c_k \log^{2k} n \cdot \frac{p^{1/2} n^{k/4}}{d^{(k-2)/4}} \right) .$$

We will prove the theorem below, in Section 3.4. First, we will see how to use this certificate, with the deleted high-multiplicity rows and columns, to strongly refute k-XOR instances.

Validity of certificate with deleted rows and columns

When we zero out the high-multiplicity rows and columns in Algorithm 3.4.3,

$$\left\langle \mathbf{T}_\Phi, x^{\otimes k} \right\rangle^d = (x^{\otimes dk/2})^\top \left(C_{(d)} + C_{\geq} \right) x^{\otimes dk/2},$$

where $C_{\geq}$ is a matrix containing only the zeroed out rows and columns. So our upper bound on $\|C_{(d)}\|$ from Theorem 3.4.4 is not enough. It is not hard to bound the ℓ_1-norm of $C_{\geq}$. However, because for our values of p, $\|C_{(d)}\|$ is close to 0, the ℓ_1-norm bound is too costly when we try to bound $\langle \mathbf{T}_\Phi, x^{\otimes k} \rangle$. For this reason, we will work with $P_\Phi(x)$, the fraction of satisfied constraints, which is bounded away from 0. We will relate $(P_\Phi(x))^d$ to the matrix norms of $C_1, \dots C_d$.

Let us write

$$P_\Phi(x) = \mathop{\mathbb{E}}_{i \sim [m]} [P_i(x)] = \mathop{\mathbb{E}}_{i \sim [m]} \left[\frac{1}{2}(1 + C_i(x)) \right],$$

where $P_1(x), \dots, P_m(x)$ are the $0-1$ valued predicates of the instance Φ, and $C_1(x), \dots, C_m(x)$ are the ± 1-valued predicates of the instance Φ. We have that

$$(P_\Phi(x))^d = \mathop{\mathbb{E}}_{i_1, \dots, i_d \sim [m]^d} \left[\prod_{\ell=1}^{d} P_{i_\ell}(x) \right].$$

We will prove that the quantity above is not changed very much if we remove sets $i_1, \dots, i_d$ corresponding to high-multiplicity rows and columns.

Proposition 3.4.5. *Let Φ be a random k-XOR formula in which each clause is sampled independently with probability p.*

Let $\mathcal{C}^d_{low} \subset [m]^d$ be the set of all ordered multisets of clauses $C_{i_1}, \dots, C_{i_d}$ from Φ with the property that if we form two multisets of variables $I, J \in [n]^{dk/2}$ with I containing the first $k/2$ variables of each C_{i_ℓ} and J containing the last $k/2$ variables of each C_{i_ℓ}, then I, J are both low-multiplicity multisets, in that both have no element of $[n]$ with multiplicity $\geq 100 \log n$.

Suppose that no variable appears in more than $m_{\max}$ clauses. Then if $d \ll n$ and $dkm_{\max} < 200 \varepsilon m \log n$,

$$P_\Phi(x) \leq \left(\mathop{\mathbb{E}}_{i_1, \dots, i_d \sim \mathcal{C}^d_{low}} \left[\prod_{\ell=1}^{d} P_{i_\ell}(x) \right] \right)^{1/d} + \varepsilon$$

for all $x \in \{\pm 1\}^n$ with high probability. Furthermore when $p \geq 200 \frac{\log n}{n^{k-1}}$, we have that $\varepsilon = o(1)$ with high probability.

Proof. Let $m_{\max}$ be an upper bound on the number of clauses any variable x_i appears in the instance Φ. We sample a uniform element $\mathcal{C} \sim \mathcal{C}^d_{low}$, $\mathcal{C} = C_1, \ldots, C_d$ in the following way:

- For $t = 1, \ldots, d$: Let $\mathcal{A}_t \subset \Phi$ be the set of clauses such that for any $C' \in \mathcal{A}$, the multiset $C_1, \ldots, C_{t-1}, C'$ is not excluded from $\mathcal{C}^t_{low}$. Choose a uniformly random $C \sim \mathcal{A}_t$ and set $C_t := C$, adding C to $\mathcal{C}$.

This sampling process clearly gives a uniformly random element of $\mathcal{C}^d_{low}$.

Claim 3.4.6. At step $t+1$ there are at least $m - t\frac{k \cdot m_{\max}}{200 \log n}$ clauses that can be added.

Proof. In order to exclude any variable $i \in [n]$, we must add at least $200 \log n$ copies of i. Further, to exclude ℓ distinct variables in $[n]$, at least $200 \log n$ copies of each variable, for a total of $200\ell \log n$ variables, which requires adding at least $200\ell \log n/k$ clauses. If ℓ distinct variables are excluded, then at most $\ell \cdot m_{\max}$ clauses are excluded. The claim now follows. □

Now, define the random variable $X_t = \prod_{j=1}^{t} P_j(x)$ to be the value of x on $\mathcal{C}_t$. We apply Claim 3.4.6, along with the observation that the total number of satisfied clauses can only drop by 1 for each clause removed regardless of the assignment x, to conclude that

$$\mathbb{E}[X_{t+1} | C_1, \ldots, C_t] \geq \left(P_\Phi(x) - t\frac{k \cdot m_{\max}}{m \cdot 200 \log n} \right) \cdot X_t$$

From this we have that $\mathbb{E}[X_t] \geq \left(P_\Phi(x) - t\frac{km_{\max}}{200m \log n} \right) \cdot \mathbb{E}[X_{t-1}]$ from which we have that as long as $dkm_{\max} \leq \varepsilon 200m \log n$,

$$\mathbb{E}[X_d] \geq \prod_{t=1}^{d} \left(P_\Phi(x) - t\frac{km_{\max}}{200m \log n} \right) \geq \left(P_\Phi(x) - \varepsilon \right)^d .$$

Which by definition of X_d gives us our first result.

Now, we can establish that $dkm_{\max} \leq \varepsilon \cdot 200m \log n$ with high probability. A Chernoff bound implies that when $p \geq 200 \log n/n^{k-1}$, $2pn^k \geq m \geq pn^k/2$ with probability at least $1 - 2\exp(-pn^k/8)$, and that $m_{\max} \leq 2pn^{k-1}$ with probability at least $1 - \exp(-pn^{k-1}/2)$, and so by a union bound and using the assumption that $pn^{k-1} \geq 200 \log n$, we have our result by taking $\varepsilon = \Theta(1/\log n)$. □

Now, we will relate the right-hand-side of Proposition 3.4.5 to the matrices from Algorithm 3.4.3. We recall that given a k-XOR instance Φ, $\mathcal{C}^d_{low} \subset [m]^d$ is the set of all ordered multisets of clauses $C_{i_1}, \ldots, C_{i_d}$ from Φ with the property that if we form two multisets of variables $I, J \in [n]^{dk/2}$ with I containing the first $k/2$ variables of each C_{i_ℓ} and J containing the last $k/2$ variables of each C_{i_ℓ}, then I, J are both low-multiplicity multisets, in that both have no element of $[n]$ with multiplicity $\geq 100 \log n$. By Proposition 3.4.5,

$$\left(P_\Phi(x) - o(1) \right)^d \leq \mathop{\mathbb{E}}_{i_1, \ldots, i_d \sim \mathcal{C}^d_{low}} \left[\prod_{\ell=1}^{d} P_{i_\ell}(x) \right] = \mathop{\mathbb{E}}_{i_1, \ldots, i_d \sim \mathcal{C}^d_{low}} \left[\prod_{\ell=1}^{d} \frac{1}{2}(1 + C_{i_\ell}(x)) \right]$$

and expanding the product on the right and applying the symmetry of the uniform distribution on $\mathcal{C}^d_{low}$,

$$= \left(\frac{1}{2}\right)^d \sum_{j=0}^{d} \binom{d}{j} \mathop{\mathbb{E}}_{i_1,\ldots,i_j \sim \mathcal{C}^j_{low}} \left[\prod_{\ell=1}^{d} C_{i_\ell}(x)\right]$$

and by definition, for any assignment $x \in \{\pm 1\}^n$,

$$= \left(\frac{1}{2}\right)^d \sum_{j=0}^{d} \binom{d}{j} \frac{1}{|\mathcal{C}^j_{low}|} \cdot (x^{\otimes jk/2})^\top C_{(j)} (x^{\otimes jk/2}),$$

where $C_{(j)}$ is the matrix output by Algorithm 3.4.3 when $d \leftarrow j$. We won't get a good bound on $\|C_{(j)}\|$ when j is too small, but we can take

$$= \left(\frac{1}{2}\right)^d \sum_{j=0}^{t} \binom{d}{j} + \left(\frac{1}{2}\right)^d \sum_{j=t+1}^{d} \binom{d}{j} \frac{1}{|\mathcal{C}^j_{low}|} \cdot (x^{\otimes jk/2})^\top C_{(j)} (x^{\otimes jk/2}) \,. \tag{3.4.2}$$

We'll take $t = \alpha \cdot d$ for some small constant α, so that the sum on the left is small, and the sum on the right we will bound by applying Theorem 3.4.4, our upper bound on $\|C_{(j)}\|$. We will also need a bound on $|\mathcal{C}^j_{low}|$, which we can easily get by modifying our proof of Proposition 3.4.5:

Lemma 3.4.7. *If Φ is a random k-XOR instance on n variables with m clauses such that no variable participates in more than $m_{\max}$ clauses, then so long as $d \ll n$ and $dkm_{\max} \leq 200\varepsilon m \log n$,*

$$|\mathcal{C}^d_{low}| \geq (1-\varepsilon)^d m^d \,.$$

Furthermore, when $p \geq 200\frac{\log n}{n^{k-1}}$, we can take $\varepsilon = o(1)$ with high probability.

The proof proceeds exactly as the proof of Proposition 3.4.5, but instead of bounding the decrease in the value as each clause is added, one bounds the probability that a clause is chosen which will make the multiplicity of some index too high.

We are now ready to prove that computing the norm of $O(d)$ matrices $C_{(\alpha d)}, \ldots, C_{(d)}$ will give us a strong refutation algorithm for random k-XOR. This concludes the proof of the refutation theorem, modulo the proof of the $C_{(d)}$ matrix norm bound from Theorem 3.4.4, which we give in the next subsection.

Theorem 3.4.8. *Let k be even, and let $d \ll n$. Then there is an algorithm that certifies with high probability that a random k-XOR instance has value at most $\frac{1}{2} + \gamma + o(1)$ for any constant $\gamma > 0$ at clause density $m/n = \tilde{O}\left(\frac{n^{k/2-1}}{d^{(k/2-1)}}\right)$ (where the $\tilde{O}$ hides a dependence on γ and k) in time $n^{O(d)}$.*

Proof. Define $\beta := \frac{dkm_{\max}}{200m \log n}$, where $m_{\max}$ is the maximum number of clauses any variable participates in. By Proposition 3.4.5 and the proceeding calculations culminating in (3.4.2), with high probability over the choice of the instance Φ, for any $x \in \{\pm 1\}^n$,

$$(P_\Phi(x) - \beta)^d \leq \frac{1}{2^d} \sum_{j=0}^{t} \binom{d}{j} + \frac{1}{2^d} \sum_{j=t+1}^{d} \binom{d}{j} \frac{1}{|\mathcal{C}^j_{low}|} \cdot (x^{\otimes jk/2})^\top C_{(j)} (x^{\otimes jk/2}) \,.$$

Setting $t = \delta d$ for some $\delta < 1$,

$$\leq \frac{1}{2^d} \cdot \left(\sum_{j=1}^{\delta d} \binom{d}{j} \right) + \sum_{j=\delta d+1}^{d} \frac{\binom{d}{j}}{2^d} \cdot \frac{1}{|\mathcal{C}^j_{low}|} \cdot (x^{\otimes jk/2})^\top C_{(j)} (x^{\otimes jk/2})$$

For the terms in the right-hand sum, we can apply our bound on $|\mathcal{C}^j_{low}|$ from Lemma 3.4.7 and the fact that $\|x\| = n^{1/2}$, to conclude that

$$\frac{1}{|\mathcal{C}^j_{low}|} \cdot (x^{\otimes jk/2})^\top C_{(j)} (x^{\otimes jk/2}) \leq \frac{\|x\|^{jk}}{|\mathcal{C}^j_{low}|} \cdot \|C_{(j)}\| = \frac{n^{jk/2}}{|\mathcal{C}^j_{low}|} \cdot \|C_{(j)}\|$$

By Lemma 3.4.7 and Theorem 3.4.4,

$$\leq \frac{n^{jk/2}}{(1-\beta)^j m^j} \cdot \left(\frac{n^{k/2} p \cdot c_k \log^{2k} n}{j^{(k/2-1)}} \right)^{j/2}$$

And since $m = \Theta(pn^k)$ w.h.p.,

$$\leq \frac{n^{jk/2}}{(1-\beta)^j (0.1pn^k)^j} \cdot \left(\frac{n^{k/2} p \cdot c_k \log^{2k} n}{j^{(k/2-1)}} \right)^{j/2}$$
$$\leq \left(\frac{c'_k \log^{2k} n}{(1-\beta)^2 pn^{k/2} \cdot j^{k/2-1}} \right)^{j/2} .$$

for some constant c'_d, where the second inequality holds with high probability from the conditions of Theorem 3.4.4, and so also holds with high probability simultaneously for all $j \in [\delta d, d]$ by a union bound.

The term comprised of the sum of binomial coefficients is at most

$$\frac{1}{2^d} \sum_{j=0}^{\delta d} \binom{d}{j} \leq 2^{(H(\delta)-1)d},$$

where $H(\cdot)$ is the binary entropy function, $H(\delta) = -\delta \log_2 \delta - (1-\delta) \log_2 (1-\delta)$.

Since the coefficients of the $C_{(j)}$ terms sum to < 1, we have that for some $\alpha \in [\delta, 1]$,

$$P_\Phi(x) - \beta \leq \left(2^{(H(\delta)-1)d} + \left(c'_k \frac{\log^{2k} n}{(1-\beta)^2 pn^{k/2} (\alpha d)^{k/2-1}} \right)^{\alpha d/2} \right)^{1/d}$$

Now, for $p \gg \tilde{O}(n^{-k/2}d^{-(k/2-1)})$, where the $\tilde{O}$ hides a dependence on k and $\alpha > \delta$, $\beta = o(1)$ and the latter quantity is $o(1)$. Thus, for sufficiently large n the term in the parenthesis is at most $(1 + o(1))2^{H(\delta)-1}$, which we can take to be a constant arbitrarily close to $\frac{1}{2}$ by choosing sufficiently small constant δ. We can certify this bound in time $n^{O(d)} \cdot d$, by running Algorithm 3.4.3 to compute $\|C_{(j)}\|$ for each $j \in [\delta d, d]$. □

Bounding the Even Certificate Spectral Norm

Here, we prove the norm bound on the matrix $\|C_{(d)}\|$ given in Theorem 3.4.4, the main theorem of this section.

Theorem (Restatement of Theorem 3.4.4)**.** *Let $k, n, d, k \in \mathbb{N}$, so that $d \log n \ll n$, k is even. Let Φ be a random instance of k-XOR on n variables with $\Theta(pn^k)$ clauses (so each constraint is sampled uniformly and independently with probability p). Let $C_{(d)}$ be the matrix formed from the instance Φ as described in Algorithm 3.4.3. Then if $p \cdot d^{(k/2-1)}n^{k/2} > 1$, there is a constant c_k depending on k such that with high probability,*

$$\|C_{(d)}\|^{1/d} \leq \left(c_k \log^{2k} n \cdot \frac{p^{1/2} n^{k/4}}{d^{(k-2)/4}} \right) .$$

The proof is similar to that of Theorem 3.3.3, except that, because the moments of the entries of $\mathbf{T}_\Phi$ depend on p, and because we rely on getting an accurate bound in terms of p, our counting arguments have to be much more precise. So we require stricter, specialized analogues of our even simple graphs count (Proposition 3.3.6) and our even hypergraph sampling probability (Lemma 3.2.8).

Proof. We will apply the trace power method (Proposition 3.2.4) to $C_{(d)}$, for which it suffices to obtain an upper bound on $\mathbb{E}[\mathrm{Tr}((C_{(d)}C_{(d)}^\top)^\ell)]$. We recall from Section 3.2 our interpretation of the (S, T)th entry of $C_{(d)}$ as the average over all k-hypergraph matchings between two multisets $S, T \in [n]^{dk/2}$; additionally, now by construction we can restrict our attention to S, T which do not have more than $R \stackrel{\text{def}}{=} 100 \log n$ copies of any one vertex (since those rows/columns are zeroed out). For convenience, we say such sets are R-multilinear.

We also recall that the trace gives us a sum over all R-multilinear vertex configurations consisting of sets $S_1, \ldots, S_{2\ell} \in [n]^{dk/2}$, and for each vertex configuration an average over all choices of sequences of hypergraph matchings. Let the set of all valid R-multilinear vertex configurations be denoted $\mathcal{V}_R$, and let the set of all hyperedge matching sequences be denoted $\mathcal{H}$. For $H \in \mathcal{H}$ and $V \in \mathcal{V}_R$, denote by (V, H) the hypergraph given by the hyperedges H on the vertex configuration V. Applying the above observations, and recalling that we have assembled $C_{(d)}$ from the random tensor $\mathbf{T} := \mathbf{T}_\Phi$, we have that

$$\mathbb{E}[\mathrm{Tr}(C_{(d)}C_{(d)}^\top)^\ell] = \sum_{V \in \mathcal{V}_R} \mathop{\mathbb{E}}_{H \in \mathcal{H}} \left[\prod_{(i_1,\ldots,i_d) \in (V,H)} \mathbf{T}_{i_1,\ldots,i_d} \right] ,$$

The expectation of each product is 0 if any hyperedge in (V, H) appears with odd multiplicity, and is p^M if exactly M distinct hyperedges appear in (V, H). Thus,

$$\begin{aligned}\mathbb{E}[\mathrm{Tr}(C_{(d)}C_{(d)}^\top)^\ell] &\leq \sum_{V\in\mathcal{V}_R}\sum_{M=1}^{d\ell} p^M \cdot \mathop{\mathbb{E}}_{H\in\mathcal{H}}[\mathbb{I}((V,H)\text{ even})\cdot\mathbb{I}((V,H)\text{ has }M\text{ hyperedges})]\\ &= \sum_{V\in\mathcal{V}_R}\sum_{M=1}^{d\ell} p^M \cdot \mathop{\mathbb{P}}_{H\in\mathcal{H}}[(V,H)\text{ even with }M\text{ hyperedges})]\ . \qquad (3.4.3)\end{aligned}$$

To bound this probability, we will again sample uniformly $H \sim \mathcal{H}$ in a two step process.

1. Sample a uniformly random perfect matching (with 2-edges rather than hyperedges) between each set $S_i, S_{i+1} \in \mathcal{V}_R$–call the edge set sampled in this manner E, so that we now have the graph (V, E).
2. Sample hyperedge matching configuration from E by choosing a uniform random grouping of the edges between S_i, S_{i+1} into groups of $k/2 = \kappa$ edges.

We invoke the following lemma, which is a very slight embellishment upon Lemma 3.2.7:

Lemma 3.4.9. *Let $h, w, \kappa, t, \tau \in \mathbb{N}$. Let $V \in \mathcal{V}_R$ be a vertex configuration with R-multilinear vertex sets $S_1, \ldots, S_w \in [n]^{\kappa h}$. Let $H \in \mathcal{H}$ be a hypergraph configuration with w 2κ-uniform hypergraph matchings between the sets $S_i, S_{i+1} \forall i \in [w]$, with κ vertices from S_i and κ vertices from S_{i+1} in each hypergraph matching.*

Suppose that (V, H) has τ distinct labeled hyperedges and the evenness property, where hyperedges on the same vertex set but with a different partition into S_i, S_{i+1} count as distinct. Suppose that we sampled H by first choosing a set of simple-edge perfect matchings E on V, then grouping them into hyperedges. Then

$$\mathbb{P}(\ (V,E)\ \textit{even with}\ t \leq \kappa\tau\ \textit{edges} \mid (V,H)\ \textit{even with}\ \tau\ \textit{edges}) \geq \left(\frac{1}{\kappa!}\right)^{wh}.$$

Proof. The proof is almost identical to that of Lemma 3.2.7–choosing a random matching within each hyperedge gives a uniformly random E from which H is sampled, and that with probability at least $(\kappa!)^{-wh}$ we choose the same matching in every copy of every hyperedge. We need only add that if a hyperedge $h_i \in (V, H)$ has multiplicity a, then if we chose the same matching in every copy of h_i, all κ of the simple edges making up h_i will have multiplicity at least a, so if t is the total number of distinct edges in (V, E), we have $t \leq \kappa\tau$ and also the evenness property. □

Letting $\mathcal{E}_H^M$ be the event that (V, H) is even with M distinct hyperedges and letting $\mathcal{E}_E^{kM/2}$ be the event that (V, E) is even with at most $kM/2$ distinct hyperedges, Lemma 3.4.9 (and the asymmetry of $\mathbf{T}_\Phi$) with $w \leftarrow 2\ell$, $h \leftarrow d$, $\tau \leftarrow M$, $\kappa \leftarrow k/2$ implies that

$$\begin{aligned}\mathop{\mathbb{P}}_{H\in\mathcal{H}}(\ (V,H)\text{ even with }M\text{ edges}) &= \frac{\mathbb{P}(\mathcal{E}_H^M, \mathcal{E}_E^{kM/2})}{\mathbb{P}(\mathcal{E}_E^{kM/2} \mid \mathcal{E}_H^M)}\\ &\leq \left(\frac{k}{2}!\right)^{2d\ell}\mathbb{P}(\mathcal{E}_H^M, \mathcal{E}_E^{kM/2}) \quad \text{(by Lemma 3.4.9)}\end{aligned}$$

$$\leq \left(\frac{k}{2}\right)^{dk\ell} \mathbb{P}(\mathcal{E}_E^M) \cdot \mathbb{P}(\mathcal{E}_H^M \mid \mathcal{E}_E^{kM/2}) \,.$$

Therefore, from (3.4.3) we have

$$\mathbb{E}[\mathrm{Tr}(C_{(d)}C_{(d)}^\top)^\ell] \leq \left(\frac{k}{2}\right)^{dk\ell} \sum_{V \in \mathcal{V}_R} \sum_{M=1}^{d\ell} \mathbb{P}(\mathcal{E}_E^{kM/2}) \cdot \mathbb{P}(\mathcal{E}_H^M | \mathcal{E}_E^{kM/2}) \cdot p^M \,,$$

Now, we use a lemma to bound the conditional probability of sampling an even hyperedge matching with M hyperedges, given that we sampled an even matching with at most $kM/2$ edges:

Lemma 3.4.10. *Suppose $h, w, \kappa, n, \tau \in \mathbb{N}$. Let $G = (V, E)$ be a graph consisting of w sets of κh vertices each with R-multilinear labels from $[n]$, where E is a set of w perfect matchings $M_1, \ldots, M_w$, so that M_i is a perfect matching between S_i and S_{i+1}, and $\alpha = a_1, \ldots, a_t$ is a list of even edge multiplicities of E on the labeled vertex set V, so that $\sum a_i = \kappa w h$.*

Suppose we sample a hyperedge matching configuration H from E by uniformly grouping the edges in each matching from S_i to S_{i+1} into hyperedges of order 2κ, and let τ be a number of distinct hyperedges that is possible to sample from (V, E) in this way. Then,

$$\mathbb{P}((V, H) \text{ even with } \tau \text{ edges} \mid (V, E) \text{ even }) \leq \frac{(2e^\kappa \kappa^\kappa R^{\kappa+1} w)^{wh}}{(\kappa h)^{(\kappa-1)(wh-\tau)}} \,.$$

We'll prove Lemma 3.4.10 below in Section 3.4. For now, we apply Lemma 3.4.10 with $w \leftarrow 2\ell$, $h \leftarrow d$, $\kappa \leftarrow k/2$ and $\tau \leftarrow M$, which for R-multilinear $V \in \mathcal{V}_R$ implies that

$$\mathbb{P}(\mathcal{E}_H^M | \mathcal{E}_E^{kM/2}) \leq \frac{(4e^{k/2} R^{k/2+1} (k/2)^{k/2} \ell)^{2d\ell}}{(dk/2)^{(k/2-1)(2d\ell - M)}} \,.$$

Combining this with the above and letting $c_1 := 4e^{k/2}(k/2)^{k/2}$ for convenience,

$$\mathbb{E}[\mathrm{Tr}(CC^\top)^\ell] \leq \left(\frac{k}{2}\right)^{dk\ell} \sum_{V \in \mathcal{V}_R} \sum_{M=1}^{d\ell} \mathbb{P}(\mathcal{E}_E^{kM/2}) \cdot \frac{(c_1 R^{k/2+1} \ell)^{2d\ell}}{(dk/2)^{(k/2-1)(2d\ell - M)}} \cdot p^M \,. \tag{3.4.4}$$

It remains for us to bound $\mathbb{P}((V, E)$ even with $\leq kM/2$ edges$)$. We now interchange the order of the summation, and bound the sum over V for a fixed value of M. Letting $\mathcal{M}$ be the set of all possible edge configurations E, we have

$$\begin{aligned} \sum_{V \in \mathcal{V}_R} \mathop{\mathbb{P}}_E(\mathcal{E}_E^{kM/2}) &= \sum_{V \in \mathcal{V}_R} \frac{|\{E \mid (V, E) \text{ even with} \leq kM/2 \text{ edges}\}|}{|\mathcal{M}|} \\ &= \frac{|\{\, E, V \mid (V, E) \text{ even with} \leq kM/2 \text{ edges}\,\}|}{|\mathcal{M}|} \,. \end{aligned} \tag{3.4.5}$$

We will bound this quantity with the following proposition, which counts the number of $V \in \mathcal{V}_R$ that yield and even graph with at most t edges on a fixed $E \in \mathcal{M}$.

Proposition 3.4.11. *Let $w, h, n \in \mathbb{N}$. Let $\alpha = \alpha_1, \ldots, \alpha_t$ be a sequence of t even numbers so that $\sum_{i=1}^{t} \alpha_i = w \cdot h$. Let $E = M_1, \ldots, M_w$ be a sequence of perfect matchings between two sets of size h.*

Let $\mathcal{G}_{w\times h}^{\alpha,E}$ be the set of all graphs which have a vertex set comprised of w R-multilinear multisets $S_1, \ldots, S_w \in [n]^h$, and have edges forming the perfect matching M_i between S_i, S_{i+1} (where the indexing is modulo w), so that the labels in $[n]$ assigned to the vertices induce exactly t distinct labelings for the edges, and the labeled edges have multiplicities $\alpha_1, \ldots, \alpha_t$. In words, $\mathcal{G}_{w\times h}^{\alpha,E}$ is the set of $w \times h$ matching cycles with matchings specified by E that have edge multiplicities α when labeled with R-multilinear labels from $[n]$.

If $w \cdot h \le n$,

$$|\mathcal{G}_{w,h}^{\alpha,E}| \le (5Rw)^{wh}(wh)^3 \cdot n^{t+h}\,.$$

We will prove this proposition below, in Section 3.4. Applying Proposition 3.4.11 with $h \leftarrow kd/2$, $w \leftarrow 2\ell$, and $t \leftarrow m$, we have that for a fixed $E \in \mathcal{M}$ and for a fixed list of edge multiplicities $a_1, \ldots, a_m$,

$$|\{V \mid (V,E) \text{ has } m \text{ edges with multiplicities } a_1, \ldots, a_m\}| \le (10R\ell)^{dk\ell} \cdot (dk\ell)^2 \cdot n^{m+dk/2}.$$

where we have used the assumption that $dk\ell \ll n$ to meet the requirements of Proposition 3.4.11. The number of possible edge multiplicity lists $a_1, \ldots, a_m$ for a given value of m is at most $\binom{m+dk\ell-1}{m-1} \le 2^{2dk\ell}$. Thus, applying (3.4.15) and noting that there are $|\mathcal{M}|$ choices for E for each V,

$$\begin{aligned}\sum_{V\in\mathcal{V}_R} \mathop{\mathbb{P}}_{E}(\mathcal{E}_E^{kM/2}) &\le \frac{1}{|\mathcal{M}|} \cdot \sum_{m=1}^{kM/2} |\mathcal{M}| \cdot \sum_{a_1,\ldots,a_m} |\{V | (V,E) \text{ has } m \text{ edges w/even mults } a_1, \ldots, a_m\}| \\ &\le \sum_{m=1}^{kM/2} 2^{2dk\ell} \cdot (10R\ell)^{dk\ell} \cdot (dk\ell)^2 \cdot n^{m+dk/2} \;\le\; (40R\ell)^{dk\ell} \cdot (dk\ell)^2 \cdot n^{kM/2+dk/2+1}\,.\end{aligned}$$

Combining (3.4.4) and the above, there is a constant c_2 depending on k so that

$$\begin{aligned}\mathbb{E}[\mathrm{Tr}(C_{(d)}C_{(d)}^\top)^\ell] &\le \left(\frac{k}{2}\right)^{dk\ell} \sum_{M=1}^{d\ell} p^M \cdot \frac{(c_1 R^{k/2+1}\ell)^{2d\ell}}{(dk/2)^{(k/2-1)(2d\ell-M)}} \sum_{V\in\mathcal{V}_R} \mathbb{P}(\mathcal{E}_E^{kM/2}) \\ &\le \left(\frac{c_2 R^{2k+2}\ell^{k+2}}{d^{2(k/2-1)}}\right)^{d\ell} \cdot (dk\ell)^2 n^{dk/2+1} \cdot \sum_{M=1}^{d\ell} \left(pd^{(k/2-1)}n^{k/2}\right)^M\end{aligned}$$

By assumption, $p \cdot (d^{k/2-1}n^{k/2}) \ge 1$, so the term $M = d\ell$ dominates:

$$\begin{aligned}&\le \left(\frac{c_2 R^{2k+2}\ell^{k+2}}{d^{2(k/2-1)}}\right)^{d\ell} \cdot (d\ell)^2 n^{dk/2+1} \cdot d\ell \left(d^{(k/2-1)}pn^{k/2}\right)^{d\ell} \\ &= \left(\frac{c_2 R^{2k+2}\ell^{k+2} \cdot pn^{k/2}}{d^{k/2-1}}\right)^{d\ell} \cdot (d\ell)^3 n^{dk/2+1}\,.\end{aligned}$$

Choosing $\ell = O(d \log n)$, recalling that $R = 100 \log n$, and invoking Proposition 3.2.4, we have that with probability $1 - n^{-100}$, for some constant $c_k := c(k)$,

$$\|C_{(d)}\|^{1/d} \leq c_k \log^{2k} n \cdot \left(\frac{p^{1/2} n^{k/4}}{d^{(k-2)/4}} \right) .$$

The conclusion follows. □

Odd k-XOR

In this section, we modify our algorithm for refuting random even k-XOR instances to handle odd k-XOR instances. The odd k-XOR algorithm is extremely similar to the algorithm for even k-XOR, save for complications introduced by the fact that an odd-order tensor has no natural matrix flattening.

The solution is to apply the Cauchy-Schwarz inequality to the objective value.[5] Let $k = 2\kappa + 1$ for some integer κ. For the tensor $\mathbf{T}_\Phi$ formed by the constraints of Φ and for its $n^\kappa \times n^\kappa$ slices T_i $\forall i \in [n]$, we have that

$$\left\langle x^{\otimes k}, \mathbf{T}_\Phi \right\rangle^2 \leq \left(\sum_{i \in [n]} x_i^2 \right) \left(\sum_{i \in [n]} \left((x^{\otimes \kappa})^\top T_i x^{\otimes \kappa} \right)^2 \right) = n \cdot (x^{\otimes 2\kappa})^\top \left(\sum_{i \in [n]} T_i \otimes T_i \right) x^{\otimes 2\kappa} .$$

Now, the first technicality arises–since the entries $(T_i \otimes T_i)_{(ab),(cd)} = \mathbf{T}_{a,c,i} \cdot \mathbf{T}_{b,d,i}$ are always squares when $a = b$ and $c = d$, we must subtract them from the matrix $\sum_i T_i \otimes T_i$, as otherwise they contribute too much to the norm. Thus, using squares$(\cdot)$ to refer to the part of the matrix for which $a = b$ and $c = d$, we instead will use that the number of constraints $m = (x^{\otimes 2\kappa})^\top$ squares $\left(\sum_{i \in [n]} T_i \otimes T_i \right) x^{\otimes 2\kappa}$, and that

$$\left\langle x^{\otimes k}, \mathbf{T}_\Phi \right\rangle^2 - mn \leq n \cdot (x^{\otimes 2d})^\top \left(\sum_{i \in [n]} T_i \otimes T_i - \text{squares}(T_i \otimes T_i) \right) x^{\otimes 2d} . \tag{3.4.6}$$

We can also view this as doing one step of resolution, so that we have gotten a 4κ-XOR instance starting from a $(2\kappa + 1)$-XOR instance. That is how we will treat our new instance from now on.

Suppose Φ has ± 1-constraint predicates $C_1, \ldots, C_m$, so that $C_a(x) = \eta_a \cdot \prod_{j \in S_a} x_j$. We create a new $2(k-1)$-XOR instance Ψ as follows. For each $a, b \in [n]$, $a \neq b$: if C_a and C_b both contain the variable i in the kth position, add the ± 1 constraint predicate $C'_{ab}(x) = \eta_a \cdot \eta_b \cdot \left(\prod_{j \in S_a} x_j \right) \left(\prod_{j \in S_a} x_j \right)$ to Ψ. Let m' be the number of clauses in Ψ.

[5] We remark that this idea is not new, and has appeared before (as early as e.g. [FG01])—however it does introduce some new challenges in our analysis.

The right-hand side of (3.4.6) is $n \cdot \sum_{ab} C'_{ab}(x) = n \cdot 2m' \cdot (P_\Psi(x) - \frac{1}{2})$, where $P_\Psi(x)$ is the fraction of clauses of Ψ satisfied by x. Combining this with the above calculations,

$$\left(2m\left(P_\Phi(x) - \frac{1}{2}\right)\right)^2 \leq nm + n \cdot 2m' \cdot \left(P_\Psi(x) - \frac{1}{2}\right),$$

$$P_\Phi(x) \leq \frac{1}{2} + \frac{1}{2m}\sqrt{nm + 2nm' \cdot \left(P_\Psi(x) - \frac{1}{2}\right)}. \tag{3.4.7}$$

Now, we will essentially apply our even-k-XOR strategy to Ψ. The only issue is that the clauses of Ψ are not independent, so we will need to zero out not only rows and columns indexed by high-multiplicity subsets of $[n]^{2\kappa}$, but also get rid of terms that contain the same slice with too high a multiplicity. So, instead of taking the dth tensor power of the matrix $\sum_i T_i \otimes T_i - \text{squares}(T_i \otimes T_i)$, we omit the cross-products in which $T_i \otimes T_i$ appears more than $100 \log n$ times for any $i \in [n]$.

Formalizing this, we introduce our matrix certificate for the odd case:

Algorithm 3.4.12 (Odd k-XOR certificate at level d)**.**
Input: A k-XOR instance for odd $k = 2\kappa + 1$ on n variables with m clauses $C_1, \ldots, C_m$, where $C_a(x) = \eta_a \cdot \prod_{j \in S_a} x_j$ for $S_j \in [n]^k$ and $\eta_a \in \{\pm 1\}$.

1. Form the tensor $\mathbf{T} := \mathbf{T}_\Phi$ by setting $\mathbf{T}_{S_a} = b_a$ for all $a \in [m]$, and setting all other entries to 0.
2. Initialize an empty $n^{2d\kappa} \times n^{2d\kappa}$ matrix Γ.
3. For each ordered multiset $U \in [n]^d$ in which no entry appears with multiplicity $> 100 \log n$:
   a) Add the squared tensor of the slices of $\mathbf{T}_\Phi$ corresponding to the indices in U:

$$\Gamma := \Gamma + \bigotimes_{i \in U} (T_i \otimes T_i - \text{squares}(T_i \otimes T_i))$$

   where squares$(\cdot)$ is the restriction to entries $(I, J), (K, L)$ such that $(I, K) = (J, L)$ as ordered multisets.
4. Letting $\hat{\mathcal{S}}_{2d\kappa}$ be the set of all permutation matrices that perform the index permutations corresponding to $\mathcal{S}_{2d\kappa}$ on the rows and columns of Γ, form

$$\Gamma_{(d)} := \mathop{\mathbb{E}}_{\Pi, \Sigma \in \hat{\mathcal{S}}_{2d\kappa}} [\Pi \Gamma \Sigma] .$$

5. Set to zero all rows and columns of $\Gamma_{(d)}$ indexed by multisets $S \in [n]^{2d\kappa}$ which contain some element of $[n]$ with multiplicity $> 100 \log n$.

Output: The value $\|\Gamma_{(d)}\|$.

The following theorem, which is the main theorem of this section, gives a bound on the value output by Algorithm 3.4.3.

Theorem 3.4.13. *Let $k, n, d \in \mathbb{N}$, so that $d \log n \ll n$, and furthermore let k be odd so that $k = 2\kappa + 1$. Let Φ be a random instance of k-XOR on n variables $x_1, \dots, x_n$, with $\Theta(pn^k)$ clauses (so each constraint is sampled uniformly and independently with probability p). Let $\Gamma_{(d)}$ be the matrix formed from the instance Φ as described in Algorithm 3.4.12. Then if $d^{(k-2)/2} n^{k/2} p > 1$, there exists a constant c_k depending on k such that with high probability over the choice of Φ,*

$$\|\Gamma_{(d)}\|^{1/d} \leq \tilde{O}\left(\frac{pn^{k/2}}{d^{(k-2)/2}}\right).$$

We will prove the theorem below, in Section 3.4, and we will now show how to take this matrix and acquire a certificate from it.

Validity of the Odd Certificate

Again, our strategy will be to work with the polynomial $(P_\Psi(x))^d$, which is not much altered by removing terms corresponding to the high-multiplicity rows, columns, or slice cross-products.

Proposition 3.4.14. *Let Φ be a random k-XOR formula in which each clause is sampled independently with probability p, and let Ψ be the $2(k-1)$-XOR instance obtained from Φ as described above, where Ψ has m' clauses $\{C_{ab}\}_{ab}$ corresponding to pairs of clauses from Φ sharing the same final variable.*

Let $\hat{\mathcal{C}}^d_{low} \subset [m']^d$ be the set of all ordered multisets of clauses $C_{a_1 b_1}, \dots, C_{a_d b_d}$ from Ψ with the property that if we form three multisets of variables, $I, J \in [n]^{d(k-1)}$ and $S \in [n]^d$, with I containing the first $(k-1)/2$ variables of each C_{a_ℓ}, C_{b_ℓ}, J containing the next $(k-1)/2$ variables of each C_{a_ℓ}, C_{b_ℓ}, and S containing the last (shared) variable of C_{a_ℓ} and C_{b_ℓ}, then I, J are both low-multiplicity multisets, in that both have no element of $[n]$ with multiplicity $> 100 \log n$.

Let $o_{\max}$ be the maximum number of clauses of Ψ any variable appears in. Then if $d \ll n$ and $d(2k-1)o_{\max} < 200\varepsilon m' \log n$,

$$P_\Psi(x) \leq \left(\mathop{\mathbb{E}}_{a_1 b_1, \dots, a_d b_d \sim \hat{\mathcal{C}}^d_{low}} \left[\prod_{\ell=1}^{d} \frac{1}{2} \left(1 - C_{a_\ell}(x) C_{b_\ell}(x)\right) \right] \right)^{1/d} + \varepsilon$$

for all $x \in \{\pm 1\}^n$ with high probability. When $p \geq 200 \frac{\log n}{n^{k-1}}$, we have that $\varepsilon = o(1)$ with high probability.

Proof. The proof is very similar to that of Proposition 3.4.14. First, let m' be the number of clauses in Ψ, and let $o_{\max}$ be the maximum number of clauses of Ψ that any variable $i \in [n]$ appears in (even if it the shared variable and is included with multiplicity 2).

By definition, we have that $P_\Psi(x)$ gives the proportion of satisfied clauses, so

$$(P_\Psi(x))^d = \mathop{\mathbb{E}}_{a_1 b_1, \dots, a_d b_d \sim [m']^d} \left[\prod_{\ell=1}^{d} \frac{1}{2} \left(1 - C_{a_\ell}(x) C_{b_\ell}(x)\right) \right]. \tag{3.4.8}$$

Since only a $o(1)$ fraction of the multisets of indices, $[n]^{dk}$, will not contain any item with multiplicity more than $100 \log n$, we will be able to prove that those terms contribute negligibly.

We sample a uniform element $\mathcal{C} \sim \hat{\mathcal{C}}^d_{low}$, $\mathcal{C} = (C_{a_1}, C_{b_2}), \ldots, (C_{a_d}, C_{b_d})$ in the following way. For $t = 1, \ldots, d$:

- Let $\mathcal{A}_t \subset \mathcal{I}$ be the set of pairs of clauses such that for any $(C', C'') \in \mathcal{A}$, $(C_{a_1}, C_{b_1}), \ldots, (C_{b_{t-1}}, C_{b_{t-1}}), (C', C'') \in \hat{\mathcal{C}}^t_{low}$, that is, the set of clauses from Ψ that maintain the low-multiplicity conditions.
- Choose a uniformly random $(C', C'') \sim \mathcal{A}_t$ and set $C_{a_t}, C_{b_t} := (C', C'')$, adding (C', C'') to $\mathcal{C}$.

This sampling process clearly gives a uniformly random element of $\hat{\mathcal{C}}^t_{low}$.

Claim 3.4.15. At step $t+1$ there are at least $m' - t\frac{(2k-1)\cdot o_{\max}}{R}$ clauses that can be added.

Proof. In order to exclude any variable $i \in [n]$, we must add at least $100 \log n$ copies of i.. Further, to exclude ℓ distinct variables in $[n]$, we must add at least $100 \log n$ copies of each variable, for a total of $100\ell \log n$ variables, which requires adding at least $100\ell \log n/(2k-1)$ pairs (since each pair contains $2k-1$ variables). If ℓ distinct variables are excluded, then at most $\ell \cdot o_{\max}$ pairs of clauses are excluded. The claim now follows. □

Now, define the random variable $X_t = \prod_{j=1}^{t} \frac{1}{2}(1 - C_{a_j}(x)C_{b_j}(x))$–this is the 0-1 value of x on $\mathcal{C}_t$. We apply Claim 3.4.15, along with the observation that $P_\Psi(x)$ can only drop by $1/m'$ for each clause pair removed, to conclude that

$$\mathbb{E}[X_{t+1} | C_{a_1 b_1}, \ldots, C_{a_t b_t}] \geq \left(P_\Psi(x) - t\frac{(2k-1)\cdot o_{\max}}{100m' \log n} \right) \cdot X_t \,.$$

From this we have that $\mathbb{E}[X_t] \geq \left(P_\Psi(x) - t\frac{(2k-1)o_{\max}}{100m' \log n} \right) \cdot \mathbb{E}[X_{t-1}]$ from which we have that as long as $d(2k-1)o_{\max} \leq \varepsilon 100m' \log n$,

$$\mathbb{E}[X_k] \geq \prod_{t=1}^{d} \left(P_\Psi(x) - t\frac{(2k-1)o_{\max}}{100m' \log n} \right) \geq \left(P_\Psi(x) - \varepsilon \right)^d \,.$$

So taking $\varepsilon = 1/\log n$, if we can establish that the inequality $d(2k-1)o_{\max} \leq 100m'$ with high probability when $p \geq \Omega(\log n / n^{k-1})$, then we are done.

A Chernoff bound implies that $pn^k/2 \leq m \leq 2pn^k$ with probability at least $1 - \exp(-\Omega(pn^k))$, and that each variable's degree m_i is $pn^{k-1}/2 \leq m_i \leq 2pn^{k-1}$ with probability at least $1 - \exp(-\Omega(pn^{k-1}))$. We have that $m' = (\sum_i m_i^2) - m$, and so by a union bound and using the assumption that $pn^{k-1} \geq \Omega(\log n)$, we have that

$$m' \geq p^2 n^{2k-1}/4.$$

Let o_i be the degree of variable i in Ψ. To bound o_{max}, we observe that o_i is made up of occurrences of pairs in which i is the shared variable, and of pairs in which i is not the shared

variable. The contribution of the first category is m_i^2, and with high probability by our union bound $m_i^2 \leq (2pn^{k-1})^2$. In the second category, we have $\sum_j m_j \cdot m_{ij}$, where m_{ij} is the number of clauses containing i and j. By our previous assumption regarding the concentration of the m_i, we have that $\sum_j m_j \cdot m_{ij} \leq 2pn^{k-1} \sum_j m_{ij}$. The quantity $\sum_j m_{ij} = m_i$, and so we can conclude that $o_{\max} \leq 4p^2 n^{2k-2}$, so that $o_{\max}/m' \leq 16/n$, yielding our result. □

The proof above can be modified to give the following lemma, which gives a lower bound on the number of low-multiplicity terms.

Lemma 3.4.16. *If Φ is a random k-XOR instance, then so long as $d \ll n$ and $d(2k-1)o_{max} < 200\varepsilon m' \log n$,*

$$|\hat{\mathcal{C}}^d_{low}| \geq ((1-\varepsilon)m')^d \cdot \ .$$

Furthermore, $\varepsilon = o(1)$ with high probability when $p \geq \Omega(n^{-k+1} \log n)$.

The proof proceeds exactly as the proof of Proposition 3.4.14, but instead of bounding the decrease in the value as each clause is added, one bounds the probability that a clause the multiplicity restriction is chosen.

Now, we have that

$$\begin{aligned}
&\mathop{\mathbb{E}}_{a_1 b_1, \ldots, a_d b_d \in \hat{\mathcal{C}}^d_{low}} \left[\prod_{\ell=1}^{d} \frac{1}{2}(1 + C_{a_\ell}(x) C_{b_\ell}(x)) \right] \\
&\qquad = \frac{1}{2^d} \sum_{S \subseteq [d]} \mathop{\mathbb{E}}_{a_1 b_1, \ldots, a_d b_d \in \hat{\mathcal{C}}^d_{low}} \left[\prod_{\ell \in S} C_{a_\ell}(x) C_{b_\ell}(x) \right]
\end{aligned}$$

and by the symmetry of the uniform distribution over $\hat{\mathcal{C}}^d_{low}$,

$$\begin{aligned}
&= \frac{1}{2^d} \sum_{j=0}^{d} \binom{d}{j} \cdot \mathop{\mathbb{E}}_{a_1 b_1, \ldots, a_j b_j \in \hat{\mathcal{C}}^j_{low}} \left[\prod_{\ell=1}^{j} C_{a_\ell}(x) C_{b_\ell}(x) \right] \\
&\leq \frac{1}{2^d} \sum_{j=0}^{t} \binom{d}{j} + \frac{1}{2^d} \sum_{j=t+1}^{d} \mathop{\mathbb{E}}_{a_1 b_1, \ldots, a_j b_j \in \hat{\mathcal{C}}^j_{low}} \left[\prod_{\ell=1}^{j} C_{a_\ell}(x) C_{b_\ell}(x) \right] \\
&= \frac{1}{2^d} \sum_{j=0}^{t} \binom{d}{j} + \frac{1}{2^d} \sum_{j=t+1}^{d} \binom{d}{j} \cdot \frac{1}{|\hat{\mathcal{C}}^j_{low}|} (x^{\otimes 2j\kappa})^\top \Gamma_{(j)} x^{\otimes 2j\kappa} \ .
\end{aligned} \tag{3.4.9}$$

Now, we can use the spectral norm of $\Gamma_{(j)}$ as a certificate, for values of $j \in [\delta d, d]$– we stitch together the details below. The following concludes the proof of the refutation theorem, modulo the proof of the $\Gamma_{(d)}$ matrix norm bound from Theorem 3.4.13, which we give in the next subsection.

Theorem 3.4.17. *Let $k = 2\kappa + 1$ be odd, and let $d \ll n$. Then for sufficiently large n there is an algorithm that with high probability certifies that a random k-XOR instance has value at most $\frac{1}{2} + \gamma + o(1)$ for any constant $\gamma > 0$ at clause density $m/n = \tilde{O}\left(\frac{n^{(k-2)/2}}{d^{(k-2)/2}}\right)$ (where the $\tilde{O}$ hides a dependence on γ and k) in time $n^{O(d)}$.*

Proof. As argued in the proof of Proposition 3.4.14, we have that $m' = \Theta(p^2 n^{2k-1})$ and $m = \Theta(pn^k)$ with high probability, as long as $p \geq \Omega(\frac{\log n}{n^{k-1}})$. Suppose that no variable appears in more than $o_{\max}$ clauses in Ψ. Then for $\beta = \frac{d(2k-1)o_{\max}}{200m' \log n}$, from Proposition 3.4.14 and (3.4.9),

$$(P_\Psi(x) - \beta)^d \leq \frac{1}{2^d} \sum_{j=0}^{t} \binom{d}{j} + \sum_{j=t+1}^{d} \frac{\binom{d}{j}}{2^d} \cdot \frac{1}{|\hat{\mathcal{C}}^j_{low}|} (x^{\otimes 2j\kappa})^\top \Gamma_{(j)} x^{\otimes 2j\kappa}.$$

If we choose $t = \delta d$ for some constant $\delta > 0$, then we can bound the jth term in the second summation by

$$\frac{1}{|\hat{\mathcal{C}}^j_{low}|} (x^{\otimes 2j\kappa})^\top \Gamma_{(d)} x^{\otimes 2j\kappa} = \frac{\|x\|^{2j(k-1)}}{|\hat{\mathcal{C}}^j_{low}|} \cdot \|\Gamma_{(j)}\|$$

By Lemma 3.4.16 and Theorem 3.4.13

$$\begin{aligned} &= \frac{n^{j(k-1)}}{|\hat{\mathcal{C}}^j_{low}|} \cdot \|\Gamma_{(j)}\| \\ &\leq \frac{n^{j(k-1)}}{(1-\beta))^j (m')^j} \cdot \tilde{O}\left(\frac{pn^{k/2}}{j^{(k-2)/2}}\right)^j, \\ &\leq \tilde{O}\left(\frac{1}{(1-\beta) pn^{k/2} j^{(k-2)/2}}\right)^j, \qquad (\text{since } m' = \Theta(p^2 n^{2k-1}) \text{ w.h.p.}) \end{aligned}$$

where the inequality holds with high probability from the conditions of Theorem 3.4.13, and therefore also holds with high probability simultaneously for all $j \in [\delta k, k]$ by a union bound.

The term comprised of the sum of binomial coefficients is at most

$$\frac{1}{2^d} \sum_{j=0}^{\delta d} \binom{d}{j} \leq 2^{(H(\delta)-1)d},$$

Where $H(\cdot)$ is the binary entropy function, $H(\delta) = -\delta \log \delta - (1-\delta)\log(1-\delta)$. Also, $\beta = o(1)$ with high probability. Therefore, for some $\alpha \in [\delta, 1]$,

$$P_\Psi(x) \leq \left(2^{(H(\delta)-1)d} + \tilde{O}\left(\frac{1}{pn^{k/2}(\alpha d)^{(k-2)/2}}\right)^{\alpha d}\right)^{1/d} + o(1).$$

Now, for $p \geq \tilde{O}(n^{-(k/2)} d^{-((k-2)/2)})$, the latter quantity is $o(1)$, where the $\tilde{O}$ hides a polylog n and a dependence on δ and k. Thus, for n sufficiently large the full term is at most $(1 +$

$o(1))2^{H(\delta)-1}$. We can choose δ sufficiently small so as to bound $P_\Psi(x) \le \frac{1}{2} + \varepsilon$ for any constant $\varepsilon > 0$.

Using the fact that c, $m' \le 4p^2n^{2k-1}$ with high probability for sufficiently large $p \ge \tilde{O}(n^{-k+1})$ (see the proof of Proposition 3.4.14), and that $m \ge pn^k/2$ with high probability, combining with (3.4.7) we have that

$$\begin{aligned} P_\Phi(x) &\le \frac{1}{2} + \frac{1}{2m}\sqrt{nm + 2nm' \cdot \left(P_\Psi(x) - \frac{1}{2}\right)} \\ &\le \frac{1}{2} + \sqrt{\Theta\left(\frac{1}{pn^{k-1}}\right) + \frac{\varepsilon nm'}{2m^2}} \\ &\le \frac{1}{2} + \sqrt{o(1) + \frac{\varepsilon 4p^2n^{2k}}{2(\frac{1}{2})^2p^2n^{2k}}} \;=\; \frac{1}{2} + \sqrt{\Theta\left(\frac{1}{pn^{k-1}}\right) + 8\varepsilon} \end{aligned}$$

and we can take the quantity within the square root to be an arbitrarily small constant by choosing a constant ε sufficiently small.

We can certify this bound in time $n^{O(d)} \cdot d$, by running Algorithm 3.4.12 to compute the top eigenvalue of $\Gamma_{(j)}$ for each $j \in [\delta d, d]$. □

Bounding the odd certificate spectral norm

Now, we bound $\|\Gamma_{(d)}\|$ for the matrices $\Gamma_{(d)}$ defined above in Algorithm 3.4.12. Before stating our theorem, we describe the hypergraphs corresponding to entries of $\Gamma_{(d)}$, and to $((\Gamma_{(d)})(\Gamma_{(d)})^\top)^\ell$.

We obtain $\Gamma_{(d)}$ by averaging over row and column symmetries of the matrix

$$\sum_{i\in[n]} \sum_{\substack{U \\ U\in[n]^d \\ \text{low-mult}}} \bigoplus_{u\in U} T_u \otimes T_u - \text{squares}(T_u \otimes T_u)\,,$$

then setting rows and columns indexed by high-multiplicity multisets to 0. Ignoring the subtracted squares for now, this can in turn be understood as the low-multiplicity restriction of the matrix

$$\left(\sum_u T_u \otimes T_u\right)^{\otimes d},$$

where the low-multiplicity restriction is occurring on the Cauchy-Schwarz'd mode u, as well as on the rows and columns. We begin with the hypergraph interpretation of the matrix $(\sum_u T_u \otimes T_u)^{\otimes k}$, from which the interpretation for $\Gamma_{(d)}$ will follow by our understanding of symmetrization over $\mathcal{S}_{2d\kappa}$ and of low-multiplicity restrictions. Let $M := (\sum_u T_u \otimes T_u)^{\otimes d}$ for convenience. We have that the $(A,B),(C,D)$th entry of M (for $A,B,C,D \in [n]^{d\kappa}$ with $A = a_1,\dots,a_d$ with $a_i \in [n]^\kappa$, and with similar decompositions defined for B,C,D) has value

$$M_{(A,B),(C,D)} = \prod_{i\in[d]} \left(\sum_{u\in[n]} \mathbf{T}_{a_i,c_i,u} \cdot \mathbf{T}_{b_i,d_i,u}\right) = \sum_{U\in[n]^d} \prod_{i\in[d]} \left(\mathbf{T}_{a_i,c_i,u_i} \cdot \mathbf{T}_{b_i,d_i,u_i}\right).$$

Interpreting the variables $\mathbf{T}_{a_i,c_i,u_i}$ as $k = (2\kappa+1)$-uniform hyperedges, we have that each entry is a sum over hypergraphs indexed by $U \in [n]^d$. For each $U \in [n]^d$, we have a hypergraph on the following vertex configuration: on the left, we have the vertices from the multiset A, B. On the right, we have the vertices from the multiset C, D. In the center, we have the vertices from U. On this vertex set, we have $2d$ hyperedges. Of these hyperedges, d form a tripartite matching on the vertices in A, C, U, with κ vertices from each of A, C and one vertex in U. The other d form a similar tripartite matching on the vertices in B, D, U. Every hyperedge on A, C, U shares exactly one vertex in U with exactly one hyperedge from B, D, U. See Figure 3.2 for an illustration.

Now, we detail the impact of subtracting the squares, and of removing high-multiplicity rows, columns, and Kronecker powers.

- The subtraction of the square terms squares$(T_u \otimes T_u)$ forces us to never have two hyperedges sharing a vertex in U if they contain vertices of the same type in $[n]$: that is, we can never have $(a_i, c_i) = (b_i, d_i)$ as ordered multisets.
- The deletion of high-multiplicity indices, both in the Cauchy-Schwarz'd mode and in the rows and columns, forces us to exclude hypergraphs with (A, B), (C, D), or U containing more than $100 \log n$ repetitions of any one vertex type.
- The averaging operation $\mathbb{E}_{\Pi,\Sigma\in\hat{\mathcal{S}}_{2d\kappa}}$ takes each such entry to an average over all allowed hyperedge configurations on the vertex set $(A, B), (C, D), U$.

When we take $\mathrm{Tr}(\Gamma_{(d)}\Gamma_{(d)}^\top)^\ell$, we are taking a sum over all "cycles" of length 2ℓ in such hypergraphs, where the vertices in the cycle are given by the (A, B) multisets, and the edges are given by the average hyperedge configuration between (A, B) and the next (C, D), with the U vertices in between.

We are now ready to prove our upper bound on $\Gamma_{(d)}$.

Theorem (Restatement of Theorem 3.4.13)**.** *Let $k, n, d \in \mathbb{N}$, so that $d \log n \ll n$, and furthermore let k be odd so that $k = 2\kappa + 1$. Let Φ be a random instance of k-XOR on n variables $x_1, \ldots, x_n$, with $\Theta(pn^k)$ clauses (so each constraint is sampled uniformly and independently with probability p). Let $\Gamma_{(d)}$ be the matrix formed from the instance Φ as described in Algorithm 3.4.12. Then if $d^{(k-2)/2}n^{k/2}p > 1$, there exists a constant c_k depending on k such that with high probability over the choice of Φ,*

$$\|\Gamma_{(d)}\|^{1/d} \leq \tilde{O}\left(\frac{pn^{k/2}}{d^{(k-2)/2}}\right).$$

Proof. We fix d and take $\Gamma := \Gamma_{(d)}$ for convenience. Also, fix $R := 100 \log n$, and call a multiset $S \in [n]^m$ R-multilinear if no element of $[n]$ appears with multiplicity more than R in S. We will apply the trace power method (Proposition 3.2.4) to Γ, for which it suffices to obtain an upper bound on

$$\mathbb{E}\left[\mathrm{Tr}\left(\Gamma_{(d)}\Gamma_{(d)}^\top\right)^\ell\right].$$

As described above, this amounts to bounding the number of hypergraph cycles of length 2ℓ, where each "vertex" of the cycle is comprised of an R-multilinear multiset $(A, B) \in [n]^{2d\kappa}$, and the "edges" in the cycle between (A, B) and (C, D) are the sum over all R-multilinear $U \in [n]^d$ of the average over all possible hypergraphs (because of the symmetries) that

contain a tripartite hypergraph matching with $2d$ edges between $(A, B), U, (C, D)$ in which every hyperedge contains d vertices from (A, B), d vertices from (C, D), and one vertex from U. Hypergraphs that contain two identical hyperedges sharing a vertex from U have contribution 0 to the sum (due to the subtraction of the "squares").

Let the set of all valid R-multilinear vertex configurations V comprising $V = S_1, \ldots, S_{2\ell} \in [n]^{2d\kappa}$ be denoted $\mathcal{V}_R$. Let the set of all R-multilinear center vertex configurations $U = U_1, \ldots, U_{2\ell} \in [n]^d$ be denoted $\mathcal{U}_R$. Let the set of all hyperedge matching sequences $H = H_1, \ldots, H_{2\ell}$ with H_i a matching between S_i, U_i, S_{i+1} be denoted $\mathcal{H}$. For $H \in \mathcal{H}, V \in \mathcal{V}_R, U \in \mathcal{U}_R$, denote by (V, U, H) the hypergraph given by the hyperedges H on the vertex configuration V. We think of the elements in the sum $\mathrm{Tr}(\Gamma\Gamma^\top)^\ell$ as being indexed by $\mathbb{E}_H[(V, U, H)]$, where the expectation over H is a result of our symmetrization/averaging operation.

Applying the above observations, and recalling that we have assembled Γ from the random tensor $\mathbf{T}$, we have that

$$\mathop{\mathbb{E}}_{\mathbf{T}}[\mathrm{Tr}(\Gamma\Gamma^\top)^\ell] = \sum_{U \in \mathcal{U}_R} \sum_{V \in \mathcal{V}_R} \mathop{\mathbb{E}}_{H \in \mathcal{H}} \left[\mathop{\mathbb{E}}_{\mathbf{T}} \left[\prod_{(i_1, \ldots, i_d) \in (V, U, H)} \mathbf{T}_{i_1, \ldots, i_d} \right] \right] ,$$

Because $\mathbf{T}_S \neq \mathbf{T}_{\pi(S)}$, our hyperedges are ordered, and so two hyperedge variables are not identical unless the vertices appear in the same order (in particular, the partition into a_i, b_i, u_i and the order within each should be the same). The expectation over $\mathbf{T}$ of a term is 0 if any ordered hyperedge in (V, U, H) appears with odd multiplicity or if two identical ordered hyperedges share the same vertex in some U_i, and is p^M if exactly M distinct hyperedges appear in (V, U, H). Thus,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{\mathbf{T}}[\mathrm{Tr}(\Gamma\Gamma^\top)^\ell] \\
&\leq \sum_{U \in \mathcal{U}_R} \sum_{V \in \mathcal{V}_R} \sum_{M=1}^{d\ell} p^M \cdot \mathop{\mathbb{E}}_{H \in \mathcal{H}} [\mathbb{I}((V, U, H) \text{ even, nonsharing}) \cdot \mathbb{I}((V, U, H) \text{ has } M \text{ hyperedges})] \\
&= \sum_{U \in \mathcal{U}_R} \sum_{V \in \mathcal{V}_R} \sum_{M=1}^{d\ell} p^M \cdot \mathop{\mathbb{P}}_{H \in \mathcal{H}} [(V, U, H) \text{ even, nonsharing with } M \text{ hyperedges})] \, . \qquad (3.4.10)
\end{aligned}$$

To bound this probability, we will sample uniformly $U \in \mathcal{U}_R$ and $H \sim \mathcal{H}$ in a three-step process.

1. Fix $V \in \mathcal{V}_R$.
2. Sample a uniformly random perfect matching (with 2-edges rather than hyperedges) between each set $S_i, S_{i+1} \in \mathcal{V}_R$–call the edge set sampled in this manner E, so that we now have the graph (V, E).
3. Sample a hyperedge matching configuration G from E by choosing a uniform random grouping of the edges between S_i, S_{i+1} into groups of d edges, to obtain the hypergraph (V, G) (when $k = 3 \implies \kappa = 1$, this step is skipped).
4. Sample a pairing H of the hyperedges in G, a center vertex for each pair in H and an order on the center vertices to form U, to obtain the hypergraph (V, U, H).

For step 2 and 3, we will employ the same Proposition 3.4.11 and Lemma 3.4.10 that we used in the proof of the even case (Theorem 3.4.4) to bound the probability that we sample a (V, E) and (V, G) with a certain edge multiplicity and the evenness property. For step 4, we will need another lemma along the same lines.

We note that if (V, U, H) has every hyperedge appearing an even number of times and there are M distinct edges, then even if all center vertices are removed to obtain a 2κ-hypergraph (V, G), every ordered hyperedge must still appear with even multiplicity, and there can only be at most M distinct hyperedges. Therefore, letting $\mathcal{E}_H^M$ be the event that (V, U, H) is even with M edges and no square/sharing hyperedges, letting $\mathcal{E}_G^m$ be the event that (V, G) is even with at most m edges, we have that

$$\mathbb{P}[\mathcal{E}_H^M] = \sum_{m \le M} \mathbb{P}[\mathcal{E}_H^M, \mathcal{E}_G^m] = \sum_{m \le M} \mathbb{P}[\mathcal{E}_H^M | \mathcal{E}_G^m] \cdot \mathbb{P}[\mathcal{E}_G^m]\,. \tag{3.4.11}$$

Now, let $\mathcal{E}_E^{\le m'}$ be the event that (V, E) is a simple graph with the evenness property and at most m' distinct edges. We use the asymmetry of $\mathbf{T}$ to invoke Lemma 3.4.9 with $h \leftarrow 2d$, $w \leftarrow 2\ell$, $\tau \leftarrow M$ and $\kappa \leftarrow \kappa$ which gives us that $\mathbb{P}(\mathcal{E}_G^m \mid \mathcal{E}_E^{\le dm}) \ge (\kappa!)^{-4d\ell}$. From this, we have

$$\begin{aligned}
\mathbb{P}[\mathcal{E}_H^M] &= \sum_{m \le M} \mathbb{P}[\mathcal{E}_H^M | \mathcal{E}_G^m] \cdot \mathbb{P}[\mathcal{E}_G^m] \qquad \text{(by (3.4.11))} \\
&= \sum_{m \le M} \mathbb{P}[\mathcal{E}_H^M | \mathcal{E}_G^m] \cdot \frac{\mathbb{P}[\mathcal{E}_G^m, \mathcal{E}_E^{\le \kappa m}]}{\mathbb{P}[\mathcal{E}_E^{\le \kappa m} \mid \mathcal{E}_G^m]} \\
&\le (\kappa!)^{4d\ell} \sum_{m \le M} \mathbb{P}[\mathcal{E}_H^M | \mathcal{E}_G^m] \cdot \mathbb{P}[\mathcal{E}_G^m, \mathcal{E}_E^{\le \kappa m}] \qquad \text{(by Lemma 3.4.9)} \\
&\le \kappa^{4d\ell\kappa} \sum_{m \le M} \mathbb{P}[\mathcal{E}_H^M | \mathcal{E}_G^m] \cdot \mathbb{P}[\mathcal{E}_G^m \mid \mathcal{E}_E^{\le \kappa m}] \cdot \mathbb{P}[\mathcal{E}_E^{\le \kappa m}]\,.
\end{aligned} \tag{3.4.12}$$

We will bound $\mathbb{P}(\mathcal{E}_G^m | \mathcal{E}_E^{\le \kappa m})$, using Lemma 3.4.10 with $h \leftarrow 2d$,$w \leftarrow 2\ell$, $\tau \leftarrow m$, which gives us that

$$\mathbb{P}(\mathcal{E}_G^m | \mathcal{E}_E^{\le dm}) \le \mathbb{P}((V, G) \text{ even with } m \text{ edges} \mid (V, E) \text{ even }) \le \frac{(4e^\kappa \kappa^\kappa R^{\kappa+1} \ell)^{4d\ell}}{(2d\kappa)^{(\kappa-1)(4d\ell - m)}}\,.$$

And so now, combining with (3.4.12), we have that for some constant c_1 depending on κ,

$$\begin{aligned}
\mathbb{P}[\mathcal{E}_H^M] &\le \kappa^{4d\ell\kappa} \sum_{m \le M} \mathbb{P}[\mathcal{E}_H^M | \mathcal{E}_G^m] \cdot \left(\frac{(4e^\kappa \kappa^\kappa R^{\kappa+1} \ell)^{4d\ell}}{(2d\kappa)^{(\kappa-1)(4d\ell - m)}} \right) \cdot \mathbb{P}[\mathcal{E}_E^{\le \kappa m}] \\
&\le \frac{(c_1 R^{d+1} \ell)^{4d\ell}}{d^{(\kappa-1)4d\ell}} \sum_{m \le M} d^{(\kappa-1)m} \cdot \mathbb{P}[\mathcal{E}_H^M | \mathcal{E}_G^m] \cdot \mathbb{P}[\mathcal{E}_E^{\le \kappa m}]
\end{aligned} \tag{3.4.13}$$

And therefore, with (3.4.10),

$$\mathbb{E}[\mathrm{Tr}(\Gamma\Gamma^\top)^\ell] = \frac{(c_1 R^{\kappa+1}\ell)^{4d\ell}}{d^{(\kappa-1)4d\ell}} \cdot \sum_{M=1}^{d\ell} p^M \cdot \sum_{m\le M} \sum_{\substack{U\in[n]^d \\ R-\text{multi}}} d^{(\kappa-1)m} \cdot \mathbb{P}[\mathcal{E}_H^M | \mathcal{E}_G^m] \sum_{V\in\mathcal{V}_R} \cdot \mathbb{P}[\mathcal{E}_E^{\le \kappa m}] \tag{3.4.14}$$

We now bound $\mathbb{P}(\mathcal{E}_E^{\le\kappa m})$. If we interchange the order of summation, sum over these probabilities for a fixed value of m, letting $\mathcal{M}$ be the set of all possible edge configurations E, we have

$$\begin{aligned}\sum_{V\in\mathcal{V}_R} \mathop{\mathbb{P}}_E(\mathcal{E}_E^{\kappa m}) &= \sum_{V\in\mathcal{V}_R} \frac{|\{E \mid (V,E) \text{ even with} \le \kappa m \text{ edges}\}|}{|\mathcal{M}|} \\ &= \frac{|\{E, V \mid (V,E) \text{ even with} \le \kappa m \text{ edges}\}|}{|\mathcal{M}|}. \end{aligned} \tag{3.4.15}$$

We will bound this quantity with Proposition 3.4.11, which counts the number of $V \in \mathcal{V}_R$ that yield and even graph with at most m edges on a fixed $E \in \mathcal{M}$. From Proposition 3.4.11 with $w \leftarrow 2\ell$, $h \leftarrow 2d$, $t \leftarrow m'$, we have that for a fixed $E \in \mathcal{M}$ and for a fixed list of edge multiplicities $a_1, \dots, a_{m'}$,

$$|\{V \mid (V,E) \text{ has } m' \text{ edges with multiplicities } a_1,\dots,a_{m'}\}| \le (c_2 R\ell)^{4d\kappa\ell} \cdot (d\ell)^2 \cdot n^{m'+2d\kappa}$$

for some constant c_2 depending on k, where we have used the assumption that $d \ll n$ to meet the requirements of Proposition 3.4.11. The number of possible edge multiplicity lists $a_1, \dots, a_{m'}$ for a given value of m' is at most $\binom{m'+4d\kappa\ell-1}{m'-1} \le 2^{4dk\ell}$. Thus,

$$\begin{aligned}&\frac{|\{E, V \mid (V,E) \text{ even with} \le \kappa m \text{ edges}\}|}{|\mathcal{M}|} \\ &\quad\le \frac{1}{|\mathcal{M}|} \cdot \sum_{m'=1}^{\kappa m} \sum_{a_1,\dots,a_{m'}} |\mathcal{M}| \cdot |\{V \mid (V,E) \text{ has } m' \text{ edges with even mult.s } a_1,\dots,a_{m'}\}| \\ &\quad\le \sum_{m'=1}^{\kappa m} 2^{4dk\ell} \cdot (c_2 R\ell)^{4d\kappa\ell} \cdot (d\ell)^2 \cdot n^{m'+2d\kappa} \;\le\; (c_3 R\ell)^{4dk\ell} \cdot (d\ell)^2 \cdot n^{\kappa m + 2d\kappa + 1}, \end{aligned}$$

for some constant c_3 depending on k. So there is a constant c_4 depending on k so that,

$$\begin{aligned}&\mathbb{E}[\mathrm{Tr}(\Gamma\Gamma^\top)^\ell] \\ &\le \frac{(c_2 R^{\kappa+1}\ell)^{4d\ell}}{d^{(\kappa-1)4d\ell}} \cdot \sum_{M=1}^{d\ell} p^M \cdot \sum_{m\le M} \sum_{\substack{U\in[n]^k \\ R-\text{multi}}} d^{(\kappa-1)m} \cdot \mathbb{P}[\mathcal{E}_H^M | \mathcal{E}_G^m] \left((c_3 R\ell)^{4d\kappa\ell} (d\ell)^2 n^{\kappa m + 2d\kappa + 1}\right) \end{aligned}$$

$$= \frac{(c_4 R\ell)^{4d\ell(\kappa+1)}}{d^{(\kappa-1)4d\ell}}(d\ell)^2 n^{2d\kappa+1} \cdot \sum_{M=1}^{k\ell} p^M \cdot \sum_{m \le M} d^{(\kappa-1)m} \cdot n^{\kappa m} \sum_{\substack{U \in [n]^k \\ R-\text{multi}}} \mathbb{P}[\mathcal{E}_H^M | \mathcal{E}_G^m] \,.$$

For a given hypergraph (V, G) and a fixed U, there are $\left(d! \cdot \left(\frac{(2d)!}{d!2^d}\right)\right)^{2\ell}$ hyperedge groupings from which we can sample (V, U, H)–first we choose a matching of the $2d$ hyperedges in each column, and then we choose an ordering on the d vertices of U to determine which belong to each hyperedge pair. We now appeal to the following lemma, which bounds the number of pairings and choices of U that can result in the evenness property for a given (V, G):

Lemma 3.4.18. *Suppose $R, w, h \in \mathbb{N}$ such that $h \le n$ and h is even. Let G be an even R-multilinear hypergraph with h hyperedges per column and w columns, and hyperedge multiplicity profile $\alpha = a_1, \ldots, a_t$. Let H be a hypgergraph we sample from G by matching edges in a column to each other, then adding a vertex in between with a label from the set $[n]$, with the additional constraint that the columns of center labels be R-multilinear, and that no two identical ordered hyperedges from G are matched to the same center vertex. Let τ be a valid number of distinct hyperedges sampleable from G. Then*

$$(\#\ H\ \textit{even with}\ \le \tau\ \textit{edges} \mid G) \le (\frac{h}{2}!)^w \cdot (2hw)^2 (4hR^2)^{hw} \cdot (hn)^{\tau/2} \,.$$

Applying Lemma 3.4.18 with $w \leftarrow 2\ell$, $h \leftarrow 2d$,$\tau \leftarrow M$, we have that since all of our U-configurations are R-multilinear, and since we forbid two identical ordered edges to share a center vertex, we sum over all possible hyperedge multiplicity profiles (at most $2^{4d\ell}$) we divide by the number of possible hyperedge groupings and get that for some constant c_5 depending on k,

$$\sum_{U \in \mathcal{U}_R} \sum_{\alpha} \mathbb{P}[\mathcal{E}_H^M \mid \mathcal{E}_G^{\le m}] \le 2^{8d\ell} \cdot \frac{(d!)^{2\ell}(8d\ell)^2(4\ell R^2)^{4d\ell} \cdot (2dn)^{M/2}}{\left(d! \cdot \frac{(2d)!}{2^d d!}\right)^{2\ell}} \le (c_5 R^2 \ell)^{4d\ell} \frac{(d\ell)^2 \cdot (dn)^{M/2}}{d^{2d\ell}}$$

And combining this with the above, there exists a constant c_6 depending on k so that

$$\begin{aligned}
&\mathbb{E}[\mathrm{Tr}(\Gamma\Gamma^\top)^\ell] \\
&\le \frac{(c_4 R\ell)^{4d\ell(\kappa+1)}}{d^{(\kappa-1)4d\ell}}(d\ell)^2 n^{2d\kappa+1} \cdot \sum_{M=1}^{2d\ell} p^M \cdot \sum_{m \le M} d^{(\kappa-1)m} \cdot n^{\kappa m} \left(\frac{(d\ell)^2 (c_5 \ell R^2)^{4d\ell} (dn)^{M/2}}{d^{2d\ell}}\right) \\
&\le \frac{(c_6 R\ell)^{4d\ell(\kappa+8)}}{d^{(\kappa-1)4d\ell}} \cdot \frac{1}{d^{2d\ell}}(d\ell)^4 \cdot n^{2d\kappa+1} \sum_{M=1}^{2d\ell} p^M (dn)^{M/2} \cdot \sum_{m \le M} d^{(\kappa-1)m} \cdot n^{\kappa m} \\
&\le \frac{(c_6 R\ell)^{4d\ell(\kappa+8)}}{d^{(\kappa-1/2)4d\ell}} \cdot (d\ell)^4 n^{2d\kappa+1} (2d\ell) \sum_{M=1}^{2d\ell} p^M \cdot d^{(\kappa-1/2)M} \cdot n^{(\kappa+1/2)M}
\end{aligned}$$

And so long as $pk^{(\kappa-1/2)}n^{(\kappa+1/2)} \geq 1$,

$$\leq \frac{(c_6 R\ell)^{4d\ell(\kappa+8)}}{d^{(\kappa-1/2)4d\ell}} \cdot n^{2d\kappa+1}(d\ell)^6 \cdot p^{2d\ell} \cdot d^{(\kappa-1/2)2d\ell} \cdot n^{(\kappa+1/2)2d\ell}$$
$$\leq n^{2d\kappa+1}(d\ell)^6 \cdot \left(c_6(R\ell)^{2(\kappa+8)} \cdot \frac{pn^{\kappa+1/2}}{d^{\kappa-1/2}}\right)^{2d\ell},$$

and now taking $\ell = O(\log n)$ and recalling that $R = 100 \log n$, with high probability by Proposition 3.2.4 we have that

$$\|\Gamma\|^{1/d} \leq \tilde{O}\left(\frac{pn^{k/2}}{d^{(k-2)/2}}\right).$$

□

This concludes our bound on $\|\Gamma\|$–in the next subsection, we prove the bounds on the sampling probabilities that we relied upon in the proofs of Theorem 3.4.4 and Theorem 3.4.13.

Bounding probabilities of sampling even hypergraphs

Our first proposition counts the number of vertex configurations with the evenness property and a given set of edge multiplicities for a fixed edge set E.

Proposition (Restatement of Proposition 3.4.11)**.** *Let $w, h, n \in \mathbb{N}$. Let $\alpha = \alpha_1, \ldots, \alpha_t$ be a sequence of t even numbers so that $\sum_{i=1}^{t} \alpha_i = w \cdot h$. Let $E = M_1, \ldots, M_w$ be a sequence of perfect matchings between two sets of size h.*

Let $\mathcal{G}_{w\times h}^{\alpha,E}$ be the set of all graphs which have a vertex set comprised of w R-multilinear multisets $S_1, \ldots, S_w \in [n]^h$, and have edges forming the perfect matching M_i between S_i, S_{i+1} (where the indexing is modulo w), so that the labels in $[n]$ assigned to the vertices induce exactly t distinct labelings for the edges, and the labeled edges have multiplicities $\alpha_1, \ldots, \alpha_t$. In words, $\mathcal{G}_{w\times h}^{\alpha,E}$ is the set of $w \times h$ matching cycles with matchings specified by E that have edge multiplicities α when labeled with R-multilinear labels from $[n]$.

If $w \cdot h \leq n$,

$$|\mathcal{G}_{w,h}^{\alpha,E}| \leq (5Rw)^{wh}(wh)^3 \cdot n^{t+h}.$$

We remark that this proposition resembles a lemma used in establishing exact bounds on the order of the deviation of the second eigenvalue of a Wigner matrix, in the work of [FK81]. Unfortunately, their statement does not directly imply the bounds we require, as they work in a slightly different setting and wished to precisely bound the constant. Our proof is similar to the exposition of [FK81] in [Tao].

Proof. We bound the number of such graphs by encoding each graph as a unique string. Since E is known, it suffices to encode enough information to recover the labels of the vertices.

We will call M_i (the matching in E between S_i and S_{i+1}) the ith *column* of edges. We choose an ordering on the edges of E: we order them first by column, and within each column

arbitrarily. Given a $G \in \mathcal{G}_{w\times h}^{\alpha,E}$, we will process the edges one at a time in this pre-specified order, and for each edge we will record either the labels of its incident vertices, or enough information to recover the labels from what we have previously recorded. To reduce the amount of recorded information, it will be helpful to specify several edge types:

- Edges we see for the first time:
  - *new-endpoint edges*: never-before seen edges e_i with $a_i = 2$ which take us to a vertex with a label we have not seen before. Let there be #new such edges.
  - *reused-endpoint edges*: never-before seen edges e_i with $a_i = 2$ which take us to a vertex with a label we have already seen.
- Edges we see for the second (and last) time:
  - *return edges*: edges e_i with $a_i = 2$ which we see for the second (and last) time.
  - *unforced edges*: return edges e_i with $a_i = 2$ which are not the only possible labeled edge we can use from the endpoint vertex of the previous edge.
- Edges we see more than twice:
  - *high-multiplicity edges*: edges e_i with $a_i > 2$.

Suppose there are #new new-endpoint edges, #reused reused-endpoint edges, #unforced forced edges, and #high high-multiplicity edges. As we process the edges in our pre-specified order, we record:

- The labels of each vertex belonging to the first column set S_1 (at most $n^{|S_1|} = n^h$ choices).
- The edge type of every edge in the graph: whether it is a new-endpoint edge, a reused-endpoint edge, a high-multiplicity edge, or a forced or unforced return edge (at most $5^{|E|} = 5^{wh}$ choices).
- For each new-endpoint edge, we record the label of its second endpoint (at most $n^{\#\text{new}}$ choices).
- For each reused-endpoint edge, we record the location of the first appearance of its second endpoint (at most $|V|^{\#\text{reused}} = (wh)^{\#\text{reused}}$ choices).
- For each unforced return edge, we record the column index of the first appearance of that edge, and the index within that column of the label involved in the edge (at most $(Rw)^{\#\text{unforced}}$ choices).
- For each high multiplicity edge e_i we record the second endpoint of its first appearance (at most $n^{\#\text{high}}$ choices). For each consequent appearance of e_i, we record the column index of the first appearance of e_i, and the index within that column of the label involved in the high multiplicity edge (in total, at most $(Rw)^{\sum_{a_i>2} a_i}$ choices).

Claim. The recorded information, along with E and α, suffices to reconstruct G.

Proof. We will prove this by induction. Our inductive claim is that in the ith step, we can reconstruct the labels of the ith column of vertices, S_i.

For the first column, we have recorded all of the labels, so we have S_1.

In any subsequent column, assuming we know S_i, we will process the edges in order, and determine their endpoint in S_{i+1}. For each edge, we can the determine from our edge information whether it is a new-endpoint, reused-endpoint, high-multiplicity, unforced return, or

forced return edge. Depending on the type of edge we use different information to discern the label of its endpoint in S_{i+1}:

- If we traversed a new-endpoint edge: we have recorded the label of the endpoint in S_{i+1}.
- If we traversed a reused-endpoint edge: we have recorded the position of the first appearance of the vertex's label, and we can look it up.
- If we traversed an unforced return edge: we have recorded the column index of the first appearance of the edge, as well as the index I within that column of the label corresponding to this edge's endpoint in S_i. We go to the column, and then choose the label in S_{i+1} by finding the Ith edge in that column with a label matching our edge's known label from S_i.
- If we traversed a forced return edge: there is only one choice for the second endpoint.
- If we traversed a high-multiplicity edge: we have recorded the column index of the other appearances of this edge. If this is the first appearance of the edge, we have recorded the label of the second endpoint. Otherwise, we have recorded the column in which this edge first appeared, as well as the index I within that column of the label corresponding to this edge. We go to the column, and then choose the label in S_{i+1} by finding the Ith edge in that column with a label matching our edge's known label from S_i.

This proves the inductive claim. □

Thus,

$$|\mathcal{G}^{\alpha,E}_{w\times h}| \le \sum_{\substack{\#\text{reused}\\ \#\text{new}\\ \#\text{unforced}}} 5^{wh}\cdot (wh)^{\#\text{reused}}\cdot (Rw)^{\sum_{a_i>2} a_i+\#\text{unforced}}\cdot n^{\#\text{new}+\#\text{high}+h}. \tag{3.4.16}$$

All that remains is for us to translate between the above quantity, which is in terms of edge types, to our desired quantity in terms of the parameters t, w, h. We do this by observing that

$$\#\text{new} = t - \#\text{high} - \#\text{reused}\,.$$

This is because there are a total of t distinct labeled edges, and of those, #high are high-multiplicity, and #reused do not introduce new labels.

We use this observation to simplify n's exponent, and we use the fact that $\sum_{\alpha_i>2} a_i + \#\text{unforced} + \#\text{reused} \le \sum_i \alpha_i = wh$ to simplify w and R's exponents, giving us that

$$|\hat{\mathcal{G}}^t_\alpha| \le (5Rw)^{wh}\cdot n^h \sum_{\substack{\#\text{reused}\\ \#\text{new}\\ \#\text{unforced}}} \left(\frac{wh}{n}\right)^{\#\text{reused}} \cdot n^t. \tag{3.4.17}$$

The number of possible combinations of values for #new, #unforced, and #reused is at most $(wh)^3$. Furthermore, because $wh \le n$, from (3.4.17) we may conclude,

$$|\mathcal{G}^{\alpha,E}_{w\times h}| \le (5Rw)^{wh}(wh)^3\cdot n^{t+h}$$

as desired. □

We now prove Lemma 3.4.10, which gives us a bound on the probability that we sample an even hypergraph cycle with the correct edge multiplicity from a simple edge cycle. Again, the proof of Lemma 3.4.10 is different from the proof of Lemma 3.2.8, and is more similar to the proof of Proposition 3.4.11 (although it is already quite different from the proof of [FK81]).

Lemma (Restatement of Lemma 3.4.10)**.** *Suppose $h, w, \kappa, n, \tau \in \mathbb{N}$. Let $G = (V, E)$ be a graph consisting of w sets of κh vertices each with R-multilinear labels from $[n]$, where E is a set of w perfect matchings $M_1, \dots, M_w$, so that M_i is a perfect matching between S_i and S_{i+1}, and $\alpha = a_1, \dots, a_t$ is a list of even edge multiplicities of E on the labeled vertex set V, so that $\sum a_i = \kappa w h$.*

Suppose we sample a hyperedge matching configuration H from E by uniformly grouping the edges in each matching from S_i to S_{i+1} into hyperedges of order 2κ, and let τ be a number of distinct hyperedges that is possible to sample from (V, E) in this way. Then,

$$\mathbb{P}((V, H) \text{ even with } \tau \text{ edges} \mid (V, E) \text{ even }) \leq \frac{(2e^\kappa \kappa^\kappa R^{\kappa+1} w)^{wh}}{(\kappa h)^{(\kappa-1)(wh-\tau)}}.$$

Proof. We will count the number of possible even H one can sample from E with at most τ unique hyperedges by encoding each such H uniquely as a string, then counting the number of strings.

First, we fix an ordering on the edges of E, ordering them first by column, then arbitrarily within each column. Now, define a *new* hyperedge to be a labeled hyperedge which we have never seen before, and define an *old* hyperedge to be a hyperedge which has already been seen. Let H be some even hypergraph sampled from G with at most M unique labeled hyperedges. We encode H in a string as follows. We will process the hyperedges of H one at a time, ordering hyperedges by the first simple edge they contain.

- For every hyperedge encountered in H, record whether it is new or old ($2^{|H|} = 2^{wh}$ options).
- For every new hyperedge encountered: record the indices of the $2, ..., d$ simple edges that it contains $\left((dh)^{(d-1)}\right)^\tau$ options). The identity of the first edge will be obvious from the edge ordering.
- For every old hyperedge encountered, record the column index j of its first appearance ($w^{wh-\tau}$ choices). If any of the simple edges $e_{i_1}, \dots, e_{i_d}$ appear with multiplicity > 1 in the old (jth) column, record the indices of those edge within the identical edges in that column (at most $(R^d)^{wh-\tau}$ choices, since no simple edge can appear more than R times in a column).

We claim that given V, E, and this encoding, we can uniquely recover H. We process the simple edges in our specified order. If the edge is contained in a new hyperedge, we can deduce the other simple edges belonging with it from what we recorded. If the edge is contained in an old hyperedge, we know the column in which the hyperedge first appears, and we can determine the other edges in the group by looking up the edge in the previous column–if there are multiple copies of the edge in the old column, we have recorded which

copy to look up. Furthermore, if the grouping is insufficient due to multiplicities within this column, we have recorded the relative indices of the relevant edges.

The number of such strings is at most

$$(\#\text{ encodings}) \leq 2^{wh} \cdot (dh)^{(d-1)\tau} \cdot (R^d w)^{wh-\tau} \tag{3.4.18}$$

giving an upper bound on the number of H we can sample from (V, E) with at most τ distinct edges. There are a total of

$$(\#\text{ sampleable graphs}) = \left(\frac{1}{h!} \prod_{j=0}^{h-1} \binom{d(h-j)}{d} \right)^w = \left(\frac{(dh)!}{(d!)^h h!} \right)^w \tag{3.4.19}$$

possible hyperedge graphs sampleable from (V, E), and from this we have that

$$\mathbb{P}(H \text{ has } \tau \text{ edges } | (V, E) \text{ even }) \leq \frac{(\#\text{ encodings})}{(\#\text{ sampleable graphs})} \leq \frac{(2d^d e^d w R^d)^{hw}}{(dh)^{(d-1)(hw-\tau)}}$$

where we have combined (3.4.18) with (3.4.19) and applied Stirling's inequality. Our conclusion follows. □

Now we prove a lemma that bounds the number of even k-hyperedge configurations with τ hyperedges, each paired and sharing a center vertex, sampleable from an even $2d$-hyperedge configuration by pairing and labeling.

Lemma (Restatement of Lemma 3.4.18)**.** *Suppose $R, w, h \in \mathbb{N}$ such that $h \leq n$ and h is even. Let G be an even R-multilinear hypergraph with h hyperedges per column and w columns, and hyperedge multiplicity profile $\alpha = a_1, \ldots, a_t$. Let H be a hypgergraph we sample from G by matching edges in a column to each other, then adding a vertex in between with a label from the set $[n]$, with the additional constraint that the columns of center labels be R-multilinear, and that no two identical ordered hyperedges from G are matched to the same center vertex. Let τ be a valid number of distinct hyperedges sampleable from G. Then*

$$(\#\ H \text{ even with } \leq \tau \text{ edges } |\ G) \leq (\frac{h}{2}!)^w \cdot (2hw)^2 (4hR^2)^{hw} \cdot (hn)^{\tau/2}\,.$$

Proof. We will count the number of such H by encoding each instance uniquely as a string. Fix an order on the hyperedges, first in column order.

- A *new* hyperedge is a hyperedge which introduces a new center vertex.
- A *reuse* hyperedge is a hyperedge which we see for the first time, and whose center vertex is the first of its type in its column, but which reuses a center vertex from a previous column.
- A *sharing* hyperedge is a hyperedge which we see for the first time, but which shares a center vertex with another hyperedge in its own column.
- A *return* hyperedge is a hyperedge which we see for the second or later time.

Our encoding is as follows:

- In the first hw positions, we record the type of every hyperedge we see (4^{hw} choices).
- In the next #new positions, we record the labels of new hyperedges ($n^{\#\text{new}}$ choices).
- In the next #reused positions, we record the position of the first appearance of the reused label ($(hw/2)^{\#\text{reused}}$ choices)
- In the next #share positions, we record the partner of the sharing edge within the column ($(h/2)^{\#\text{share}}$ choices, since there cannot be more than h new labels in a column).
- In the next #return positions, we record the column of the first appearance of the hyperedge (a total of $w^{\#\text{return}}$ choices), the index of the previous occurrence of the hyperedge among hyperedges with the same labels in the previous column (a total of $R^{\#\text{return}}$ choices), and the index of the hyperedge's center vertex among center vertices of the same label within the current column (a total of $R^{\#\text{return}}$ choices).
- We record the permutation of the middle labels in each column ($(\frac{h}{2}!)^w$ choices).

Given this information, we can uniquely reconstruct an H from G. For every surprise hyperedge we encounter, we have recorded the center label. For every recycle hyperedge we encounter, we can determine the center label by looking at the previous occurrence. For every sharing hyperedge, we can determine its partner. For every return hyperedge, we can determine the label of the center vertex by looking at the previous occurrence, and we can determine partnership by knowing the index of the center label's occurrence within the column.

We thus have

$$(\#H) \leq \sum_{\substack{\#\text{new} \\ \#\text{reused}}} \left(\frac{h}{2}!\right)^w (4w)^{wh} \cdot R^{2\#\text{return}} \cdot n^{\#\text{new}} \cdot h^{\#\text{reused}+\#\text{share}}\,.$$

We use some observations about these quantities to simplify the above expression. We have that

$$\tau = \#\text{new} + \#\text{share} + \#\text{reused}\,,$$

since every hyperedge must appear for the first time. Furthermore, we have that #new $\leq$ #share, since every surprising hyperedge must be paired with a sharing hyperedge, since it introduces a new vertex label which hasn't been seen before, so it's partner must be a sharing edge since we forbid two hyperedges with the same labels to share a center vertex. It follows that

$$\#\text{new} \leq \tau/2\,.$$

Putting these together,

$$\begin{aligned}
(\#H) &\leq \left(\frac{h}{2}!\right)^w (4wR^2)^{hw} \sum_{\substack{\#\text{new} \\ \#\text{reused}}} n^{\#\text{new}} \cdot h^{\tau-\#\text{new}} \\
&= \left(\frac{h}{2}!\right)^w (4wR^2)^{hw} h^\tau \cdot \sum_{\substack{\#\text{new} \\ \#\text{reused}}} \left(\frac{n}{h}\right)^{\#\text{new}} \\
&\leq \left(\frac{h}{2}!\right)^w (2hw)^2 (4hR^2)^{hw} \cdot (hn)^{\tau/2}\,,
\end{aligned}$$

where in the last line we have assumed that $h < n$, and have used that #new and #reused take on at most $\tau < wh/2$ values. The conclusion follows. □

3.5 Strong Refutation for All CSPs

In this section, we consider the problem of refuting Boolean CSP's with arbitrary predicates.

Problem 3.5.1 (Refuting CSP's with predicate P)**.** Let $P : \{\pm 1\}^k \to \{0, 1\}$ be a predicate on k variables. Then we sample a *random instance of CSP-P*, Φ, with clauses $C_1, \ldots, C_m$, as follows: for each $I \in [n]^k$:

- With probability p, sample a uniformly random $\sigma \in \{\pm 1\}^k$ and add the constraint $P(x_I \oplus \sigma) = 1$ to Φ as clause C_I, where $\oplus$ denotes the entry-wise product and x_I denotes the ordered subset of variables x_i for each $i \in I$.

The problem of *strongly refuting CSP-P* is to devise an algorithm that given an instance Φ sampled as above, with high probability over Φ, outputs a certificate that

$$\mathrm{opt}(\Phi) \leq 1 - \gamma$$

for an absolute constant $\gamma > 0$.

Our result is the following:

Theorem 3.5.2. *Let Φ be a random instance of a k-CSP with predicate P, with clause density $m/n \geq \tilde{O}(n^{(k/2-1)(1-\delta)})$. Then with high probability over the choice of Φ, there is a spectral algorithm which strongly refutes Φ in time $\exp(\tilde{O}(n^\delta))$, certifying that*

$$\mathrm{opt}(\Phi) \leq \mathop{\mathbb{E}}_{x \sim \{\pm 1\}^k}[P(x)] + \varepsilon$$

for any constant $\varepsilon > 0$. Furthermore, the degree-$O(n^\delta)$ SoS relaxation also certifies this bound.

We will employ the framework of Allen et al. [AOW15] to prove that we can strongly refute any k-CSP with predicate P at densities as low as $\tilde{O}(1)$, given sufficient time. The strategy is as follows: given some random k-CSP on $x \in \{\pm 1\}^n$ with clauses $C_1, \ldots, C_m$, $C_i : \{\pm 1\}^k \to \{\pm 1\}$:

- Expand $C_1(x), \ldots, C_m(x)$ using the Fourier expansion.
- Split the Fourier expansions of $C_1, \ldots, C_m$ into XOR instances.
- Refute each XOR instance.

Because a k-CSP predicate has a Fourier expansion of degree at most k, the above strategy in combination with our k-XOR refutation results will allow us to tightly strongly refute any k-CSP in time $\exp(n^\delta)$ at densities $\geq \tilde{O}(n^{(k/2-1)(1-\delta)})$. However, as a result of the work of [AOW15], we are able to show that for predicates satisfying some additional properties, it is possible to strongly refute more quickly at lower densities. We elaborate further:

Definition 3.5.3. Let $1 \le t \le k$. A predicate $P : \{\pm 1\}^k \to \{0,1\}$ is δ-far from t-wise supporting if every distribution $\mathcal{D}$ on $\{\pm 1\}^k$ which has uniform marginals on all subsets of t variables only satisfies P with probability at most $1-\delta$, i.e.

$$\mathop{\mathbb{E}}_{x\sim\mathcal{D}}[P(x)] \le 1-\delta.$$

Allen et al. give the following characterization of δ-far from t-wise supporting predicates:

Theorem 3.5.4 (Lemma 3.16 and Theorems 4.9 and 6.6 of [AOW15])**.** *Let the predicate $P : \{\pm 1\}^k \to \{0,1\}$ be δ-far from t-wise supporting, for $0 \le t \le k$. Then there exists a multilinear polynomial $Q : \{\pm 1\}^t \to \mathbb{R}$ such that $\mathbb{E}_{x\sim\{\pm\}^t}[Q(x)] = 0$ and $P(x) \le (1-\delta)+Q(x)$ for any $x \in \{\pm 1\}^k$. Furthermore, Q can be obtained by solving a linear program of constant size, and there is a degree-k SoS proof that $\tilde{\mathbb{E}}[P(x)] \preceq (1-\delta) + \tilde{\mathbb{E}}[Q(x)]$.*

We will use this theorem to extend the results of [AOW15] for δ-far from t-wise independent predicates below the spectral threshold.

Theorem 3.5.5. *Let the predicate $P : \{\pm 1\}^k \to \{0,1\}$ be δ-far from t-wise supporting, for $0 \le t \le k$ and a constant $\delta > 0$. Let Φ be a random instance of a k-CSP with predicate P, with clause density $m/n \ge \tilde{O}(n^{(t/2-1)(1-\delta)})$. Then with high probability over the choice of Φ, there is a spectral algorithm which strongly refutes Φ in time $\exp(\tilde{O}(n^\delta))$, certifying that*

$$\mathrm{opt}(\Phi) \le 1-\delta+\varepsilon$$

for any constant $\varepsilon > 0$. Furthermore, the degree-$O(n^\delta)$ SoS relaxation also certifies this bound.

We now prove Theorem 3.5.2, and then below we will describe the mild changes needed to prove Theorem 3.5.5. We will utilize our own Theorem 3.1.3, as well as the following theorem which has appeared in [AOW15] (and also partially in [BM16]). We cite the exact form of the theorem given in [AOW15].

Theorem 3.5.6 ([AOW15], Theorem 4.1)**.** *For $k \ge 2$, $q \ge n^{-k/2}$, let $\{w_S\}_{S\in[n]^k}$ be independent random variables such that for each $S \in [n]^k$,*

$$\mathbb{E}[w_S] = 0, \qquad \mathbb{P}[w_S \ne 0] \le q, \qquad \text{and} \qquad |w_S| \le 1.$$

Then there is an efficient algorithm certifying that

$$\left|\sum_{S\in[n]^k} w_S \prod_{i\in S} x_i\right| \le 2^{O(k)}\sqrt{q}n^{3k/4}\log^{3/2} n$$

for all x with $\|x\|_\infty \le 1$ with high probability.

Remark 3.5.7. In [AOW15], the theorem appears without the absolute value–however the statement for the absolute value is implied by the fact that the negated variables $-w_S$ also satisfy all of the constraints.

Given the above theorem and our results for refuting XOR instances (Theorem 3.1.3), the result for arbitrary binary CSPs follows easily.

Proof of Theorem 3.5.2. Let the Fourier expansion of P on $y \in \{\pm 1\}^k$ be $P(y) = \sum_{S \subseteq [k]} \hat{P}(S) \cdot \chi_S(y)$. Let Φ have constraints $C_1, \ldots, C_m$, chosen independently on each $I \in [n]^k$ with probability p, so that the constraint C_I asserts that $P(x_I \oplus \sigma^I) = 1$ for a uniformly chosen signing $\sigma^I \in \{\pm 1\}^k$. We have that

$$P_\Phi(x) = \sum_{I \in [n]^k} \mathbb{I}(C_I \in \Phi) \cdot P(x_I \oplus \sigma^I) = \sum_{I \in [n]^k} \mathbb{I}(C_I \in \Phi) \cdot \sum_{S \subseteq [k]} \hat{P}(S) \cdot \prod_{i \in I} x_i \cdot \prod_{i \in S(I)} \sigma_i^I \,,$$

where we have used $\oplus$ to denote the entry-wise product and $S(I)$ to denote the entries of I corresponding to the subset S of $[k]$. We will move the sum over ordered subsets $S \subseteq k$ outwards, then simplify further

$$\begin{aligned} &= \sum_{S \subseteq k} \sum_{I \in [n]^k} \mathbb{I}(C_I \in \Phi) \cdot \hat{P}(S) \cdot \prod_{i \in S(I)} x_i \cdot \prod_{i \in S(I)} \sigma_i^I \\ &= \hat{P}(\emptyset) + \sum_{\substack{S \subseteq k \\ |S| \geq 1}} \hat{P}(S) \sum_{I \in [n]^k} \mathbb{I}(C_I \in \Phi) \cdot \prod_{i \in S(I)} x_i \cdot \prod_{i \in S(I)} \sigma_i^I \,. \end{aligned}$$

Now, we will see that for each $S \subseteq [k]$, $|S| \geq 1$, we have a random weighted XOR instance Ψ_S on $|S|$ variables. Letting $b_L^I \stackrel{\text{def}}{=} \prod_{\ell \in L} \sigma_\ell^I$, we define

$$\begin{aligned} \Psi_S(x) &\stackrel{\text{def}}{=} \sum_{I \in [n]^k} \mathbb{I}(C_I \in \Phi) \cdot \prod_{i \in S(I)} x_i \cdot \prod_{i \in S(I)} \sigma_i^I \\ &= \sum_{L \in [n]^S} \prod_{i \in L} x_i \cdot \left(\sum_{J \in [n]^{k \setminus S}} \mathbb{I}(C_{J \cup L} \in \Phi) \cdot b_L^I \right) , \end{aligned}$$

where we have abused notation by allowing $J \cup L$ to denote the ordered multiset with L in the exact positions corresponding to S and J in the positions corresponding to $k \setminus S$. Furthermore, by definition,

$$\frac{P_\Phi(x)}{m} = \hat{P}(\emptyset) + \sum_{\substack{S \subseteq k \\ |S| \geq 1}} \hat{P}(S) \cdot \frac{\Psi_S(x)}{m} \leq \hat{P}(\emptyset) + \sum_{\substack{S \subseteq k \\ |S| \geq 1}} \hat{P}(S) \cdot \frac{|\Psi_S(x)|}{m}$$

and so bounding the values of the Ψ_S suffices to get a bound on the value of Φ.

We list some properties of Ψ_S. For convenience, we now use the notation $\chi_L(x) \stackrel{\text{def}}{=} \prod_{\ell \in L} x_\ell$ and the notation x_L to denote a string of elements of x indexed by L. For each $S \subseteq k, S \neq \emptyset$, Ψ_S is an instance of $|S|$-XOR with independent constraints on each $\chi_L(x)$ for $L \in [n]^{|S|}$–the independence is because the constraints for $\chi_L(x)$ depend only on the presence of clauses in $C_I \in \Phi$ for $I \in [n]^k$ including x_L in the positions corresponding to S. The weight on

each $\chi_L(x)$ is distributed according to a sum of $n^{k-|S|}$ independent random variables, each of which is 0 with probability $1-p$, and uniformly ± 1 with probability p. For convenience, call the distribution over such sums $\hat{k}(n^{k-|S|}, p)$, so that the coefficient c_L of $\chi_L(x)$ is distributed according to $c_L \sim \hat{k}(n^{k-|S|}, p)$.

We now perform a case analysis on p and $|S|$, which allows us to bound the contributions of each $\Psi_S(x)$ individually.

Case 1: $pn^{k-|S|} \geq 1$**.** If $pn^{k-|S|} \geq 1$ then with high probability, every constraint c_L has $|c_L| \leq O(\sqrt{pn^{k-|S|} \log n})$ (where we have combined Lemma 3.5.8 with a union bound). Furthermore the c_L are distributed symmetrically about 0, and they are nonzero with probability at most $1 \leq pn^{k-|S|}$.

- If $|S| = 1$, we have that with high probability,

$$\frac{|\Phi(x)|}{m} \leq \frac{n \cdot \max_{\ell \in [n]} |c_\ell|}{m} \leq \frac{n \cdot O(\sqrt{pn^{k-1} \log n})}{\Theta(pn^k)} \leq O\left(\sqrt{\frac{\log n}{pn^{k-1}}}\right),$$

  where we have applied a Chernoff bound to use that $m = \Theta(pn^k)$. By our assumption on the clause density, $p \geq n^{-(k-1)} \cdot \operatorname{polylog} n$, and therefore it follows that $|\Psi_S(x)|/m = o(1)$.

- Otherwise, if $|S| \geq 2$, we can divide each c_L by $\beta = O(\sqrt{pn^{k-|S|} \log n})$ to obtain a polynomial with coefficients bounded in absolute value by 1, with independent symmetrically distributed coefficients and probability at most $1 \leq pn^{k-|S|}$ of being nonzero. We can thus apply Theorem 3.5.6 to get that with high probability we can certify in polynomial time that,

$$\frac{|\Psi_S(x)|}{\beta} \leq O(n^{3|S|/4} \log^{3/2} n),$$

  Implying that with high probability,

$$\begin{aligned}\frac{|\Psi_S(x)|}{m} \leq \frac{\beta \cdot O(n^{3|S|/4} \log^{3/2} n)}{\Theta(pn^k)} &\leq \frac{O(\sqrt{pn^{k-|S|} \log n}) \cdot O(n^{3|S|/4} \log^{3/2} n)}{\Theta(pn^k)} \\ &\leq O\left(\frac{\log^2 n}{p^{1/2} n^{k/2-|S|/4}}\right) \leq O\left(\frac{\log^2 n}{n^{|S|/4}}\right).\end{aligned}$$

  where the last inequality follows by the assumption that $pn^{k-|S|} \geq 1$.

Case 2: $pn^{k-|S|} < 1$**.** If $pn^{k-|S|} < 1$, then with high probability all $|c_L| \leq O(\log n)$ (where we have combined Lemma 3.5.8 with a union bound). We now split into cases in which we can apply Theorem 3.5.6 and cases in which we must apply Theorem 3.1.3.

- If $|S| < k$ and $p \geq n^{-|S|/2}$, then again letting $\beta = O(\log n)$, we can divide Ψ_S by β to obtain a polynomial with coefficients that are symmetrically distributed about 0, bounded by 1 in absolute value, and are nonzero with probability at most $pn^{k-|S|}$. By Theorem

3.5.6 and by a Chernoff bound on m, it follows that we can certify in polynomial time that

$$\frac{|\Psi_S(x)|}{m} \leq \frac{\beta \cdot O(\sqrt{pn^{k-|S|}} \cdot n^{3|S|/4} \log^{3/2} n)}{m} \leq \frac{O(\log n) \cdot O(\sqrt{pn^{k-|S|}} \cdot n^{3|S|/4} \log^{3/2} n)}{\Theta(pn^k)}$$
$$\leq O\left(\frac{\log^{3/2} n}{p^{1/2} n^{k/2-|S|/4}}\right)$$
$$\leq O\left(\frac{\log^{3/2} n}{n^{(k-|S|)/2}}\right) = o(1),$$

where the second-to-last inequality follows by our assumption that $pn^{|S|/2} \geq 1$.

- If $|S| = k$, then Ψ_S is a k-XOR instance with each constraint present with probability p. By Theorem 3.1.3, we can certify in time $\exp(\tilde{O}(n^\delta))$ that $\frac{|\Psi_S|}{m} \leq \gamma$ for any constant $\gamma > 0$.
- If $|S| < k$ and $p < n^{-|S|/2}$, we must apply Theorem 3.1.3. We must modify the instances slightly first, since Theorem 3.1.3 applies to unweighted instances.
  To obtain an unweighted instance, we split Ψ_S further into $r = \log^2 n$ instances, $\Psi_S^{(1)}, \ldots, \Psi_S^{(r)}$. We split as follows: let $c_L^{(i)}$ denote the coefficient of $\chi_L(x)$ in $\Psi_S^{(i)}$. For each nonzero c_L, we choose $|c_L|$ uniformly random indices $i_1, \ldots, i_{|c_L|} \in [r]$, and assign $c_L^{(i_j)} = \frac{c_L}{|c_L|}$ for each $j = 1, \ldots, |c_L|$ (recall that with high probability $|c_L| \leq O(\log n) < r$).
  Let m_i be the number of constraints in $\Phi_S^{(i)}$, so that we have $m \geq \sum_i m_i$ (since c_L may be a sum of negative and positive constraints from the full instance Φ).
  First, note that

$$\frac{|\Psi_S(x)|}{m} \leq \frac{\left|\sum_{i=1}^r \Psi_S^{(i)}(x)\right|}{\sum_i m_i} \leq \max_{i \in [r]} \frac{|\Psi_S^{(i)}(x)|}{m_i}.$$

  It remains to argue that each instance $\Psi_S^{(i)}$ has bounded value with high probability. Towards this, consider the properties of $\Psi_S^{(i)}$. First, we note that the constraints of $\Psi_S^{(i)}$ are independent of one another and are distributed symmetrically about zero–this is because the c_L are independent of one another and symmetrically distributed about zero. Furthermore, we have that each $c_L^{(i)}$ is nonzero with probability $\hat{q}$:

$$\hat{q} \stackrel{\text{def}}{=} \mathbb{P}[c_L^{(i)} \neq 0] = \sum_{j=1}^r \mathbb{P}[|c_L| = j,\ i \text{ chosen}]$$
$$= \sum_{j=1}^r \mathbb{P}[i \text{ chosen} \mid |c_L| = j] \cdot \mathbb{P}[|c_L| = j]$$
$$= \sum_{j=1}^r \frac{j}{r} \cdot \mathbb{P}[|c_L| = j],$$

where we have taken the sum up to r because we implicitly condition on $|c_L| \leq O(\log n)$ (which occurs with high probability). We thus have that

$$\hat{q} = \sum_{j=1}^{r} \frac{j}{r} \cdot \mathbb{P}[|c_L| = j] \leq \mathbb{P}[|c_L| > 0] \leq pn^{k-|S|},$$

and also that

$$\hat{q} = \sum_{j=1}^{r} \frac{j}{r} \cdot \mathbb{P}[|c_L| = j] \geq \frac{1}{r} \cdot \mathbb{P}[|c_L| > 0] \geq \frac{1}{\log^2 n} \cdot \frac{pn^{k-|S|}}{2}.$$

The last inequality follows by observing that with probability at least $pn^{k-|S|}$, at least one of the chosen constraints in Φ contributes to c_L, and conditioned on this event, the contributions to c_L sum to zero with probability at most $1/2$. Therefore, $\hat{q} = \delta \cdot pn^{k-|S|}$ for some $\delta \in [\frac{1}{2\log^2 n}, 1]$.

Thus, each $\Phi_S^{(i)\prime}$ is a random $|S|$-XOR instance in which each clause is revealed with probability $\hat{q}$.

From Theorem 3.1.3, with high probability we can refute a random $|S|$-XOR instance in which each clause is present with probability $\hat{q}$ in time $\exp(\tilde{O}(n^\delta))$ so long as $\hat{q}n^{|S|-1} \geq \tilde{O}(n^{(|S|/2-1)(1-\delta)})$, certifying that the instance satisfies at most $\frac{1}{2} + \gamma + o(1)$ clauses for any constant $\gamma > 0$. This condition on $\hat{q}n^{|S|-1}$ holds by our assumption that $p \geq \tilde{O}(n^{-k(1+\delta)/2+\delta})$ (as long as we make the correct adjustments of logarithmic factors on p to account for the value of $\hat{q}$). We can also certify with high probability that the fraction of satisfied constraints is at least $\frac{1}{2} - \gamma$ for any constant γ, by applying the same argument with the negations of the $c_L^{(i)}$.

Thus, in time $\exp(\tilde{O}(n^\delta))$, with high probability we can certify that $\frac{|\Phi_S^{(i)}(x)|}{m_j} \leq \gamma$, implying by a union bound over $i \in [r]$ that $\frac{|\Psi_S(x)|}{m} \leq \gamma$.

Using this case analysis, we have that

$$\frac{P_\Phi(x)}{m} \leq \hat{P}(\emptyset) + \gamma \cdot \sum_{\substack{S \subseteq k \\ |S| \geq 1}} \hat{P}(S)$$

and since $\hat{P}(S)$ can depend only on k, for large enough n, with high probability over Φ we can certify that $\frac{P_\Phi(x)}{m} \leq \hat{P}(\emptyset) + \gamma'$ for any constant γ' in time $\exp(\tilde{O}(n^\delta))$ when $m/n \geq \tilde{O}(n^{(k/2-1)(1-\delta)})$.

The same conclusion holds in the degree-$O(n^\delta)$ SoS relaxation, as every step of this proof holds within the SoS proof system (because Theorem 3.5.6 and Theorem 3.1.3 hold within the SoS proof system). □

The proof of Theorem 3.5.5 proceeds almost identically, except that instead of using the Fourier expansion of the predicate P, we use the degree-t polynomial $Q(x)$ given by the work

of Allen et al. [AOW15] (Theorem 3.5.4). Because $P(x) \leq (1-\delta) + Q(x)$, and since $Q(x)$ has no constant term, the proof we applied to the degree ≥ 1 terms of the Fourier expansion of P applies to $Q(x)$, and this completes the proof.

Lemma 3.5.8. *For any $0 \leq q \leq 1$, define $\hat{k}(N,q)$ to be the distribution over scalars such that $X \sim \hat{k}(N,q)$ is a sum of N independent variables, each 0 with probability $1-q$, -1 with probability $q/2$, and 1 with probability $q/2$. Then if $Nq \geq 1$, for any constant c, there exists a constant c' such that*

$$\mathbb{P}_{X \sim \hat{k}(N,q)}[|X| \leq \sqrt{c' Nq \log N}] \geq 1 - N^{-c},$$

and if $Nq < 1$, for any constant c there exists a constant c' such that

$$\mathbb{P}_{X \sim \hat{k}(N,q)}[|X| \leq c' \log N] \geq 1 - N^{-c}.$$

Proof. By definition, for $X \sim \hat{k}(N,q)$, $X = \sum_{i=1}^{N} x_i$ for x_i distributed according to $\hat{k}(1,q)$.

It is easy to see that $\mathbb{E}[X] = 0$, and to calculate that $\mathbb{V}(X) = qN$, and we note also that $|x_i| \leq 1$. Therefore when $Nq \geq 1$, from a Bernstein inequality we have that

$$\mathbb{P}[|X| \geq t] \leq \exp\left(\frac{-t^2/2}{t/3 + qN}\right),$$

and taking $t = \sqrt{4cqN \log N}$, and using that $qN \geq 1$, we have the desired result.

When $qN < 1$, we apply the same bound with $t = 4c \log N$ to obtain our second result. □

3.6 Sum-of-Squares Algorithms

In this section, we use our spectral algorithms to certify SoS upper bounds.

Relaxations for tensor norm and k-XOR

The natural SoS relaxations for computing the tensor norm and for maximizing k-CSPs are very similar to each other. Both correspond to polynomial maximization problems, where the constraint is that the maximizing solution $x \in \mathbb{R}^n$ lie on the unit sphere or on the Boolean hypercube. Save for these "normalization" constraints and the natural SDP constrains SoS_d, there are no other constraints.

Definition 3.6.1 (d-round SoS relaxation for tensor norm)**.** Given an order-k tensor $\mathbf{T}$, for any $d \geq \lceil k/2 \rceil$, the d-round SoS relaxation for the injective tensor norm is

$$\begin{aligned} &\max \ \tilde{\mathbb{E}}\left[\langle \mathbf{T}, x^{\otimes k} \rangle\right] \\ &s.t. \quad \tilde{\mathbb{E}}\left[\sum_{i \in [n]} x_i^2\right] = \sum_{i \in [n]} X_{i,i} = 1, \end{aligned} \tag{3.6.1}$$

With the addition of the standard d-round SoS constraints.

Definition 3.6.2 (d-round SoS relaxation for k-XOR)**.** Given an instance Φ of d-XOR with constraint tensor $\mathbf{T}_\Phi$ defined as described in Section 3.4, for any $d \geq \lceil k/2 \rceil$, the d-round SoS relaxation is given by

$$\begin{aligned} \max \ & \tilde{\mathbb{E}} \left[\langle \mathbf{T}_\Phi, x^{\otimes k} \rangle \right] \\ s.t. \quad & \tilde{\mathbb{E}} \left[x_i^2 \right] = 1 \qquad \forall i \in [n] \,, \end{aligned} \tag{3.6.2}$$

With the addition of the standard d-round SoS constraints.

Bounds for Tensor Norm

In this subsection, we show how bound the objective value of the d-round SoS relaxation for a polynomial optimization problem when $\tilde{\mathbb{E}}(\sum_i x_i^2)$ is known, in terms of the operator norm of a specific matrix. We will use $\succeq$ and $\preceq$ to denote inequalities that are sum-of-squares identities.

Proposition 3.6.3. *Let $\mathbf{T}$ be an order-k tensor for even $k = 2\kappa$. Let $R, d \in \mathbb{N}$ such that d is a power of 2, $R \leq dk$ and k is even. Consider the dk-round SoS relaxation for the problem $\mathcal{P}$,*

$$\max \ \tilde{\mathbb{E}} \left[\langle \mathbf{T}, x^{\otimes k} \rangle \right] \qquad s.t. \quad \tilde{\mathbb{E}} \left[\|x\|_2^2 \right] = \alpha \,, \tag{3.6.3}$$

Furthermore, let T be the natural flattening of $\mathbf{T}$ to an $n^{k/2} \times n^{k/2}$ tensor, let $\hat{\mathcal{S}}_{dk/2}$ be the set of matrices that permute rows and columns of matrices in $[n]^{dk/2} \times [n]^{dk/2}$ according to actions of $\mathcal{S}_{dk}$ on the coordinates in $[n]$ Then in the dk-round SoS relaxation,

$$\tilde{\mathbb{E}} \left[\left\langle \mathbf{T}, x^{\otimes k} \right\rangle \right] \leq \alpha^{k/2} \cdot \left\| \mathop{\mathbb{E}}_{\Pi, \Sigma \in \hat{\mathcal{S}}_{dk/2}} \left[\Pi T^{\otimes d} \Sigma \right] \right\|^{1/d} .$$

Proof. We have that

$$\begin{aligned} \left(\tilde{\mathbb{E}} \left[\langle \mathbf{T}, x^{\otimes k} \rangle \right] \right)^d &\leq \tilde{\mathbb{E}} \left[\left(\langle \mathbf{T}, x^{\otimes k} \rangle \right)^d \right] && \text{(by Fact 2.5.3)} \\ &= \tilde{\mathbb{E}} \left[\langle \mathbf{T}^{\otimes d}, x^{\otimes dk} \rangle \right] && \text{(by the symmetry constraints (2.3.2))} \\ &= \tilde{\mathbb{E}} \left[\left\langle \left(\mathop{\mathbb{E}}_{\Pi, \Sigma \in \hat{\mathcal{S}}_{dk}} \left[\Pi(\mathbf{T}^{\otimes k}) \Sigma \right] \right), x^{\otimes dk} \right\rangle \right] && \text{(by (2.3.2))} \\ &\leq \left\| \mathop{\mathbb{E}}_{\Pi, \Sigma \in \hat{\mathcal{S}}_{dk/2}} \left[\Pi(\mathbf{T}^{\otimes k}) \Sigma \right] \right\| \cdot \tilde{\mathbb{E}} \left[\|x\|^{dk} \right] && \text{(by Lemma 2.5.1)} \end{aligned}$$

The conclusion follows from (3.6.3). □

As an immediate corollary of the above and of Theorem 3.3.3, we have Theorem 3.1.7 for even k. To get Theorem 3.1.7 for odd k, we can apply Cauchy-Schwarz before applying Fact 2.5.3, so that we are working with

$$\begin{aligned}
&\langle \mathbf{T}, x^{\otimes k}\rangle^2 \\
&\preceq \left(\sum_{i\in[n]} x_i^2\right)\cdot\left\langle \sum_{i\in[n]} T_i\otimes T_i, x^{\otimes 2(k-1)}\right\rangle \\
&= \left(\sum_{i\in[n]} x_i^2\right)\cdot\left(\left\langle \sum_{i\in[n]} T_i\otimes T_i - \text{squares}(T_i\otimes T_i), x^{\otimes 2(k-1)}\right\rangle + \sum_{i\in[n]}\sum_{A[n]^{k-1}} T_{A,i}^2 \prod_{j\in A} x_j^2\right),
\end{aligned}$$

where T_i is the ith slice of $\mathbf{T}$, and squares$(T_i\otimes T_i)$ corresponds to the entries of $T_i\otimes T_i$ which are squares of the base variables $\mathbf{T}$. The right-hand term is bounded by obtaining a high-probability bound of $O(\log n)$ on the maximum coefficient T_{iA}^2, and the left-hand term is bounded by following the same steps as in the proof of Proposition 3.6.3, then applying Theorem 3.3.11.

Bounds for k-XOR

For the case of k-XOR, the proof is a bit more complicated than for the case of tensor norms, because the matrix certificates we used have certain rows and columns deleted. Still, the arguments are similar to our proof from Section 3.4. All steps in the proofs from Section 3.4 and Section 3.4 we can make into SoS proofs in an analogous way to the tensor norm SoS proofs above, except for the steps in which the high-multiplicity rows and columns are deleted. This too is not difficult to see, and we will prove it for the even case. We will require the following SoS fact:

Claim 3.6.4. Suppose $\tilde{\mathbb{E}}$ is a degree $2d$ pseudoexpectation functional with Boolean constraints, i.e., $\tilde{\mathbb{E}}[x_i^2\cdot r(x)] = \tilde{\mathbb{E}}[r(x)]$ for all $r(x)$ with $\deg(r)\le 2d-2$.

Let $q=\sum_\sigma \hat{q}_\sigma x_\sigma$ and r be polynomials such that $\deg(qr^2)\le 2d$. Then,

$$\tilde{\mathbb{E}}[q(x)r^2(x)] \le \tilde{\mathbb{E}}[\|\hat{q}\|_1\cdot r^2]\,.$$

Proof. Note that for each monomial x_σ, $\tilde{\mathbb{E}}[(1-x_\sigma)] = \tilde{\mathbb{E}}[(1-x_\sigma)^2]\ge 0$. Using this inequality for each of the monomials in q, the claim follows immediately. □

Now, we prove an SoS analogue of Proposition 3.4.5, which allows us to use the low-multiplicity restrictions of our certificate matrices to get our upper bounds.

Proposition 3.6.5. *Let Φ be a random k-XOR formula in which each clause is sampled independently with probability p. Let $\mathcal{C}_{low}^d\subset[m]^d$ be the set of all ordered multisets of clauses $C_{i_1},\dots,C_{i_d}$ from Φ with the property that if we form two multisets of variables $I,J\in[n]^{dk/2}$ with I containing the first $k/2$ variables of each C_{i_ℓ} and J containing the last $k/2$ variables of each C_{i_ℓ}, then I,J are both low-multipicity multisets, in that both have no element of $[n]$ with multiplicity $\ge 100\log n$.*

Then if $p \geq 200\frac{\log n}{n^{k-1}}$ and $d \ll n$, and if $\tilde{\mathbb{E}}$ is a pseudoexpectation of degree at least $2dk$, then

$$\tilde{\mathbb{E}}[P_\Phi(x)] \leq \left(\mathop{\mathbb{E}}_{i_1,\ldots,i_d \sim \mathcal{C}^d_{low}} \left[\prod_{\ell=1}^{d} P_{i_\ell}(x) \right] \right)^{1/d} + o(1)$$

for all $x \in \{\pm 1\}^n$ with high probability.

Proof. We sample a uniform element $\mathcal{C} \sim \mathcal{C}^d_{low}$, $\mathcal{C} = C_1, \ldots, C_d$ in the following way:

- For $t = 1, \ldots, d$: Let $\mathcal{A}_t \subset \mathcal{I}$ be the set of clauses such that for any $C' \in \mathcal{A}$, $C_1, \ldots, C_{t-1}, C' \in \mathcal{C}^t_{low}$. Choose a uniformly random $C \sim \mathcal{A}_t$ and set $C_t := C$, adding C to $\mathcal{C}$.

This sampling process clearly gives a uniformly random element of $\mathcal{C}^d_{low}$.

Let $P_i(x)$ be the $0-1$ predicate corresponding to whether x satisfies the clause C_i. Let $m_{\max}$ be the maximum number of clauses any variable in Φ participates in. Because $\tilde{\mathbb{E}}\left[(P_i^2(x) - P_i(x))r(x)\right] = 0 \ \forall r(x), \deg(r) \leq 2d$ we can write,

$$\begin{aligned}
\mathop{\mathbb{E}}_{C_1,\ldots,C_d \sim \mathcal{C}^d_{low}} \tilde{\mathbb{E}} \left[\prod_{i\in[d]} P_i(x) \right] &= \mathop{\mathbb{E}}_{C_1,\ldots,C_d \sim \mathcal{C}^d_{low}} \tilde{\mathbb{E}} \left[\prod_{i\in[d]} P_i^2(x) \right] \\
&= \mathop{\mathbb{E}}_{C_1,\ldots,C_{d-1}} \tilde{\mathbb{E}} \left[\left(\prod_{i\in[d-1]} P_i^2(x) \right) \cdot \left(P_\Phi(x) + \Delta_{C_1,\ldots,C_{d-1}}(x) \right) \right]
\end{aligned}$$

where $\Delta_{C_1,\ldots,C_{d-1}}(x) \stackrel{\text{def}}{=} \mathbb{E}\left[P_d^2(x) | C_1, \ldots, C_{d-1}\right] - \mathbb{E}[P_d^2(x)] = \mathbb{E}\left[P_d^2(x) | C_1, \ldots, C_{d-1}\right] - P_\Phi(x)$. By definition, the ℓ_1-norm of the coefficients of the polynomial $\Delta_{C_1,\ldots,C_{k-1}}$ is at most $dm_{max}/100m \log n < o(1)$ with high probability, by concentration argument for m and $m_{\max}$ (see the proof of Proposition 3.4.5) and by our requirement that $d \ll n$. Using Claim 3.6.4, this implies that

$$\begin{aligned}
\mathop{\mathbb{E}}_{C_1,\ldots,C_{d-1}} \tilde{\mathbb{E}} \left[\prod_{i\in[d-1]} P_i^2(x) \cdot \Delta_{C_1,\ldots,C_{d-1}}(x) \right] &\preceq \mathop{\mathbb{E}}_{C_1,\ldots,C_{d-1}} \tilde{\mathbb{E}} \left[\prod_{i\in[d-1]} P_i^2(x) \cdot \|\Delta_{C_1,\ldots,C_{d-1}}(x)\|_1 \right] \\
&\preceq o(1) \cdot \mathop{\mathbb{E}}_{C_1,\ldots,C_{d-1}} \tilde{\mathbb{E}} \left[\prod_{i\in[d-1]} P_i^2(x) \right]
\end{aligned}$$

Therefore,

$$\mathop{\mathbb{E}}_{C_1,\ldots,C_d \sim \mathcal{C}^d_{low}} \tilde{\mathbb{E}} \left[\prod_{i\in[d]} P_i^2(x) \right] \succeq \mathop{\mathbb{E}}_{C_1,\ldots,C_{d-1}} \tilde{\mathbb{E}} \left[(P_\Phi(x) - o(1)) \prod_{i\in[d-1]} \cdot P_i^2(x) \right]$$

Repeating the argument d times, we can conclude that for even d,

$$\mathop{\mathbb{E}}_{C_1,\ldots,C_d\sim\mathcal{C}^d_{low}} \tilde{\mathbb{E}}\left[\prod_{i\in[d]} P_i^2(x)\right] \succeq \tilde{\mathbb{E}}\left[(P_\Phi(x)-o(1))^d\right] \geq \left(\tilde{\mathbb{E}}\left[P_\Phi(x)-o(1)\right]\right)^d$$

Where the last inequality follows from Fact 2.5.3. This concludes the argument. □

This proposition, plugged into the argument from Section 3.4 along with the SoS-ifying steps used for the tensor norm upper bound, gives Theorem 3.1.3 for the even k case. The odd k case can be obtained in a similar way.

Chapter 4

Degree-4 Sum-of-Squares Lower Bounds for Planted Clique

4.1 Introduction

Let $G(n,p)$ be the Erdős-Rényi random graph where each edge is present in G with probability p independently of others. By a standard calculation, the largest clique in $G \sim G(n,\frac{1}{2})$ is of size $(2+o(1)) \cdot \log(n)$ with high probability [GM75, Mat76, BE76]. Recovering such a clique using an efficient algorithm is a long standing open question in theoretical computer science. As early as 1976, Karp [Kar76] suggested the impossibility of finding cliques of size even $(1+\varepsilon)\log(n)$ for any constant $\varepsilon > 0$ in polynomial time. (A greedy approach growing a clique from a random vertex finds a clique of size $(1+o(1))\log n$.) Karp's conjecture was remarkably prescient, and stands its ground after nearly four decades of research.

Lack of algorithmic progress on the question motivated Jerrum [Jer92] and Kucera [Kuc95] to consider a relaxed version known as the *planted clique* problem. In this setting, we are given a graph G obtained by planting a clique of size ω on a graph sampled according to $G(n,\frac{1}{2})$, and we would like to recover the clique efficiently. This variant is also connected to the question of finding large communities in social networks and the problem of *signal finding* in molecular biology [PS000]. The heart of the problem is captured by a decision version: distinguish this planted distribution from $G(n,\frac{1}{2})$. The goal is to do so via an efficient algorithm for as small an ω as possible.

For $\omega > (2+\varepsilon)\log n$, there is a simple quasi-polynomial time algorithm that distinguishes between the two distributions. The algorithm simply tries all subsets of $(2+\varepsilon)\log n$ vertices, looking for a clique. For a random graph $G(n,\frac{1}{2})$, there are no cliques of size $(2+\varepsilon)\log n$, but there is one in the planted distribution. Clearly, the planted clique problem becomes easier as the clique size ω increases. Yet despite much effort there are no polynomial-time algorithms known for this problem for any $\omega < o(\sqrt{n})$. For $\omega = \Omega(\sqrt{n})$, the maximum eigenvalue of the (mean-zero) adjacency matrix suffices to distinguish between the distributions [AKS98]. The Lovász-Schrijver+ (LS+) semi-definite programming hierarchy leads to the state of the art time/clique-size trade-off: it distinguishes between the distributions for $\omega \approx \sqrt{\frac{n}{2^d}}$ in time $n^{O(d)}$ for any $d = O(\log n)$.

While algorithmic progress has been slow, there has been success in proving strong lower

bounds for the planted clique problem within specific algorithmic frameworks. The first such bound was given by Jerrum, who showed that a class of Markov Chain Monte Carlo algorithms require a super-polynomial number of steps to find a clique of size $(1+\varepsilon)\log n$, for any fixed $\varepsilon > 0$, in an instance of $G(n, \frac{1}{2})$ [Jer92]. Feige and Krauthgamer showed that r-levels of the Lovász-Schrijver semi-definite programming hierarchy are needed to find a planted clique of size $\omega \geq \tilde{\Omega}(\sqrt{n}/2^r)$ [FK00, FK03]. Feldman et al. show (for the planted bipartite clique problem) that any "statistical algorithm" cannot distinguish in a polynomial number of queries between the random and planted cases for $\omega < \tilde{O}(\sqrt{n})$ [FGR+12].

Recently, this difficulty of finding cliques of size $\omega \ll \sqrt{n}$ has led to an increasing confidence in planted clique being a candidate for an average case hard problem and has inspired new research directions in cryptography [ABW10b], property testing [AAK+07], machine learning [BR13], algorithmic game theory [HK09, ABC11] and mathematical finance [ABBG10].

In this chapter, we are interested in understanding the performance of the SoS Hierarchy for the planted clique problem. There are quite a few examples where SoS provides better algorithms than the weaker hierarchies—Lovász-Schrijver in particular—to which the best known lower bounds for planted clique apply. To provide just one, while there are instances of unique games that are hard for $\mathrm{poly}(\log\log n)$-rounds of the Lovász-Schrijver SDP hierarchy [KS09, RS09b], recent work has shown that these instances are resolved by degree-8 SOS hierarchy [BBH+12].

Moreover, even the degree-4 SOS relaxation proves to be surprisingly powerful in a few applications:

- First, the work of Barak et al. [BBH+12] shows that a degree-4 SOS relaxation can certify 2-to-4 hypercontractivity of low-degree polynomials over the hypercube. This argument is the reason that hard instances for Lovász-Schrijver and other SDP hierarchies constructed via the *noisy hypercube gadgets* are easily refuted by the SOS hierarchy.

- A degree-4 SOS relaxation can certify that the 2-to-4 norm of a random subspace of dimension at most $o(\sqrt{n})$ is bounded by a constant (with high probability over the choice of the subspace) [BBH+12]. This average-case problem has superficial similarities to the planted clique problem.

Thus a natural question is: can the SoS hierarchy can yield improved algorithms for the planted clique problem? The first published work along these lines was of Meka, Potechin and Wigderson [MPW15] who showed that for every $d \geq 2$, the degree-$2d$ SoS hierarchy cannot find planted cliques of size smaller than $\approx n^{\frac{1}{2d}}$.[1] Deshpande and Montanari [DM15b] independently proved a tighter lower bound of $\approx n^{1/3}$ for the case of degree-4. In the main result of this chapter, we extend the prior works and show that the degree-4 SoS algorithm cannot find cliques of size $\approx \sqrt{n}$, a bound optimal within $\mathrm{poly}\log(n)$ factors. Our lower bound for degree-4 is obtained by a careful "correction" to the certificate used by [MPW15] and [DM15b] in their lower bounds.

[1] We use $\approx$ to denote equality up to factors polylogarithmic in n (the size of the graph) and with an arbitrary dependence on the degree parameter d.

Theorem 4.1.1 (Degree Four)**.** *The canonical degree-4 SoS relaxation of the planted clique problem (4.2.1) has an integrality gap of at least $\tilde{\Omega}(\sqrt{n})$ with high probability.*

We note that this lower bound cannot be obtained using the certificate from [MPW15, DM15b]. An argument of Kelner shows that this certificate fails to show a lower bound on the degree-4 SoS hierarchy when the size of the planted clique is larger than $\approx n^{\frac{1}{3}}$ so the analysis of Deshpande and Montanari is tight. To prove our lower bound, we will modify this certificate to take Kelner's argument into account.

Remark 4.1.2. We note that both Deshpande and Montanari's analysis and Kelner's argument can be generalized to higher degrees, showing that the degree-$2d$ sum of squares hierarchy cannot find planted cliques of size smaller than $\approx n^{\frac{1}{d+1}}$ and this is the best that can be done with this certificate. For details, see the preprint [HKP15].

Other Related Work

The earliest works on proving SoS lower bounds were due to Grigoriev, who showed that degree-$\Omega(n)$ SoS does not beat the random assignment for 3SAT or 3XOR even on random instances from a natural distribution [Gri01b]. Some of these lower bounds were rediscovered by Schoenebeck [Sch08]. Lower bounds for SOS generally rely on gadget reductions from 3SAT or 3XOR and this approach has been studied in some detail [Tul09, BCV+12]. An exception to this methodology is the recent work of Barak et al. in proving SoS lower bounds for pairwise independent CSPs [BCK15].

Turning now to the planted clique problem, the works [FK08, BV09] show that, if one were able to efficiently calculate the injective tensor norm of a certain random order-m tensor, then by extending the spectral algorithm of [AKS98] one would have a polynomial-time algorithm for $k > n^{1/m}$. However, there is no known algorithm that efficiently computes the injective tensor norm of an order-m tensor, and in fact computing the injective tensor norm is hard to approximate in the general case [HM13].

There has also been recent work on Gaussian variants of the problem showing, for example, strong indistinguishability results about the spectrum of the associated matrices with and without planting [MRZ14]. There has also been a line of recent works improving the speed and size of the constant C of algorithms finding $C\sqrt{n}$-size planted cliques [DGGP14, FR10, DM15a].

Finally, the present work builds heavily on independent papers of Meka, Potechin, and Wigderson [MPW15] and Deshpande and Montanari [DM15b], which we have already thoroughly discussed. Since the initial publication of this work, the performance of the SoS Hierarchy has been fully characterized at every degree-d by [BHK+16]. Their result implies that the SoS Hierarchy cannot solve the planted clique problem with $\omega \ll \sqrt{n}$ with $d = o(\log n)$, and they introduce new techniques for constructing SoS certificates for average-case problems.

Organization of This Chapter

In Section 4.2, we give an intuitive account in the language of *pseudo-distributions* of the new feasible solution needed to prove Theorem 4.1.1 and how it fixes the problems with

previously-constructed feasible solutions. In Section 4.3, we give an overview of the analysis required for the degree-4 lower bound building on the work of [DM15b]. Finally, in Section 4.4 we give the complete proof deferring some of the more technical lemmas to Section 4.5.

4.2 Sum of Squares, Simple Moments, and Why They Don't Work

To show that the degree-4 SoS program *fails* to solve the planted clique problem with parameter ω, we show that with high probability there is a solution with objective value at least ω for the program $\mathcal{P}_G$ (known as a *certificate* or *witness*) even when G is a random graph from $G(n, 1/2)$ (which in particular will not contain a clique of size $\gg \log n$).

In previous SoS lower bounds for problems such as random 3XOR/3SAT, Knapsack, and random constraint satisfaction problems [Gri01a, Gri01b, Sch08, BCK15], the certificate X was obtained in a fairly natural way, and the bulk of the work was in the analysis. In fact, in all those cases the certificate used in the SoS lower bounds was the same one that was used before for obtaining lower bounds for weaker hierarchies. The same holds for the previous works for the planted clique problem, where the works of [MPW15, DM15b] used a natural certificate that is a close variant of the certificate used by Feige and Krauthgamer [FK03] for LS+ lower bounds and showed that it satisfies the stronger conditions of the SoS program.

It can be shown that this "natural" certificate of the previous works does *not* satisfy the conditions of the SoS program when $\omega \gg n^{1/3}$, and thus cannot work to obtain a $\approx \sqrt{n}$ lower bound for an SoS program of degree 4 or higher (see Section 4.2 below). Hence, to obtain our tight lower bound for the degree 4 SoS program, we introduce and analyze a more complicated certificate, which can be thought of as making a global "correction" to the simple certificate of [MPW15, DM15b].

We now give an informal overview of the SoS program for planted clique, the [MPW15, DM15b] certificate, our correction to it, and our analysis.

The SoS Program for Max Clique

Let $G = G([n], E)$ be any graph with the vertex set $[n]$ and edge set E. The following polynomial equations ensure that any assignment $x \in \mathbb{R}^n$ must be the characteristic vector of an ω-sized clique in G:

$$\begin{aligned} x_i^2 &= x_i \text{ for all } i \in [n] \\ x_i \cdot x_j &= 0 \text{ for all } \{i,j\} \notin E \\ \sum_{i=1}^{n} x_i &= \omega \,. \end{aligned} \tag{4.2.1}$$

As described in Chapter 2, Section 2.3, we can take the degree-d SoS relaxation for (4.2.1) and obtain a pseudodistribution satisfying:

$$\tilde{\mathbb{E}} \text{ satisfies the equations } \{\forall i, x_i^2 = x_i, \forall i \not\sim j, x_i x_j = 0, \sum_{i \le n} x_i = \omega\} \,.$$

Our subject will be the SoS relaxation with degree $d = 4$.

The "Simple Moments"

To show a lower bound of ω for the SoS relaxation for planted clique, it suffices to show that for a random graph G, we can find a degree d pseudo-expectation operator that satisfies (4.2.1). Both previous papers [MPW15] and [DM15b] utilize essentially the same operator, which we call here the "simple moments". It is arguably the most straightforward way to satisfy the conditions of (4.2.1), and the bulk of the work is then in showing the positivity constraint that $\tilde{\mathbb{E}}\, P^2 \geq 0$ for every P of degree ≤ 2 (in the degree 4 case). [DM15b] shows that this will hold as long as $\omega \ll n^{1/3}$ and an argument of Kelner (see Section 4.2 below) shows that this is tight and in fact these simple moments fail to satisfy the positivity conditions for $\omega \gg n^{1/3}$.

To define a degree d pseudo-expectation operator $\tilde{\mathbb{E}}$, we need to choose some basis $\{P_1, \ldots, P_N\}$ for the set of polynomials of degree at most d and define $\tilde{\mathbb{E}}\, P_i$ for every i. The simplest basis is simply the monomial basis. Moreover, since our pseudo-expectation satisfies the constraints $\{x_i^2 = x_i\}$, we can restrict attention to *multilinear* monomials, of the form $x_S = \prod_{i \in S} x_i$ for some $S \subseteq [n]$. Note also that the constraints $x_i x_j = 0$ for $\{i, j\} \notin E$ imply that we must define $\tilde{\mathbb{E}}\, x_S = 0$ for every S that is not a clique in G. Indeed, the pseudo-distribution $\{x\}$ is supposed to mimic an actual distribution over the characteristic vectors of ω-sized cliques in G, and note that in any such distribution it would hold that $\tilde{\mathbb{E}}\, x_S = 0$ when S is not a clique.

The simplest form of such a pseudo-distribution is to set

$$\tilde{\mathbb{E}}\, x_S = \begin{cases} 0 & S \text{ is not a clique} \\ \alpha_{|S|} & \text{otherwise} \end{cases}$$

where $\alpha_{|S|}$ is a constant depending only on the size of S. We can compute the value $\alpha_{|S|}$ by noting that we need to satisfy $\tilde{\mathbb{E}}(\sum_i x_i)^\ell = \sum_{i_1, \ldots, i_\ell} \tilde{\mathbb{E}}\, x_{i_1} \cdots x_{i_\ell} = \omega^\ell$ for every $\ell = 1, \ldots, d$. Since there would be about $\binom{n}{\ell} 2^{-\binom{\ell}{2}}$ ℓ-sized cliques in the graph G, the value α_ℓ will be $\approx \left(\frac{\omega}{n}\right)^\ell$.[2]

There is Such a Thing as too Simple

The simple moments are essentially the same ones used by Feige and Krauthgamer [FK03] for LS+ (a weaker convex hierarchy than SoS), where they were shown to be valid for the constraints of this problem as long as $\omega < \sqrt{n/2^{d+1}}$. Initially, Meka and Wigderson conjectured that a similar bound holds for the SoS program, or in other words, that the $\binom{n}{\leq d/2} \times \binom{n}{\leq d/2}$ matrix X where $X_{S,T} = \tilde{\mathbb{E}}\, x_S x_T$ for every $S, T \subseteq [n]$ of size $\leq d/2$ is positive semidefinite as long as $\omega \ll \sqrt{n}$. However, this conjecture was later shown to be false–an

[2] These moments must be modified slightly to exactly satisfy the constraint $\sum x_i = \omega$. However, these corrections have very small magnitudes and so all the observations below apply equally well to the modified moments, and so we ignore this issue in this informal overview.

argument due to Jonathan Kelner, described in [Bar14], shows that (for $d = 4$) the matrix X is *not* positive semidefinite as long as $\omega \gg n^{1/3}$. We review this argument below, as it is instructive for our correction.

The Kelner argument notices that the simple moments turn out to be "too random" in that they fail to account for some structure that the graph possesses. The idea is as follows. Intuitively, 4-cliques which have a higher number of common neighbors are more likely to be in the planted clique, so we should have a higher value for $\tilde{\mathbb{E}}[x_S]$ when S has a higher number of common neighbors. However, the simple moments give the same value of $\tilde{E}[x_S]$ for all 4-cliques S.

This causes a discrepancy between the simple moments and what we would expect if we actually had a planted clique. To see this, we define a value r_S for each set of vertices S which measures how S is connected to the other vertices in the graph. Using these values r_S, we will define a polynomial $g(x)$ so that $\tilde{\mathbb{E}}\, g(x)^2 \ll 0$, which shows that the matrix X is not positive semi-definite.

Definition 4.2.1. Define $r_{i,j}$ so that $r_{i,j}$ equals $+1$ when $\{i,j\} \in E$, equals -1 when $\{i,j\} \notin E$, and equals 0 when $i = j$

Definition 4.2.2. Given a set of vertices S, define $r_S = \sum_{i \notin S} \prod_{j \in S} r_{i,j}$

We now consider the expression $\tilde{\mathbb{E}}[\sum_{S \in \binom{n}{4}} r_S x_S]$. With our simple moments, this has expected value 0 over the input graph because for each S, r_S is independent of x_S and has mean 0. However, if we instead consider the planted distribution where we plant a clique of size ω at random, choose the rest of the graph randomly, and then take $x_S = 1$ if S is part of the planted clique and 0 otherwise, then we obtain a different expected value for this expression.

To see this, first randomly choose where the planted clique is. There will be $\binom{\omega}{4}$ different S which are part of the planted clique. For each of these S, there are $(\omega - 4)$ $i \notin S$ which are also part of the planted clique. For each of these i, $\prod_{j \in S} r_{i,j} = 1$. For the remaining i, $\prod_{j \in S} r_{i,j}$ has expected value 0 over the remainder of the graph. Thus, the expected value of r_S is $\omega - 4$. Putting everything together, $\sum_{S \in \binom{n}{4}} r_S x_S$ has expected value $(\omega - 4)\binom{\omega}{4} \approx \omega^5$

This discrepancy can be exploited with the following polynomial. Consider the polynomial $g_i = Cx_i - \sum_{j,k:j<k} r_{i,j} r_{i,k} x_j x_k$ where C is a constant that will be chosen later.

$$g_i^2 = C^2 x_i - 2C \sum_{j,k:j<k} r_{i,j} r_{i,k} x_i x_j x_k + \sum_{j,k,l,m:j<k,l<m} r_{i,j} r_{i,k} r_{i,l} r_{i,m} x_j x_k x_l x_m$$

We now consider the expected value of $\tilde{\mathbb{E}}[g_i^2]$ on a random graph. Recall that for the simple moments, $\tilde{\mathbb{E}}[x_S] = \alpha_S \approx \left(\frac{\omega}{n}\right)^{|S|}$ if S is a clique and 0 otherwise.

1. For all graphs G, $\tilde{\mathbb{E}}[C^2 x_i] = C^2 \alpha_1 \approx \frac{C^2 \omega}{n}$

2. For the second term, note that if i, j, k are distinct, $\tilde{\mathbb{E}}[x_i x_j x_k] = 0$ unless $r_{i,j} = r_{i,k} = 1$. Thus, all of the terms in the sum add together and we have that

$$E_G\left[\tilde{\mathbb{E}}[2C \sum_{j,k:j<k} r_{i,j} r_{i,k} x_i x_j x_k]\right] = E_G\left[\tilde{\mathbb{E}}[2C \sum_{j,k:j<k, j\neq i, k\neq i} x_i x_j x_k]\right]$$
$$= E_G\left[2C(\#\text{ of 3-cliques containing } i)\alpha_3\right] \approx \frac{C\omega^3}{n}$$

3. For the third term, note that when we take the expectation over the randomness of G, the only terms that don't have expectation 0 are the ones where $l = j \neq i$ and $m = k \neq i$. Thus,

$$E_G\left[\tilde{\mathbb{E}}[\sum_{j,k,l,m:j<k,l<m} r_{i,j} r_{i,k} r_{i,l} r_{i,m} x_j x_k x_l x_m]\right]$$
$$= E_G\left[\tilde{\mathbb{E}}[\sum_{j,k:j<k, j\neq i, k\neq i} x_j x_k]\right]$$
$$= E_G\left[(\#\text{ of 2-cliques in G not containing } i)\alpha_2\right]$$
$$\approx \omega^2$$

Putting everything together, if $\omega >> n^{\frac{1}{3}}$ then we can take $C \approx \omega^2$ and we will have that $E_G\left[\tilde{\mathbb{E}}[g_i^2]\right] < 0$.

Note that this argument breaks down if x_S depends on the variables $r_{i,j}$. Further note that if we consider the terms in $\sum_{j,k,l,m:j<k,l<m} r_{i,j} r_{i,k} r_{i,l} r_{i,m} x_j x_k x_l x_m$ where j, k, l, m are all distinct and sum these terms over all i, we will obtain a multiple of $\sum_{S\in\binom{\omega}{4}} r_S x_S$. Thus, the failure of this polynomial to have non-negative expectation for the simple moments is closely related to the fact that the expected value of $\tilde{\mathbb{E}}[\sum_{S\in\binom{\omega}{4}} r_S x_S]$ is too low for the simple moments.

Fixing the Simple Moments

Our fix for the simple moments is motivated by the example above. We want to ensure that $\tilde{\mathbb{E}}[\sum_{S\in\binom{n}{4}} r_S x_S] \approx \omega^5$ to match what we would expect if there was a planted clique. The idea is to break the symmetry between different equal-sized cliques and give a significantly higher pseudo-expectation to 4-cliques S which for which r_S is high. Roughly speaking, the corrected moments will set

$$\tilde{\mathbb{E}}[x_S] = \alpha_{|S|}(1 + r_S\omega/n)$$

for every 4-clique S. Note that when $\omega = \varepsilon\sqrt{n}$, the correction factor would typically be of the form $1 \pm \Theta(\varepsilon)$.[3]

Computing the expected value of $\tilde{\mathbb{E}}[\sum_{S\in\binom{n}{4}} r_S x_S]$ under the new moments we obtain that

$$E_G\left[\tilde{\mathbb{E}}[\sum_{S\in\binom{n}{4}} r_S x_S]\right] = E_G\left[\sum_{S\in\binom{n}{4}} r_S(1+\frac{\omega}{n}r_S)\alpha_{|S|}1[S \text{ is a clique}]\right]$$
$$= E_G\left[\sum_{S\in\binom{n}{4}} \frac{\omega}{n}(r_S)^2\alpha_{|S|}1[S \text{ is a clique}]\right]$$

There are $\approx n^4$ choices for S. For each S, r_S is the sum of $n-4$ independent ± 1 variables so with high probability $r_S^2 \approx n$. $\alpha_4 \approx \left(\frac{\omega}{n}\right)^4$. Putting everything together, this expression has expected value $\approx \omega^5$, as desired.

This gives some intuition why the corrected moments might be better than the simple moments for a particular family of polynomials. In the remainder of the chapter, we turn to the technical arguments establishing that the correction does in fact yield valid pseudomoments.

4.3 Overview of our Analysis

In this section, we describe the degree-4 SoS relaxation for the max-clique SDP and discuss the aforementioned simple moments (used by Meka, Potechin, and Wigderson and Deshpande and Montanari) formally. We then describe our own modified moments, and give an overview of the proof that they form a feasible solution to the program (the difficult part being showing that they are PSD).

Recalling our previous informal discussion, we first construct a solution to the relaxation which satisfies only some of the constraints (we will ignore some aspects of the constraint $\sum_i x_i = \omega$ from (4.2.1)). Then in the final proof we show how to satisfy $\sum_i x_i = \omega$ as a constraint for some $\omega = \tilde{\Omega}(\sqrt{n})$.

Degree-4 SoS Relaxation for Max Clique

We consider the following semidefinite program.

$$\max_{\omega\geq 0} \omega \quad \text{such that there exists a degree-4 } \tilde{\mathbb{E}} \text{ satisfying}$$
$$\{x_i^2 = x_i \text{ for all } i\in[n],\ x_S = 0 \text{ for } S \text{ not a clique in } G, \text{ and } \sum_{i\in[n]} x_i = \omega\}. \tag{4.3.1}$$

[3] While it might seem that there is a chance for these pseudo-expectations to be negative, if $\omega < \sqrt{n}/\operatorname{polylog}(n)$ then it is exceedingly unlikely that there will exists an S such that $|r_S| > n/\omega$, and so we ignore this issue in this overview.

If $\mathsf{sdpval}(G)$ denotes the optimum value of the SDP relaxation on graph G, then clearly $\mathsf{sdpval}(G)$ is at least the size of the maximum clique in G. In order to prove a lower bound for degree-4 SoS relaxation on $\mathbb{G}(n, \frac{1}{2})$, it is sufficient to argue that with overwhelming probability, $\mathsf{sdpval}(G)$ is significantly larger than the size of the maximum clique in a random graph. This amounts to exhibiting a feasible SDP solution with large objective value for an overwhelming fraction of graphs sampled from $\mathbb{G}(n, \frac{1}{2})$. Formally, we show the following:

Theorem 4.3.1 (Formal version of Theorem 4.1.1)**.** *There exists an absolute constant $c \in \mathbb{N}$ such that*

$$\mathop{\mathbb{P}}_{G \sim \mathbb{G}(n,\frac{1}{2})} \left\{ \mathsf{sdpval}(G) \geq \frac{\sqrt{n}}{\log^c n} \right\} \geq 1 - O(n^{-4})$$

The Simple Moments, Formally

Henceforth, fix a graph G that is sampled from $\mathbb{G}(n, \frac{1}{2})$. Both the work of Meka, Potechin and Wigderson [MPW15] and that of Deshpande and Montanari [DM15b] construct essentially the same SDP solution for the degree-4 SoS relaxation.

This SDP solution assigns to each clique of size $1, \ldots, d$, a value that depends only on its size (in our case, $d = 4$). More formally, the SDP solution in [DM15b] is specified by four parameters $\underline{\alpha} = \{\alpha\}_{i=1}^{4}$ as,

$$\tilde{\mathbb{E}}[x_S] = \alpha_{|S|} \cdot \mathcal{G}_S \,,$$

where for a set of vertices $A \subseteq V$, $\mathcal{G}_A$ is the indicator that the subgraph induced on A is a clique. The parameters $\{\alpha\}_{i \in [4]}$ determine the value of the objective function, and the feasibility of the solution. As a convention, we will define $\alpha_0 = 1$. It is easy to check that this solution satisfies all the linear constraints of the SoS program (4.3.1) except for $\sum_i x_i = \omega$, which will be handled later, since it assigns non-zero values only to cliques in G. The key difficulty is in showing positive semi-definiteness for an appropriate choice of $\underline{\alpha}$.

For this purpose, we switch notation to explicitly discuss the moment matrix $M = M(G, \underline{\alpha})$ associated to $\tilde{\mathbb{E}}$:

$$M(G, \underline{\alpha})_{A,B} = \tilde{\mathbb{E}}\, x_{A \cup B} = \alpha_{|A \cup B|} \cdot \mathcal{G}_{A \cup B} \,.$$

In order to show that $M(G, \underline{\alpha}) \succeq 0$ it is sufficient to show that $N(G, \underline{\alpha}) \succeq 0$, where

$$N_{A,B} = \alpha_{|A \cup B|} \cdot \prod_{i \in A \setminus B, j \in B \setminus A} \mathcal{G}_{ij}.$$

(Here $\mathcal{G}_{ij}$ is the indicator for the presence of the edge (i, j).) In words, N is the matrix where the entry $\{a, b, c, d\}$ is proportional not to the indicator of whether $\{a, b, c, d\}$ is a clique, but to the indicator of whether G has as a subgraph the bipartite clique with bipartitions $\{a, b\}$ and $\{c, d\}$. It is easy to see that the matrix M is obtained by dropping from N the rows and columns corresponding $\{a, b\} \in \binom{n}{2}$ where $(a, b) \notin E(G)$. Hence $N \succeq 0 \implies M \succeq 0$. This avoids the many all-zero rows in M corresponding to non-edges in G. (This removes some spurious variance in the spectrum of the matrix.)

Notice that N is a random matrix whose entries depend on the edges in the random graph G. At the risk of over-simplification, the approach of both the previous works [MPW15] and [DM15b] can be broadly summarized as follows:

1. (Expectation) Show that the expected matrix $\mathbb{E}[N]$ has sufficiently large positive eigenvalues.

2. (Concentration) Show that with high probability over the choice of G, the *noise* matrix $N - \mathbb{E}[N]$ has bounded eigenvalues, so as to ensure that $N = \mathbb{E}[N] + (N - \mathbb{E}[N]) \succeq 0$

Here we will sketch a few key details of the argument in [DM15b]. The matrix $N \in \mathbb{R}^{\binom{n}{\leq 2}\times\binom{n}{\leq 2}}$ can be decomposed in to blocks $\{N_{ab}\}_{a,b\in\{0,1,2\}}$ where $N_{a,b} \in \mathbb{R}^{\binom{n}{a}\times\binom{n}{b}}$. Deshpande and Montanari use the Schur complements to reduce the problem of proving that $N \succeq 0$ to facts about the blocks $\{N_{ab}\}_{a,b\in\{0,1,2\}}$. Specifically, they show the following lemma:[4]

Lemma 4.3.2. *Let $\mathcal{A} \in \mathbb{R}^{\binom{n}{\leq 2}\times\binom{n}{\leq 2}}$ be the matrix defined so that $\mathcal{A}_{A,B} = \alpha_{|A|}\alpha_{|B|}$. For $a, b \in \{0, 1, 2\}$, let $\tilde{H}_{a,b}$ be the submatrix of $N(G, \alpha) - \mathcal{A}$ corresponding to moments of order $a + b$. Then $N(G, \alpha)$ is PSD if and only if*

$$\tilde{H}_{11} \succeq 0, \tag{4.3.2}$$

$$\tilde{H}_{22} - \tilde{H}_{12}^{\top}\tilde{H}_{11}^{-1}\tilde{H}_{12} \succeq 0 \tag{4.3.3}$$

The most significant challenge is to argue that (4.3.3) holds with high probability. In fact, the inequality only holds for the Deshpande-Montanari SDP solution with high probability for parameters α for which the objective value is $\omega \ll n^{1/3}$.

Definition 4.3.3 (The matrix H_{22})**.** We use a somewhat different approach (pioneered in [MPW15]) to handle the blocks of N corresponding to monomials of degree less than 4. Departing slightly from [DM15b], we define the matrix H_{22} as the submatrix of $N(G, \alpha)$ with rows and columns indexed by size-2 subsets $\{i, j\} \subseteq \binom{[n]}{2}$.

Our main goal now is to show that $H_{22} \succeq 0$ in a robust way. (Some robustness is necessary to satisfy a condition similar to (4.3.3).)

Expected matrix. The expected matrix $\mathbb{E}[H_{22}]$ is symmetric with respect to permutations of the vertices. It forms an *association scheme* (see [MPW15, DM15b]), by virtue of which its eigenvalues and eigenspaces are well understood. In particular, the following proposition in [DM15b] is an immediate consequence of the theory of association schemes.

Proposition 4.3.4 (Adapted from Proposition 4.16 in [DM15b])**.** *$\mathbb{E}[H_{22}]$ has three eigenspaces, V_0, V_1, V_2 such that*

$$\mathbb{E}[H_{22}] = \lambda_0\Pi_0 + \lambda_1\Pi_1 + \lambda_2\Pi_2,$$

[4] We will use an alternate approach from [MPW15] for our final PSDness argument, but we use the analysis of the matrix H_{22} from [DM15b].

where Π_0, Π_1, Π_2 are the projections to the spaces V_0, V_1, V_2 respectively. The eigenvalues are given by,

$$\lambda_0(\underline{\alpha}) \stackrel{\text{def}}{=} \alpha_2 + (n-2)\alpha_3 + \frac{(n-2)(n-3)}{32} \cdot \alpha_4 \tag{4.3.4}$$

$$\lambda_1(\underline{\alpha}) \stackrel{\text{def}}{=} \alpha_2 + \frac{(n-4)}{2}\alpha_3 - \frac{(n-3)}{16}\alpha_4 \tag{4.3.5}$$

$$\lambda_2(\underline{\alpha}) \stackrel{\text{def}}{=} \alpha_2 - \alpha_3 + \frac{\alpha_4}{16} \tag{4.3.6}$$

Further the eigenspaces are given by,

$$\begin{aligned} V_0 &= \text{Span}\{\mathbb{1}\}, \\ V_1 &= \text{Span}\{u \mid \langle u, \mathbb{1}\rangle = 0,\ u_{i,j} = x_i + x_j \text{ for } x \in \mathbb{R}^n\}, \\ V_2 &= R^{\binom{n}{\leq 2}} \setminus (V_0 \cup V_1). \end{aligned}$$

where the final subspace V_2 we define simply as the subspace of $\mathbb{R}^{\binom{n}{2}}$ orthogonal to V_0, V_1.

Deviation from Expectation. Given the lower bound on eigenvalues of the expected matrix $\mathbb{E}[H_{22}]$, the next step would be to bound the spectral norm of the noise $H_{22} - \mathbb{E}[H_{22}]$. However, since the eigenspaces of $\mathbb{E}[H_{22}]$ are stratified (for the given $\underline{\alpha}$), with one large eigenvalue and several much smaller eigenvalues, standard matrix concentration does not suffice to give tight bounds. To overcome this, Deshpande and Montanari split H_{22} and $H_{12}^\top H_{11}^{-1} H_{12}$ along the eigenspaces of $\mathbb{E}[H_{22}]$. (We will use ideas from [MPW15] to avoid handling $H_{12}^\top H_{11}^{-1} H_{12}$ explicitly.)

Before splitting up H_{22} along eigenspaces of $\mathbb{E}[H_{22}]$, we remove all of the entries $H_{22}(A, B)$ with $A \cap B \neq \emptyset$, to make the analysis easier. Let

$$H_{22} - \mathbb{E}[H_{22}] = Q + K$$

where Q includes all multilinear entries and K includes all non-multilinear entries, i.e., entries $K(A, B)$ where $A \cap B \neq \emptyset$. Formally,

$$Q(A, B) = \begin{cases} H_{22}(A, B) - \mathbb{E}[H_{22}](A, B) & \text{if } A \cap B = \emptyset \\ 0 & \text{otherwise} \end{cases}$$

The spectral norm of the matrix Q over the eigenspaces V_0, V_1, V_2 is carefully bounded in [DM15b]. (We reprove this in Section 4.5 for completeness.)

Lemma 4.3.5 (Adapted from Propositions 4.20, 4.25 in [DM15b], reproved with these parameters in Section 4.5)**.** *With probability at least $1 - O(n^{-4})$, all of the following bounds hold:*

$$\|\Pi_a Q \Pi_b\| \lesssim \alpha_4 n^{3/2} \log^3(n) \qquad \forall (a, b) \in \{0, 1, 2\}^2 \tag{4.3.7}$$

$$\|\Pi_2 Q \Pi_2\| \lesssim \alpha_4 n \log^3(n) \tag{4.3.8}$$

$$\|K\| \lesssim \alpha_3 n^{1/2} \log^3(n) \tag{4.3.9}$$

Proposition 4.3.4 and Lemma 4.3.5 are sufficient to conclude that $H_{22} \succeq 0$ for parameter choices of α that correspond to planted clique of size up to $\omega \ll n^{1/3}$. More precisely, to argue that with high probability $H_{22} \succeq 0$, it is sufficient to argue that $\mathbb{E}[H_{22}] \succeq \mathbb{E}[H_{22}] - H_{22}$, i.e.,

$$\begin{bmatrix} \lambda_0 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} \succeq \begin{bmatrix} \|\Pi_0 Q \Pi_0\| & \|\Pi_0 Q \Pi_1\| & \|\Pi_0 Q \Pi_2\| \\ \|\Pi_1 Q \Pi_0\| & \|\Pi_1 Q \Pi_1\| & \|\Pi_1 Q \Pi_2\| \\ \|\Pi_2 Q \Pi_0\| & \|\Pi_2 Q \Pi_1\| & \|\Pi_2 Q \Pi_2\| \end{bmatrix} + \alpha_3 n^{1/2} \log^3(n) \cdot \mathrm{Id} \, .$$

Deshpande and Montanari fix $\alpha_1 = \kappa$, $\alpha_2 = 4\kappa^2$, $\alpha_3 = 8\kappa^3$ and $\alpha_4 = 512\kappa^4$ for a parameter κ (we think of $\kappa \approx \omega/n$ for the clique-size parameter ω). Using Proposition 4.3.4 and Lemma 4.3.5 to substitute in the corresponding bounds for $\lambda_0, \lambda_1, \lambda_2$ and $\|\Pi_a(\mathbb{E}[H_{22}] - H_{22})\Pi_b\|$ for all $a, b \in \{0, 1, 2\}$, the above matrix inequality becomes

$$\begin{bmatrix} n^2\kappa^4 & 0 & 0 \\ 0 & n\kappa^3 & 0 \\ 0 & 0 & \kappa^2 \end{bmatrix} \succeq \kappa^4 \begin{bmatrix} n^{3/2} & n^{3/2} & n^{3/2} \\ n^{3/2} & n^{3/2} & n^{3/2} \\ n^{3/2} & n^{3/2} & n \end{bmatrix} , \tag{4.3.10}$$

which can be shown to hold for $\kappa \ll n^{-2/3}$. Eventually, it is necessary to show (4.3.3), which is stronger than $H_{22} \succeq 0$. This is again achieved by showing bounds on the eigenvalues of H_{11}^{-1} and H_{12}. We refer the reader to [DM15b] for more details of the arguments.

Problematic Subspace

The SDP solution described above ceases to be PSD at $\kappa \simeq n^{-2/3}$, which corresponds to an objective value of $\omega = \Theta(n^{1/3})$. The specific obstruction to $H_{22} \succeq 0$ arises out of (4.3.10). More precisely, the bottom 2×2 principal minor which yields the constraint

$$\begin{bmatrix} \lambda_1 & -\|\Pi_1 Q \Pi_2\| \\ -\|\Pi_2 Q \Pi_1\| & \lambda_2 \end{bmatrix} \approx \begin{bmatrix} n\kappa^3 & -n^{3/2}\kappa^4 \\ -n^{3/2}\kappa^4 & \kappa^2 \end{bmatrix} \succeq 0 \, ,$$

forcing $\kappa \ll n^{-2/3}$. When κ is larger than this, the problematic vectors $x \in \mathbb{R}^{\binom{n}{2}}$ for which $x^\top H_{22} x < 0$ are precisely those for which $|x^\top \Pi_2 Q \Pi_1 x|$ is large, i.e., $\Pi_2 x$ aligns $Q\Pi_1 x$. More generally, $\Pi_2 x$ should have large projection into the image of $V_0 \oplus V_1$ under Q. Such a vector only exists because the eigenspaces of Q are not the same as those of $\mathbb{E} H_{22}$.

In fact, we identify a specific subspace W that is problematic for the [DM15b] solution. This subspace is the formal counterpart of the bad polynomials $g_i(x)$ described in Section 4.2. To describe the subspace, let us recall some notation. Define the random variable $r_i(j)$ to be -1 if $(i, j) \notin E$, and $+1$ otherwise. We follow the convention that $r_i(i) = 0$.

Lemma 4.3.6. *Let the vectors $r_1^{\otimes 2}, \ldots, r_n^{\otimes 2} \in \mathbb{R}^{\binom{n}{2}}$ be defined so that $r_i^{\otimes 2}(k, \ell) \stackrel{\text{def}}{=} r_i(k) r_i(\ell)$, and let $W \stackrel{\text{def}}{=} \mathrm{Span}\{r_1^{\otimes 2}, \ldots, r_n^{\otimes 2}\}$. (Note that the polynomials discussed in Section 4.2 are given by $r_i(x)^2 = \langle x^{\otimes 2}, r_i^{\otimes 2} \rangle$.) Then with probability at least $1 - O(n^{-4})$,*

$$\|\Pi_2 Q - \Pi_2 \Pi_W Q\| \lesssim \alpha_4 n \log^3(n)$$

Proof. This is an immediate observation from the various matrix norm bounds in [DM15b]. The full proof is in [RS15]. □

Since $\|\Pi_2 Q \Pi_1\| \gg \alpha_4 n \log^3(n)$, the above lemma implies that all the vectors with large singular values for Q are within the subspace W. Furthermore, we will show the following lemma, which clearly articulates that W is the sole obstruction to $H_{22} \succeq 0$.

Lemma 4.3.7. *Suppose $\underline{\alpha} \in \mathbb{R}^4_+$ satisfies*

$$\min(\lambda_0(\underline{\alpha}), \lambda_1(\underline{\alpha}), \lambda_2(\underline{\alpha})) \gg \alpha_3 n^{1/2} \log^3(n)\,, \tag{4.3.11}$$

$$\lambda_0(\underline{\alpha}) > \lambda_1(\underline{\alpha}) \gg \alpha_4 n^{3/2} \log^3(n)\,, \tag{4.3.12}$$

$$\lambda_2(\underline{\alpha}) \gg \alpha_4 n \log^3(n) \tag{4.3.13}$$

then with probability $1 - O(n^{-4})$,

$$H_{22} \succeq \frac{1}{4} \cdot \mathbb{E}[H_{22}] - \frac{16\|Q\|^2}{\lambda_1} \cdot \Pi_2 \Pi_W \Pi_2.$$

Proof. Fix $\theta = \frac{16\|Q\|^2}{\lambda_1}$. Recall that $H_{22} - \mathbb{E}[H_{22}] = Q + K$. We can write the matrix

$$H_{22} + \theta \cdot \Pi_2 \Pi_W \Pi_2 = B_{W^\perp} + B_W + B_K + \frac{1}{4}\mathbb{E}[H_{22}]\,,$$

where if $\Pi_{W^\perp} \stackrel{\text{def}}{=} \mathrm{Id} - \Pi_W$,

$$B_{W^\perp} = \frac{1}{4}\mathbb{E}[H_{22}] + \begin{bmatrix} \Pi_0 Q \Pi_0 & \Pi_0 Q \Pi_1 & \Pi_0 Q \Pi_{W^\perp} \Pi_2 \\ \Pi_1 Q \Pi_0 & \Pi_1 Q \Pi_1 & \Pi_1 Q \Pi_{W^\perp} \Pi_2 \\ \Pi_2 \Pi_{W^\perp} Q \Pi_0 & \Pi_2 \Pi_{W^\perp} Q \Pi_1 & \Pi_2 Q \Pi_2 \end{bmatrix}$$

and

$$B_W = \frac{1}{4}\mathbb{E}[H_{22}] + \begin{bmatrix} 0 & 0 & \Pi_0 Q \Pi_W \Pi_2 \\ 0 & 0 & \Pi_1 Q \Pi_W \Pi_2 \\ \Pi_2 \Pi_W Q \Pi_0 & \Pi_2 \Pi_W Q \Pi_1 & \theta \cdot \Pi_2 \Pi_W \Pi_2 \end{bmatrix}$$

and $B_K = K + \frac{1}{4}\mathbb{E}[H_{22}]$.

It is sufficient to show that $B_{W^\perp}, B_W$ and $B_K \succeq 0$. Using Proposition 4.3.4 and (4.3.9), $B_K \succeq (\frac{1}{4}\lambda_0 - \alpha_3 n^{1/2}\log^3(n))\Pi_0 + (\frac{1}{4}\lambda_1 - \alpha_3 n^{1/2}\log^3(n))\Pi_1 + (\frac{1}{4}\lambda_2 - \alpha_3 n^{1/2}\log^3(n))\Pi_2 \succeq 0$ when condition (4.3.11) holds. Using Proposition 4.3.4, Lemma 4.3.5, and Lemma 4.3.6 we can write,

$$B_{W^\perp} \succeq \frac{1}{4} \cdot \begin{bmatrix} \lambda_0 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} - \alpha_4 \log^3(n) \cdot \begin{bmatrix} n^{3/2} & n^{3/2} & n \\ n^{3/2} & n^{3/2} & n \\ n & n & n \end{bmatrix}$$

which is PSD given the bounds on $\lambda_1, \lambda_2, \lambda_3$ in conditions (4.3.12) and (4.3.13). To see this, one shows that all the 2×2 principal minors are PSD.

Putting everything together, for any $x \in \mathbb{R}^{\binom{n}{2}}$, we can write

$$x^\top B_W x \geq \lambda_0 \|\Pi_0 x\|^2 + \frac{\theta}{2}\|\Pi_W \Pi_2 x\|^2 - 2\|Q\| \cdot \|\Pi_W \Pi_2 x\| \cdot \|\Pi_0 x\| + \lambda_1 \|\Pi_1 x\|^2 + \frac{\theta}{2}\|\Pi_W \Pi_2 x\|^2 - 2\|Q\| \cdot \|\Pi_W \Pi_2 x\| \cdot \|\Pi_1 x\|$$

Now we will appeal to the fact that a quadratic $f(p,q) = ap^2 + 2bpq + cq^2 \geq 0$ for all $p, q \in \mathbb{R}$ if $b^2 < 4ac$ and $a > 0$. Since $\theta\lambda_1, \theta\lambda_0 \geq 16\|Q\|^2$ by condition (4.3.12), it is easily seen that the above quadratic form is always non-negative, implying that $B_W \succeq 0$. □

An immediate corollary of the proof of the above lemma is the following.

Corollary 4.3.8. *Under the hypothesis of Lemma 4.3.7, with probability* $1 - O(n^{-4})$,

$$H_{22} - K \succeq \frac{1}{2} \cdot \mathbb{E}[H_{22}] - \frac{16\|Q\|^2}{\lambda_1} \cdot \Pi_2 \Pi_W \Pi_2.$$

The above corollary is a consequence of the fact that $H_{22} - K = B_W + B_{W^\perp} + \frac{1}{2}\mathbb{E}[H_{22}]$.

The Corrected Witness

We now offer an alternative, more technical motivation for our correction to the witness. Suppose we have an unconstrained matrix M that we wish to modify as little as possible so as to ensure $M \succeq 0$. Given a test vector w so that $w^\top M w < 0$, the natural update to make is to take $M' = M + \beta \cdot ww^\top$ for a suitably chosen β. In this way, one would hope to add enough positive spectral mass in the direction of w to eliminate the negative eigenvalue. This would suggest creating a new SDP solution by setting $H'_{22} = H_{22} + \beta \sum_{i \in [n]} (r_i^{\otimes 2})(r_i^{\otimes 2})^\top$.

Unfortunately, the SoS SDP relaxation has certain hard constraints, namely that the non-clique entries are fixed at zero. Moreover, $\tilde{\mathbb{E}}[x_{S_1} x_{S_2}]$ must depend only on $S_1 \cup S_2$. Setting the SDP solution matrix to $H_{22} + \beta \sum_{i \in [n]} (r_i^{\otimes 2})(r_i^{\otimes 2})^\top$ would almost certainly violate both these constraints. It is thus natural to consider multiplicative updates to the entries of the matrix, which clearly preserve the zero entries of the matrix.

Specifically, the idea is to consider an update of the form $M' = M + \beta D_w M D_w$ where D_w is the diagonal matrix with entries given by the vector w. If the matrix M has a significantly large eigenvalue along $\mathbb{1}$ (the vector of all 1 entries), i.e., $M \succeq \lambda_0 \cdot \tilde{\mathbb{1}}\tilde{\mathbb{1}}^\top$, then this multiplicative update has a similar effect as an additive update,

$$M' \succeq M + \beta \cdot \lambda_0 \cdot ww^\top.$$

Recall that, in our setting, the Deshpande-Montanari SDP solution matrix N does have a large eigenvalue along $\tilde{\mathbb{1}}$. We now formally describe the matrix M' which will form the basis for our final SDP solution $\tilde{\mathbb{E}}$.

Definition 4.3.9. Let $\hat{r}_1, \ldots, \hat{r}_n \in \mathbb{R}^{\binom{n}{\leq 2}}$ be defined so that

$$\hat{r}_i(A) = \begin{cases} 0 & |A| < 2 \\ r_i^{\otimes 2}(A) & |A| = 2. \end{cases}$$

Define $\hat{D}_i \in \mathbb{R}^{\binom{n}{\leq 2}}$ to be the diagonal matrix with $\hat{r}_i$ on the diagonal. Define $\hat{K}$ to be the restriction of $N(G, \underline{\alpha})$ to the non-multilinear entries. Also let

$$N'(G, \underline{\alpha}) = N(G, \underline{\alpha}) + \beta \cdot \sum_{i \in [n]} \hat{D}_i \left(N(G, \underline{\alpha}) - \hat{K} \right) \hat{D}_i,$$

where $\beta = \frac{1}{100\sqrt{n}\log n}$. Then define the matrix M' so that

$$M'(G, \underline{\alpha}) = \mathcal{P}\left(N'(G, \underline{\alpha}) \right),$$

where $\mathcal{P}$ is the projection that zeros out rows and columns corresponding to pairs $(i, j) \notin E$.

We will show that $N'(G, \underline{\alpha}) \succeq 0$ in a somewhat strong sense, which will be enough to establish our final lower bound. Analogous to the submatrix H_{22}, one can consider the corresponding submatrix H'_{22} of N'. The expression for H'_{22} is as follows:

$$H'_{22} \stackrel{\text{def}}{=} H_{22} + \beta \cdot \sum_{i \in [n]} D_i (H_{22} - K) D_i,$$

Here D_i is the matrix with $(r_i^{\otimes 2})$ on the diagonal, and K is the matrix corresponding to the non-multilinear entries (entries corresponding to monomials like $x_a^2 x_b x_c$). For our final lower bound, we show how to use strong PSDness of H'_{22} to produce a witness for (4.3.1).

4.4 Degree 4 Lower Bound: Proof

In this section, we will demonstrate that $H'_{22} \succeq 0$ and use this to prove our final lower bound.

Parameters. Before we proceed further, it will be convenient to set α_2 and α_3 in terms of α_4. We choose

$$\alpha_3 = \alpha_4^{3/4} \cdot \tfrac{\sqrt{2}}{4} \qquad \alpha_2 = \alpha_4^{1/2} \cdot \tfrac{1}{8}\,. \tag{4.4.1}$$

For now, we leave α_4 unset (eventually we will choose it to be $\frac{1}{n^2 \log(n)^c}$ for some constant c).

Proving that $H'_{22} \succeq 0$

Here we will show that H'_{22} is PSD in a strong sense.

Theorem 4.4.1. *For $\beta = \frac{1}{100\sqrt{n}\log n}$ there is $\alpha_4 \geq \frac{1}{n^2 \log(n)^{O(1)}}$, so that the following holds with probability at least $1 - O(n^{-4})$,*

$$H'_{22} \succeq \frac{1}{8}\mathbb{E}[H_{22}] + \frac{\beta\lambda_0}{16} \cdot \Pi_W$$

Proof. Fix $\theta = \frac{16\|Q\|^2}{\lambda_1}$. By definition of H'_{22}, we have

$$H'_{22} = H_{22} + \beta \cdot \sum_{i\in[n]} D_i(H_{22} - K)D_i.$$

Define $P_W = \sum_{i\in[n]} r_i^{\otimes 2}(r_i^{\otimes 2})^\top$. We can apply Lemma 4.3.7 to the H_{22} term and Corollary Corollary 4.3.8 for $H_{22} - K$ (the reader may easily check that our choice of $\alpha_2, \alpha_3, \alpha_4$ satisfies their hypotheses) to obtain

$$H'_{22} \succeq \frac{1}{4}\mathbb{E}[H_{22}] - \theta \cdot \Pi_2\Pi_W\Pi_2 + \beta \cdot \sum_{i\in[n]} D_i\big(\frac{1}{2}\mathbb{E}[H_{22}] - \theta\Pi_2\Pi_W\Pi_2\big)D_i\,.$$

which, dropping Π_1, Π_2, becomes

$$\succeq \frac{1}{4}\mathbb{E}[H_{22}] - \theta \cdot \Pi_2\Pi_W\Pi_2 + \beta \cdot \sum_{i\in[n]} D_i\left(\frac{\lambda_0}{2}\Pi_0 - \theta\Pi_2\Pi_W\Pi_2\right)D_i\,.$$

which, using $D_i\Pi_0 D_i = r_i r_i^\top / \binom{n}{2}$, becomes

$$\succeq \frac{1}{4}\mathbb{E}[H_{22}] - \theta \cdot \Pi_2\Pi_W\Pi_2 + \beta\frac{\lambda_0}{4n^2}P_W - \beta\theta\sum_{i\in[n]} D_i\Pi_2\Pi_W\Pi_2 D_i\,. \tag{4.4.2}$$

Now we will appeal to a few matrix concentration results (whose proofs may be found in Section 4.5). First, with probability $1 - O(n^{-5})$, the vectors $\{r_i^{\otimes 2}\}$ are nearly orthogonal, and therefore form a well-conditioned basis for the subspace W.

Lemma 4.4.2. *If $P_W \stackrel{\text{def}}{=} \sum_i r_i^{\otimes 2}(r_i^{\otimes 2})^\top$ then with probability at least $1 - O(n^{-5})$,*

$$\tfrac{1}{n^2}(1 + o(1)) \cdot P_W \succeq \Pi_W \succeq (1 - o(1))\tfrac{1}{n^2} \cdot P_W\,.$$

Also, the vectors $\{r_i^{\otimes 2}\}$ have negligible projection on to the eigenspaces V_0, V_1,

Lemma 4.4.3. *Let $\Pi_{01} = \Pi_0 + \Pi_1$. With probability at least $1 - O(n^{-\omega(\log n)})$,*

$$\|\Pi_{01}\Pi_W\Pi_{01}\| \leq O\left(\frac{\log^2 n}{n}\right).$$

This implies that with overwhelming probability,

$$\Pi_W + O\left(\frac{\log^2 n}{n}\right) \cdot \mathrm{Id} \succeq \Pi_2 \Pi_W \Pi_2 \succeq \Pi_W - O\left(\frac{\log^2 n}{n}\right) \cdot \mathrm{Id}\ .$$

Finally, W has dimension n. Each $D_i \Pi_2 \Pi_W \Pi_2 D_i$ has only n non-zero singular values, each of which is $O(1)$. Moreover, multiplying on the left and right by D_i acts as a random linear transformation/ random change of basis. Intuitively, this suggests that $\sum_i D_i \Pi_2 \Pi_W \Pi_2 D_i$ has n^2 eigenvalues all of which are roughly $O(1)$. In fact,

Lemma 4.4.4. *With probability* $1 - O(n^{-5})$,

$$\sum_w D_w \Pi_2 \Pi_W \Pi_2 D_w \preceq O(n) \cdot \Pi_0 + O(\log^2 n) \cdot \mathrm{Id}_n\ .$$

Substituting these bounds into (4.4.2) we get,

$$H'_{22} \succeq \frac{1}{4}\,\mathbb{E}[H_{22}] + \left(\frac{\beta\lambda_0}{8} - \theta\right) \cdot \Pi_W - O\left(\frac{\theta \log^2 n}{n} + \beta\theta \log^2 n\right) \cdot \mathrm{Id} - \beta\theta \cdot O(n) \cdot \Pi_0\ .$$

By Lemma 4.3.5, with probability at least $1 - O(n^{-4})$, $\|Q\| \lesssim \alpha_4 n^{3/2} \log^3(n)$. Substituting this bound for $\theta = \frac{16\|Q\|^2}{\lambda_1}$ along with (4.3.4),(4.3.5),(4.3.6), finishes the proof for our choice of parameters in (4.4.1).

□

Putting Things Together

In this section we give our final lower bound for the sum-of-squares SDP (4.3.1). The following reduction uses standard techniques, appearing first in this context in [MPW15].

Lemma 4.4.5 (Corollary 2.4 in [MPW15])**.** *Suppose* $\mathcal{L}$ *is a real-valued linear function on* n*-variate real polynomials of degree at most* 4*, and suppose that* $\mathcal{L}$ *satisfies the constraint* $\{\sum_i x_i = \omega\}$ *for some* $\omega \in \mathbb{R}$*. If for every homogeneous degree-*2 *polynomial* p *it holds that* $\mathcal{L}p^2 \geq 0$*, then for every* p *of degree at most* 2 *we have* $\mathcal{L}p^2 \geq 0$*.*

With this in mind, we define an operator $\tilde{\mathbb{E}}$ on polynomials of degree at most 4 as follows.

Definition 4.4.6. Let $G \sim G(n, 1/2)$. For a multilinear degree-4 monomial x_S where S is a clique in G, define $\tilde{\mathbb{E}}\, x_S = H'_{2,2}(S_1, S_2)$ for any pair $S_1, S_2 \subseteq [n]$ of size 2 such that $S = S_1 \cup S_2$. If S is not a clique in G, set $\tilde{\mathbb{E}}\, x_S = 0$. Let ω' be the greatest real solution to the following equation

$$\sum_{S \text{ a 4-clique in } G} \tilde{\mathbb{E}}\, x_S = \binom{\omega'}{4}\ .$$

For every $S \subseteq [n]$ of size $0 \leq s \leq 3$, define

$$\tilde{\mathbb{E}}\, x_S = \frac{1}{\omega' - s} \sum_{\ell \in [n] \setminus S} \tilde{\mathbb{E}}\, x_{S \cup \ell}\ .$$

Extend $\tilde{\mathbb{E}}$ to non-multilinear polynomials to satisfy the constraints $\{x_i^2 = x_i\}$ for all i.

Now we state our main theorem of this section.

Theorem 4.4.7 (Final restatement of Theorem 4.3.1)**.** *With probability $1 - O(n^{-4})$ over $G \sim G(n, 1/2)$, the operator $\tilde{\mathbb{E}}$ satisfies the linear constraints 4.3.1 and has $\tilde{\mathbb{E}}\, p^2 \geq 0$ for every p of degree at most 2, for some $\omega' = \Omega(\sqrt{n}/\log(n)^{O(1)})$.*

By construction, $\tilde{\mathbb{E}}$ satisfies the constraints $x_i = x_i^2$ and $\tilde{\mathbb{E}}\, x_S = 0$ for non-cliques S of (4.3.1). The next lemma shows that $\tilde{\mathbb{E}}\, 1 = 1$; the proof that $\tilde{\mathbb{E}}$ satisfies $\sum_{i \leq n} x_i = \omega'$ for ω' as in the definition above is similar.

Lemma 4.4.8. $\tilde{\mathbb{E}}\, 1 = 1$.

Proof. We expand $\tilde{\mathbb{E}}\, 1$ as

$$\tilde{\mathbb{E}}\, 1 = \tfrac{1}{\omega'(\omega'-1)(\omega'-2)(\omega'-3)} \sum_i \sum_{j \neq i} \sum_{k \neq i,j} \sum_{\ell \neq i,j,k} \tilde{\mathbb{E}}\, x_i x_j x_k x_\ell \,.$$

Recalling that ω' is the solution to $\binom{\omega'}{4} = \omega'(\omega'-1)(\omega'-2)(\omega'-3)/24 = \sum_{S \text{ a 4-clique in } G} \tilde{\mathbb{E}}\, x_S$ reveals that the above must equal 1. □

We will need the following concentration result from [MPW15] (similar results appear elsewhere in the literature, see [MPW15] for references), which bounds the deviation from expectation of the number of 4-cliques in G containing a particular smaller clique.

Theorem 4.4.9 (Special Case of Theorem 10.1 in [MPW15])**.** *For $I \subseteq [n]$ of size $i \leq 4$, let $N_4(i)$ be the number of 4-cliques in $G \sim G(n, 1/2)$ containing I. For large enough n,*

$$\mathbb{P}\left\{ \left| N_4(I) - \frac{2^{\binom{i}{2}}}{64} \cdot \binom{n-i}{4-i} \right| > 200 n^{4-i-1/2} \log(n) \;\middle|\; I \text{ is a clique} \right\} \leq O(n^{-10}) \,.$$

The next lemma shows that the objective value ω' is $\widetilde{\Omega}(\sqrt{n}/\operatorname{polylog} n)$.

Lemma 4.4.10. *If $\beta \ll n^{3/2-\delta}$ for some constant δ, then for $G \sim G(n, 1/2)$ with high probability $\omega' = \alpha_4^{1/4}(\frac{n}{2\sqrt{2}} \pm \sqrt{n}\log(n)^{O(1)})$.*

Proof. Let N_4 be the number of 4-cliques in G. Then by Theorem 4.4.9, with probability $1 - O(n^{-10})$,

$$N_4 = \tfrac{n^4}{1536}\left(1 \pm O(\log n/\sqrt{n})\right) \,.$$

Thus for large enough n, by standard concentration (e.g. the method of bounded differences) applied to the correction term,

$$\sum_{S \text{ a 4-clique in } G} \tilde{\mathbb{E}}\, x_S = \alpha_4 N_4 + \alpha_4 \beta \sum_{S \text{ a 4-clique in } G} \sum_{i \in n} \prod_{j \in S} r_i(j)$$

$$= \alpha_4 N_4 \pm \widetilde{O}(\alpha_4 \cdot \beta \cdot n^{5/2})$$
$$= \tfrac{\alpha_4 n^4}{1536}(1 \pm O(\log n/\sqrt{n}))\,.$$

Thus, $\omega' = \alpha_4^{1/4}(n/(2\sqrt{2}) \pm \sqrt{n}\log(n)^{O(1)})$. □

To prove Theorem 4.4.7 we just have to ensure that $\tilde{\mathbb{E}}\, p^2 \geq 0$ for all degree-2 polynomials p. In light of Lemma 4.4.5, we can restrict attention to homogeneous degree-2 polynomials. Let $\mathcal{H} \in \mathbb{R}^{\binom{n}{2}\times\binom{n}{2}}$ be the matrix with

$$\mathcal{H}(S_1, S_2) = \begin{cases} \tilde{\mathbb{E}}\, x_{S_1} x_{S_2} \text{ if } S_1 \cup S_2 \text{ is a clique in } G \\ H'_{2,2}(S_1, S_2) \text{ otherwise}\,. \end{cases}$$

To show that $\tilde{\mathbb{E}}\, p^2 \geq 0$ for degree-2 homogeneous polynomials, it will be enough to show that $\mathcal{H} \succeq 0$.

Let $\Delta = \mathcal{H} - H'_{2,2}$, so that $\mathcal{H} = H'_{2,2} + \Delta$. The last lemma we need is a bound on Δ.

Lemma 4.4.11. *$\|\Delta\| \leq \alpha_3 \cdot O(\log(n)^{O(1)}\sqrt{n}) + \alpha_2 \cdot O(\log(n)^{O(1)}/\sqrt{n})$, with probability $1 - O(n^{-5})$,*

Before we prove Lemma 4.4.11, we prove Theorem 4.4.7.

Proof of Theorem 4.4.7. Lemma 4.4.8 shows that the constraints 4.3.1 are satisfied by $\tilde{\mathbb{E}}$. As long as $\alpha \geq \frac{1}{n^2 \log(n)^c}$ for some c, by Lemma 4.4.10 the operator $\tilde{\mathbb{E}}$ shows the objective value of (4.3.1) is at least $\sqrt{n}/\log(n)^{c'}$ for a constant c'. By Lemma 4.4.5, it is enough to show that $\mathcal{H} \succeq 0$ (since then the submatrix of rows and columns indexed by cliques in G is also PSD, implying that $\tilde{\mathbb{E}}\, p^2 \geq 0$ for all degree-2 polynomials p). By Theorem 4.4.1 and our choice of parameters, with probability at least $1 - O(n^{-4})$,

$$H'_{2,2} \succeq \frac{\alpha_2}{16} \cdot \mathrm{Id}$$

where Id is the identity matrix. By a union bound, with probability $1 - O(n^{-4})$, it also occurs that $\|\Delta\| \leq \alpha_2/16$ (so long as $\alpha_4 \leq \frac{1}{n^2 \log(n)^{O(1)}}$). In that case, $\mathcal{H} \succeq 0$ as desired. □

It remains to prove Lemma 4.4.11.

Proof of Lemma 4.4.11. By definition, $\Delta(S_1, S_2) = 0$ unless $|S_1 \cup S_2| \in \{2, 3\}$ and $S_1 \cup S_2$ is a clique in G. Let $\Delta_i \in \mathbb{R}^{\binom{n}{2}\times\binom{n}{2}}$ for $i = 2, 3$ be given by

$$\Delta_i(S_1, S_2) = \begin{cases} \Delta(S_1, S_2) \text{ if } |S_1 \cup S_2| = i \\ 0 \text{ otherwise.} \end{cases}$$

We start by analyzing Δ_3. By Theorem 4.4.9 and standard concentration (using either the moment method or the method of bounded differences), every nonzero entry satisfies

$$\Delta_3(S_1, S_2) = \tilde{\mathbb{E}}\, x_{S_1} x_{S_2} - \alpha_3 = \tfrac{1}{\alpha_4^{1/4}(n \pm \sqrt{n}\log(n)^{O(1)})} \cdot \alpha_4 \cdot \left(\tfrac{n}{8} \pm O(\sqrt{n}\log(n))\right) - \alpha_3$$

with probability at least $1 - O(n^{-7})$. Our choice of α_3 ensures that this is at most $O(\alpha_3 \log(n)^{O(1)}/\sqrt{n})$ in magnitude. There are at most n nonzero entries in any row of Δ_3, so taking a union bound, with probability at most $1 - O(n^{-5})$, we get

$$\|\Delta_3\| \leq O(\alpha_3 \log(n)^{O(1)} \sqrt{n}) \,.$$

The matrix Δ_2 is diagonal, and once again by Theorem 4.4.9 and standard concentration, each nonzero entry satisfies

$$|\Delta_2(S,S)| = |\,\tilde{\mathbb{E}}\, x_S - \alpha_2| = |\alpha_2(1 \pm O(\log(n)^{O(1)}/\sqrt{n}))| - \alpha_2 \leq \alpha_2 \cdot O(\log(n)^{O(1)}/\sqrt{n})$$

with probability at least $1 - O(n^{-7})$. By a union bound, with probability at least $1 - O(n^{-5})$, we get that $\|\Delta_2\| \leq \alpha_2 \cdot O(\log(n)^{O(1)}/\sqrt{n})$. This concludes the proof. □

4.5 Concentration for Locally Random Matrices over $G(n, \frac{1}{2})$

The goal of this section is to prove strong concentration bounds for the matrices will encounter in our analysis. Some of these bounds are established first in [DM15b], but for completeness we prove all the bounds we need here. We will begin by establishing a set of theorems which will give us concentration bounds for a large subset of the matrices we encounter, then apply these theorems to obtain our bounds.

General Approach and Tools

Our main tool for bounding the spectral norm of random matrices is once again the trace power method (Proposition 3.2.4). It reduces the question "what is the typical spectral norm of the random matrix X?" to combinatorial questions about certain labeled graphs associated to X. Because the bounds we require may be loose by poly-logarithmic factors, the combinatorics that usually underlie the trace power method will be relatively simple in our cases.

Fact 4.5.1. *Suppose an $n \times n$ random matrix M satisfies $\mathbb{E}[\mathrm{Tr}(M^k)] \leq n^{\alpha k + \beta} \cdot (\gamma k)!$ for any even integer k, where α, β, γ are constants. Then*

$$\mathbb{P}\left(\|M\| \lesssim n^{\alpha} \cdot \log \frac{n}{\eta}\right) \geq 1 - \eta \,.$$

Proof. The proof follows from an application of Proposition 3.2.4. We have that for even k,

$$\mathbb{P}[\|M\| \geq t] \;\leq\; \frac{1}{t^k}\,\mathbb{E}[\mathrm{Tr}(M^k)] \;\leq\; \frac{\sqrt{\pi\gamma k}}{t^k}\left(\frac{\gamma k}{e}\right)^k n^{\alpha k + \beta} \,,$$

where we have applied Stirling's approximation in the last step. Choosing $k = O(\log n - \log \eta)$ and $t = O\left(k \cdot n^{\alpha}\right)$ completes the proof. □

The expression $\mathbb{E}[\mathrm{Tr}(X^k)]$ is a sum over products along closed paths of length k in the entries of X. In our case, the entries of the random matrix X are themselves low-degree polynomials in random variables $\{A_{ij}\}_{i\in[n],j\in[n]}$ where A_{ij} is the centered random variable that indicates whether the edge (i,j) is part of the random graph G. (That is, $A_{ij} = 1$ if $(i,j) \in G$ and -1 otherwise.) Thus $\mathrm{Tr}(X^k)$ can be written out as a polynomial in the random variables $\{A_{ij}\}_{i,j\in[n]}$. Since the random variables $\{A_{ij}\}_{i,j\in[n]}$ are centered (i.e., $\mathbb{E}[A_{ij}] = 0$), almost all of the terms in $\mathbb{E}[\mathrm{Tr}(X^k)]$ vanish to zero. The nonzero terms are precisely those monomials in which every variable appears with even multiplicity.

Graph-Theoretic Definitions and Lemmas

In this section, we set up some notation and definitions helpful in our proofs of the main results of this section.

For the purpose of moment calculations, we borrow much of our terminology from the work of Deshpande and Montanari [DM15b]. Every monomial in random variables $\{A_{ij}\}_{i,j\in[n]}$ corresponds to a *labeled graph* $(R = (V,E), F)$ that consists of a graph $R = (V,E)$ and a labeling $F : V \to [n]$ that maps its vertices to $[n]$. A labeling of R *contributes* (is nonzero in expectation), if and only if every pair $\{i,j\}$ appears an even number of times as a label of an edge in R.

The following definitions and notation are generalizations of the ones used in [DM15b] to general degrees d and are useful in the proof of our first main norm bound for random matrices, Lemma 4.5.7. (The generalization to higher degrees d is useful in establishing the improved bounds on the simple moments in the preprint [HKP15], and provides evidence that obtaining norm bounds is not the main roadblock to obtaining an $\approx \sqrt{n}$ lower bound for higher degrees d.)

Definition 4.5.2. Let U be a bipartite graph on vertices $\{1,2,\ldots,d\} \times \{1',2',\ldots,d'\}$. A *$U$-ribbon* of length 2ℓ is a graph R on $2\ell d$ vertices

$$a_1^1, \ldots, a_d^1, \ldots, a_1^\ell, \ldots, a_d^\ell$$
$$b_{1'}^1, \ldots, b_{d'}^1, \ldots, b_{1'}^\ell, \ldots, b_{d'}^\ell .$$

We install edges in R by placing a copy of U on vertices $1,2,\ldots,d$ and $1',2',\ldots,d'$ (with the label i or i' matching the upper index of as and bs respectively) on $a_1^i, \ldots, a_d^i, b_1^{i-1}, \ldots, b_d^{i-1}$ for every $i \le d$. We also place a copy of the mirror image of U on $a_1^i, \ldots, a_d^i, b_1^i, \ldots, b_d^i$ For $i = 0$, we treat $i-1$ as d (modular addition), so that the graph wraps around. Often we will omit the length parameter 2ℓ when it is clear from context.

Definition 4.5.3. Let G be a graph. A *labeled* U-ribbon R is a tuple (R,F) where R is a U-ribbon and $F : R \to G$ is a map labeling each vertex of R with a vertex in G. We require that for (u,v) an edge in R, $F(u) \neq F(v)$.

Definition 4.5.4. Let (R,F) be a labeled U-ribbon where U has $2d$ vertices. We say (R,F) is *disjoint* if for every i,

$$|\{F(a_1^i), \ldots, F(a_d^i), F(b_1^i), \ldots, F(b_d^i)\}| = |\{F(a_1^i), \ldots, F(a_d^i), F(b_1^{i-1}), \ldots, F(b_d^{i-1})\}| = 2d\,.$$

Definition 4.5.5. Let (R, F) be a labeled U-ribbon where U has $2d$ vertices. We say that (R, F) is *contributing* if no element of the multiset $\{(F(u), F(v)) : (u, v) \in R\}$ occurs with odd multiplicity.

The following combinatorial lemma will serve as a tool in the proofs of the main results for this section.

Lemma 4.5.6. *Let (R, F) be a contributing labeled U-ribbon of length 2ℓ. Recall that R has vertex set a^i_j, b^i_j for $i \in \ell$ and $j \in [d]$. Let $k \leq d$. Suppose that the sets*

$$\{F(a^i_1), F(b^i_1)\}_{i\in[\ell]}, \ldots, \{F(a^i_k), F(b^i_k)\}_{i\in[\ell]}, \quad \{F(a^i_j), F(b^i_j)\}_{i\in[\ell], j\in[k+1,d]}$$

are disjoint. Then if U contains the edges $\{(1,1), \ldots, (k,k)\}$ (where we identify the vertex set of U with $[d] \times [d]$), $\{F(u) : u \in R\}$ has size at most $(2d-k)\ell + k$.

Proof. The assumption on U implies that R contains the cycles

$$\begin{aligned} C_1 &\stackrel{\text{def}}{=} (a^1_1, b^1_1, \ldots, b^\ell_1, a^1_1) \\ &\ldots \\ C_k &\stackrel{\text{def}}{=} (a^1_k, b^1_k, \ldots, b^\ell_k, a^1_k)\,. \end{aligned}$$

In order for (R, F) to be contributing, every edge $(u, v) \in R$ must have a partner $(u', v') \neq (u, v)$ so that $F(u') = F(u)$ and $F(v') = F(v)$. By our disjointness assumption, every edge in cycle C_i must be partnered with another edge in C_i. Thus, now temporarily identifying edges when they are labeled identically, each C_i is a connected graph with at most ℓ unique edges (since each of the 2ℓ edges must be partnered). It therefore has at most $\ell + 1$ unique vertex labels. Among the cycles $C_1, \ldots, C_k$, there are thus at most $k(\ell + 1)$ unique vertex labels. In the rest of the ribbon R there can be at most $2\ell(d - k)$ unique vertex labels, because once the cycles $C_1, \ldots, C_k$ are removed there are only that many vertices left in R. So in total there are at most $k(\ell + 1) + 2\ell(d - k) = (2d - k)\ell + k$ unique labels. □

Concentration of Bipartite-Patterned Matrices

For a large class of matrices that we encounter, we will be able to prove a general spectral norm bound. Let U be a bipartite graph with bipartitions Q, P, each of size d, and with edge set $E(U)$. Let $G \sim G(n, \frac{1}{2})$, and define the $\binom{[n]}{d} \times \binom{[n]}{d}$ matrix $M = M_U(G)$ so that for $I, J \in \binom{[n]}{d}$,

$$M(I, J) = \prod_{(i,j)\in E(U)} A_{ij}\,,$$

where the indices of I have been identified with the vertices of Q and the indices of J have been identified with the vertices of P. We shall call such an M a bipartite-patterned matrix. For bipartite-patterned matrices, we can obtain a spectral norm bound based on characteristics of the bipartite graph U:

Lemma 4.5.7. *For $d \geq 2$, $d = O(1)$, and a bipartite graph U with bi-partitions of size d, let $M = M_U \in \mathbb{R}^{\binom{[n]}{d} \times \binom{[n]}{d}}$ be a bipartite patterned matrix such that for any $I, J \in \binom{[n]}{d}$,*

$$M(I,J) = \begin{cases} \prod_{(i,j)\in E(B)} A_{ij} & \text{if } I \cap J = \emptyset \\ 0 & \text{otherwise} \end{cases},$$

Then:

1. *When U contains a 2-matching, then $\mathbb{P}(\|M\| \geq O(n^{d-1}(\log n)^2)) \leq O(n^{-10})$.*
2. *When U is not the empty graph, $\mathbb{P}(\|M\| \geq O(n^{d-1/2}(\log n)^2)) \leq O(n^{-10})$.*

The proof of Lemma 4.5.7 is similar to the proofs via the trace power method for bounding the norms of matrices as presented in [DM15b]. The general format we present here will come in handy for multiple applications to various matrices which appear in Section 4.4 and in Section 4.3.

Useful Tools

We begin by presenting some general-purpose tools that we will employ in our analysis to reduce bounding $\|M\|$ to bounding simpler patterned matrices.

For analyzing the spectral norm of a matrix $M \in \mathbb{R}^{\binom{[n]}{d} \times \binom{[n]}{d}}$, the first tool allows us to analyze instead a related matrix $M' \in \mathbb{R}^{n^d \times n^d}$. That is, instead of rows and columns being indexed by subsets of vertices as in M, M' has rows and columns indexed by ordered tuples of vertices of size d. This transformation is not hard as one can find M as a principal submatrix of M'.

Lemma 4.5.8 (Sets to Ordered Tuples)**.** *For any $M \in \mathbb{R}^{\binom{[n]}{d} \times \binom{[n]}{d}}$ define the matrix $M' \in \mathbb{R}^{n^d \times n^d}$ such that for any ordered tuple $S = (a_1, a_2, \ldots, a_d), T = (b_1, b_2, \ldots, b_d) \in [n]^d$, $M'(S,T) = M(\{a_1, a_2, \ldots, a_d\}, \{b_1, b_2, \ldots, b_d\})$. Then, $\|M\| \leq \|M'\|$.*

Proof. It is enough to show that M' occurs as a principal submatrix of M. For this, take the submatrix of rows and columns of M indexed by tuples $(a_1, \ldots, a_d)$ in sorted order, i.e., with $a_1 \leq a_2 \leq \ldots a_d$. □

We will use the following lemma to break dependencies in certain random matrices by decomposing them into matrices whose entries, while still dependent, have additional structure.

Lemma 4.5.9 (Random Partitioning)**.** *For $d \in \mathbb{N}$, let $M \in \mathbb{R}^{n^d \times n^d}$ be indexed by subsets $I, J \in [n]^d$. Suppose $M(I,J) = 0$ when $I \cap J \neq \emptyset$. Let $(S_1^1, \ldots, S_k^1), \ldots, (S_1^r, \ldots, S_k^r)$ be a sequence of partitions of $[n]$ into k bins. Each partition induces a matrix based on M as follows: for indices $A, B \in [n]^d$ with $A = (a_1, \ldots, a_d)$ and $B = (b_1, \ldots, b_d)$,*

$$M_i(A,B) = \begin{cases} M(A,B) & \text{if } a_j, b_j \in S_j^i \text{ for } j < k, \text{ and } a_j, b_j \in S_k \text{ for } j \geq k, \\ & \text{and for all } i' < i,\ M_{i'}(A,B) = 0\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Then, there is a family of partitions $(S_1^1, \dots, S_k^1), \dots, (S_1^r, \dots, S_k^r)$ *such that* $M = \sum_{i=1}^r M_i$ *with* $r \le k^{O(d)} \log n$.

Proof of Lemma 4.5.9. For r to be chosen later, we pick partitions $(S_1^1, \dots, S_k^1), \dots, (S_1^r, \dots, S_k^r)$ uniformly at random and independently so that each is partition of $[n]$ into sets of size n/k each.

Call A, B *good* at step i if $a_j, b_j \in S_j^i$ for every $j < k$ and $a_j, b_j \in S_k^i$ if $j \ge k$. It is enough to show that after $r \le O(k^k \log n)$ steps the probability that every set $A \cup B$ of size $2d$ is good at some step $i \le r$.

Fix some A, B with $|A \cup B| = 2d$. It is good at step i with probability at least k^{-2d}. Since the steps are independent, after r steps

$$\begin{aligned}\mathbb{P}(A, B \text{ is good}) &\ge (1 - \tfrac{1}{k^{2d}})^r \\ &= ((1 - \tfrac{1}{k^{2d}})^{k^{2d}})^{r/k^{2d}} \\ &\le (\tfrac{1}{e})^{r/k^{2d}}\end{aligned}$$

which is at most $1/n^{10d}$ for some $r = O(k^{2d} \log n)$.

Taking a union bound over all $O(n^{2d})$ tuples A, B with $|A \cup B| = 2d$ completes the proof. □

Proof of Lemma 4.5.7

Proof of Lemma 4.5.7. By Lemma 4.5.8 it is enough to prove the analogous claims for the $n^d \times n^d$ matrix M with entries given by, for any multisets of indices $A = (a_1, \dots, a_d), B = (b_1, \dots, b_d) \in [n]^d$,

$$M(A, B) = \begin{cases} \prod_{(a_i, b_j) \in E(U)} x_{a_i, b_j} & \text{if } |A \cup B| = 2d \\ 0 & \text{otherwise} \end{cases},$$

By multiplying M by suitable permutation matrices P, P' to give PMP', we may assume in the 2-matching case above that the matching is $\{(1,1), (2,2)\}$ and in the nonempty graph case that the edge contained is $(1,1)$ (where we think of the vertex set of B as $[d] \times [d]$). Note that $\|M\| = \|PMP'\|$.

We apply Lemma 4.5.9 to obtain a family of matrices $\{M_i\}_{i \in [r]}$ for some $r = O(\log n)$ satisfying $M = \sum_i M_i$. On any entry (A, B) on which M_i is nonzero it is equal to M at that entry, and furthermore for each M_i there is a partition (S_1^i, S_2^i, S_3^i) of $[n]$ so that if $M_i(A, B) \ne 0$ then $a_1, b_1 \in S_1^i, a_2, b_2 \in S_2^i$, and $a_j, b_j \in S_3^i$ for all $j > 2$.

We show that every matrix $\|M_i\|$ has bounded spectral norm. To save on indices, let $N = M_i$. Let (S_1, S_2, S_3) be the partition of $[n]$ corresponding to N. We bound $\mathbb{E} \operatorname{Tr}(NN^\top)^\ell$ for some ℓ to be chosen later.

Let $\mathcal{R}(N)$ be the set of contributing disjoint labeled U-ribbons (R, F) of length 2ℓ with $F(a_1^i), F(b_1^i) \in S_1, F(a_2^i), F(b_2^i) \in S_2$ and $F(a_j^i), F(b_j^i) \in S_3$ for $j > 2$. Then $\mathbb{E} \operatorname{Tr}(NN^\top)^\ell \le O(\ell^\ell)|\mathcal{R}(N)|$. (Here we have an inequality rather than an equality because some elements of $\mathcal{R}(N)$ may correspond to entries of N which are zero because they appeared in some other part of the partitioning scheme and $\ell^\ell \ge \ell!$ accounts for reorderings of the labels.)

Supposing that U contains a 2-matching, by Lemma 4.5.6, each $(R, F) \in \mathcal{R}(N)$ contains at most $(2d-2)\ell + 2$ unique $\{F(u) : u \in R\}$. So there are at most $n^{2\ell(d-1)+2}$ elements of $\mathcal{R}(N)$.

It follows that $\mathbb{E}[\mathrm{Tr}(NN^\top)^\ell] \leq (2d\ell)! n^{2\ell(d-1)+2}$. Now by Fact 4.5.1,

$$\mathbb{P}(\|N\| \lesssim n^{(d-1)} \cdot \log \frac{n}{\eta}) \geq 1 - \eta,$$

And applying the triangle inequality to $\|M\| = \|\sum_i M_i\|$ and taking a union bound over the $c \cdot \log n = O(\log n)$ matrices M_i, we get that

$$\mathbb{P}(\|M\| \lesssim \log n \cdot n^{(d-1)} \cdot \log \frac{n}{\eta}) \geq 1 - c \cdot \log n \cdot \eta,$$

and setting $\eta = (n^{10} \log n)^{-1}$, we have our desired result.

The case that B contains only a 1-matching is similar, replacing the $(2d-2)\ell + 2$ unique vertices in a contributing B-ribbon with $(2d-1)\ell + 1$, again by Lemma 4.5.6. □

Concentration Bounds for Relevant Matrices

In this section, we give bounds on the spectra of the specific random matrices that appear in our proofs. Several of our bounds will employ the tools developed in the previous subsection. For others, we will obtain our bounds by employing the trace power method (which we also used in the previous section's proofs).

Lemma 4.5.10 (Adapted from Propositions 4.20, 4.25 in [DM15b], Restatement of Lemma 4.3.5)**.** *With probability at least $1 - O(n^{-4})$, all of the following bounds hold:*

$$\|\Pi_a Q \Pi_b\| \lesssim \alpha_4 n^{3/2} \log^3(n) \qquad \forall (a,b) \in \{0,1,2\}^2 \tag{4.5.1}$$

$$\|\Pi_2 Q \Pi_2\| \lesssim \alpha_4 n \log^3(n) \tag{4.5.2}$$

$$\|K\| \lesssim \alpha_3 n^{1/2} \log^3(n) \tag{4.5.3}$$

Proof. Recall that Q is the matrix defined so that for any $I, J \in \binom{[n]}{2}$,

$$Q(I,J) = \begin{cases} 0 & \text{if } I \cap J \neq \emptyset \\ \alpha_4 \cdot \left(\prod_{i \in I, j \in J} \frac{(1+A_{ij})}{2} - \left(\frac{1}{2}\right)^4 \right) & \text{otherwise}. \end{cases}$$

We can further split Q into 15 matrices $Q_1, \ldots, Q_{15}$, one for each term in the expansion of the polynomial in the entries, $\left(\prod_{i \in I, j \in J} \frac{(1+A_{ij})}{2} - \left(\frac{1}{2}\right)^4 \right)$. (Notice that there will be no constant term.) For any $1 \leq i \leq 15$, the matrix $\frac{1}{\alpha_4} \cdot Q_i$ is equal to $\alpha_4 \cdot M_U$ as in Lemma 4.5.7, for some U which is not the empty graph. Furthermore, $\|\Pi_a\| \leq 1$ for all $a \in \{0,1,2\}$. Together, these imply that that $\|\Pi_a Q \Pi_b\| \leq \sum_{i=1}^{7} \|\Pi_a Q_i \Pi_b\| \leq O(\alpha_4 n^{3/2} \log^3 n)$ for all $a, b \in \{0,1,2\}$ with probability at least $1 - O(n^{-9})$, and this proves our first claim.

For the second claim, the strategy is somewhat subtle. Recall that Q is zero on entries (I, J) with nontrivial intersection $|I \cap J| > 0$. Such matrices do not interact well with the projector Π_2. Lemma 4.5.11 analyzes how such matrices behave under projection V_2.

Notice first that for any Q_i corresponding to M_U for U containing a 2-matching, we already have $\|\Pi_2 Q_i \Pi_2\| \le O(\alpha_4 n \log^3 n)$ with probability $1 - O(n^{-9})$ by Lemma 4.5.7. We just have to deal with the the Q_i's corresponding to U's with only a 1-matching. Let

$$Z = \sum_{i\,:\,\text{corresponding } U \text{ contains only a 1 matching}} Q_i \,.$$

Applying Lemma 4.5.11 to Z, we find that $\Pi_2 Z \Pi_2 = \Pi_2 D \Pi_2$ for a matrix D as in the lemma, which is nonzero only on non-multilinear entries. (Here there is an implicit change of scale by a multiplicative factor α_4 to apply the lemma.) We can now afford the bound $\|\Pi_2 D \Pi_2\| \le \|D\|$.

It is straightforward to check that on non-multilinear but non-diagonal entries, D is a rescaling of K. Our last claim about K can be used to finish the proof of the second claim.

Finally, we prove our last claim, about the matrix K. We split K into n matrices $K^{(1)}, \dots, K^{(4)}$, so that $K^{(1)}$ contains the nonzero entries of K where $|I \cap J| = 1$ and furthermore I and J intersect on their lexicographically first element. Similarly, $K^{(2)}$ contains those entries where $I \cap J = \{i\}$ and i is the lexicographically first element of I and the lexicographically second element of J, and so on.

Each $K^{(s)}$ we now further divide into $K^{(s)}_i = K_i$ for $i \in [n]$ (saving some indices), so that $K^{(s)}_i(I, J) = K^{(s)}(I, J)$ if $I \cap J = \{i\}$ and is 0 otherwise. Then $K^{(s)} = \sum_i K_i$ and the matrices K_i are nonzero on disjoint sets of rows and columns, so $\|K\| \le \max_i \|K_i\|$. Because H_{22} is built from the indicator of bipartite cliques in G, each K_i is a scaled copy of the centered adjacency matrix of $G \sim G(n, 1/2)$, so it follows that with probability $1 - O(n^{-9})$, each K_i satisfies $\|K_i\| \le \alpha_3 n^{1/2} \log(n)$. The lemma follows. □

Lemma 4.5.11. *Let A be any $n \times n$ matrix with ± 1 entries. Let $M \in \mathbb{R}^{\binom{[n]}{2} \times \binom{[n]}{2}}$ be the matrix such that for any $\{a, b\}, \{c, d\} \in \binom{[n]}{2}$ with $|\{a, b, c, d\}| = 4$,*

$$M(ab, cd) = A_{ad} + A_{ac} + A_{bc} + A_{bd} + A_{ac}A_{ad} + A_{bc}A_{bd} + A_{ad}A_{bd} + A_{ac}A_{bc} \,.$$

Let

$$\begin{aligned} R_1(ab, cd) &= A_{ac} + A_{ad} \\ C_1(ab, cd) &= A_{bc} + A_{bd} \\ C_2(ab, cd) &= A_{ac}A_{ad} + A_{bc}A_{bd} \\ R_2(ab, cd) &= A_{ac}A_{bc} + A_{ad}A_{bd} \\ D(ab, cd) &= -1 \cdot \mathbb{I}(|\{a, b, c, d\}| < 4) \cdot ((R_1 + R_2 + C_1 + C_2)(ab, cd)) \,. \end{aligned}$$

Then

$$\Pi_2 M \Pi_2 = \Pi_2 (M_{\le 3} + D) \Pi_2 \,,$$

where $M_{\le 3}$ is the portion of M corresponding to indices with repetition, i.e. $|\{a, b, c, d\}| \le 3$.

Proof. We will prove that such an M must have rows or columns spanned by $V_0 \cup V_1$. Clearly, $M_4 = R_1 + C_1 + C_2 + R_2 + D$, where M_4 is the matrix containing the multilinear entries of M. Our proof proceeds by noting that for each of these matrices, either the rows or columns are spanned by $V_0 \cup V_1$.

Consider the (a,b)th row of R_1–we can write the (c,d)th entry of the row as $R_1(ab,cd) = A_{ac} + A_{ad}$. Thus, if we define the vector $u \stackrel{\text{def}}{=} A_a \in \mathbb{R}^n$ to be equal to the ith row of A, we have that $R(ab,cd) = u_c + u_d$. It follows that the rows of R_1 are spanned by $V_0 \cup V_1$. The proof for C_1 is nearly identical.

Now, consider the (c,d)th column of C_2–here, define the vector $u \in \mathbb{R}^n$ so that $u_i = A_{ci}A_{d_i}$, then we have that $C_2(ab,cd) = u_a + u_b$. Thus, the (c,d)th column of C_2 is spanned by $V_0 \cup V_1$, as desired. Again, the proof for R_2 is almost identical. The conclusion follows. □

Lemma (Restatement of Lemma 4.4.2)**.** *If* $P_W \stackrel{\text{def}}{=} \sum_i r_i^{\otimes 2}(r_i^{\otimes 2})^\top$ *then with probability at least* $1 - O(n^{-5})$,

$$\tfrac{1}{n^2}(1+o(1)) \cdot P_W \succeq \Pi_W \succeq (1-o(1))\tfrac{1}{n^2} \cdot P_W \,.$$

Proof. For convenience, we write $a_i = r_i^{\otimes 2}$. Recall also that A is the ± 1 adjacency matrix of the graph G. By definition, the vectors $a_1, \ldots, a_n$ form a basis for the subspace W.

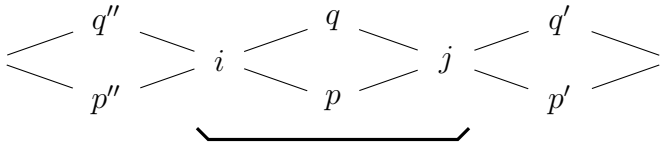
Let $\mathcal{A}$ be the matrix whose ith row is a_i. We will use matrix concentration to analyze the eigenvalues of $\mathcal{A}\mathcal{A}^T$, which are identical to the nonzero eigenvalues of $P_W = \mathcal{A}^T\mathcal{A}$.

The (i,j)th entry of $\mathcal{A}\mathcal{A}^T$ is $\langle a_i, a_j\rangle = \frac{1}{2}\langle A_i^{\otimes 2}, A_j^{\otimes 2}\rangle = \frac{1}{2}\langle A_i, A_j\rangle^2$. When $i = j$, this is precisely $\frac{1}{2}(n-1)^2$, and so $2\mathcal{A}\mathcal{A}^T = (n-1)^2 \cdot \mathrm{Id}_n + B$, where B is a matrix that is 0 on the diagonal and equal to $\langle A_i^{\otimes 2}, A_j^{\otimes 2}\rangle$ in the (i,j)th entry for $i \neq j$.

Let $M = B - \mathbb{E}[B] = B - (n-2)(J_n - \mathrm{Id}_n)$. We will use the trace power method to prove that $\|M\| = O(n^{3/2})$. The (i,j)th entry of M is given by 0 for $i = j$, and when $i \neq j$

$$M(i,j) = \langle A_i, A_j\rangle^2 - (n-2) = \left(\sum_{p,q} A_{ip}A_{iq}A_{jp}A_{jq}\right) - (n-2) = \sum_{p \neq q} A_{ip}A_{iq}A_{jp}A_{jq}.$$

The expression $\mathrm{Tr}(M^k)$ is a sum over monomial products over variables $\{A_{ip}\}_{i,p\in[n]}$, where each monomial product corresponds to a labeling $F : R \to [n]$ of a ribbon R. Each entry in M_{ij} corresponds to a sum over *links* of the ribbon R, where each link is a cycle of length 4, with the vertices i, j on opposite ends of the cycle, and the necessarily distinct vertices p, q are on the other opposite ends of a cycle. We will refer to i, j as the *center vertices* and p, q as the *peripheral* vertices of the link. Each edge (u, v) of the link is weighted by A_{uv}. Since $A_{ii} = 0$ for all $i \in [n]$, for every *contributing labeling*, it can never be the case that one of $p, q = i$. Each monomial product in the summation $\mathrm{Tr}(M^k)$ corresponds to a labeling $(F, \mathcal{L})$ of the graph F, where F is a cycle with k links. F has $4k$ edges, and in total it has $3k$ vertices.



The quantity $\mathrm{Tr}(M^k)$ is equal to the sum over all labelings of R. Taking the expectation, terms in $\mathbb{E}[\mathrm{Tr}(M^k)]$ which contain a variable A_{uv} with multiplicity 1 have expectation 0. Thus, $\mathbb{E}[\mathrm{Tr}(M^k)]$ is equal to the number of labelings of F in which every edge appears twice.

We prove that any such contributing labeling $F : R \to [n]$ has at most $3k/2 + 1$ unique vertex labels. We proceed by induction on k, the length of the cycle. In the base case, we have a cycle on two links; by inspection no such cycle can have more than 5 labels, and the base case holds.

Now, consider a cycle of length k. If every label appears twice, then we are done, since there are $3k$ vertices in F. Thus there must be a vertex that appears only once.

There can be no *peripheral* vertex whose label does not repeat, since the two center vertices neighboring a single peripheral vertex cannot have the same label in a contributing term, as $M(i, i) = 0$. Now, if there exists a *center vertex* i whose label does not repeat, it must be that there is a matching between its p, q neighbors so that every vertex is matched to a vertex of the same label. We identify these same-label vertices and remove i and two of its neighbors from the graph. This leaves us with a cycle of length $k - 1$, having removed at most one label from the graph. The induction hypothesis now applies, and we have a total of at most $3(k - 1)/2 + 2 \le 3k/2 + 1$ labels, as desired.

Thus, there are at most $3k/2 + 1$ unique labels in any contributing term of $\mathbb{E}[\mathrm{Tr}(M^k)]$. We may thus conclude that $\mathbb{E}[\mathrm{Tr}(M^k)] \le n^{3k/2+1} \cdot (3k/2 + 1)^{3k}$, and applying Fact 4.5.1, we have that $\|M\| \lesssim (n \log n)^{3/2}$ with probability at least $1 - O(n^{-5})$.

Therefore, $2\mathcal{A}\mathcal{A}^T = ((n - 1)^2 - n + 2)\,\mathrm{Id}_n + (n - 2)J_n + M$, and we may conclude that all eigenvalues of $\mathcal{A}\mathcal{A}^T$ are $(1 \pm o(1)) \cdot n^2$, which implies the same of $P_W = \mathcal{A}^T\mathcal{A}$. Since the range of P_W and Π_W is the same, we finally have that with probability $1 - o(1)$

$$(1 + o(1))/n^2 \cdot P_W \succeq \Pi_W \succeq (1 - o(1))/n^2 \cdot P_W,$$

as desired. □

The following proposition will be helpful in one of our norm bounds.

Proposition 4.5.12. *Let $F = (V, E)$ be a multigraph and let $\ell : V \to [n]$ be a labeling such that each pair (i, j) appears an even number of times as the label of an edge in E. Then,*

$$|\{\ell(v) | v \in V\}| \le \frac{|E|}{2} + (\#\ \textit{connected components of } F)$$

Proof. From F, we form a new graph F' by identifying all the nodes with the same label; thus, the number of nodes in F' is the number of labels in F. We then collapse the parallel edges in F' to form the graph H; since each labeled edge appears at least twice, the number of edges in H is at most half that in F. The number of nodes in H (and thus labels in F) is at most the number of edges in H plus the number of connected components; this is tight when H is a forest. Thus the number of distinct labels in F is at most $|E|/2 + c$, where c is the number of components in F. □

Lemma (Restatement of Lemma 4.4.3)**.** *Let $\Pi_{01} = \Pi_0 + \Pi_1$. With probability at least $1 - O(n^{-\omega(\log n)})$,*

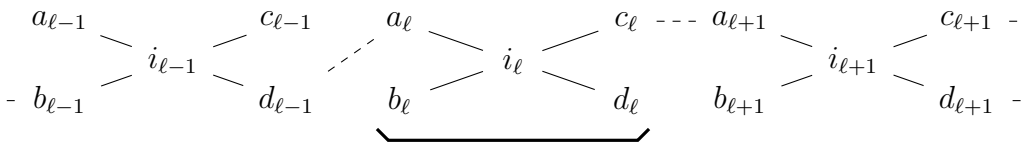
$$\|\Pi_{01}\Pi_W\Pi_{01}\| \le O\left(\frac{\log^2 n}{n}\right).$$

Proof. Call $M = \Pi_{01}\Pi_W\Pi_{01}$. We will apply the trace power method to M. By Lemma 4.5.13 and Lemma 4.4.2, we may exchange Π_W for $\frac{(1+o(1))}{n^2}\sum_i a_i a_i^T$ and Π_{01} for P_{01}, losing only constant factors. Letting $M'^k = (\frac{(1+o(1))}{n^2}P_{01}P_W)^k$, we have by the cyclic property of the trace that $\mathbb{E}[\mathrm{Tr}(M^k)] \le \mathrm{Tr}(M'^k)$.

We consider the expression for $\mathbb{E}[\mathrm{Tr}(M'^k)]$. Let a *chain* consist of a set of quadruples $\{a_\ell, b_\ell, c_\ell, d_\ell\}_{\ell\in[k]} \in [n]^4$ such that for each $\ell \in [k]$, we have $\|\{a_\ell, b_\ell\} \cap \{c_{\ell-1}, d_{\ell-1}\}| \ge 1$ (where we identify a_ℓ with $a_{\ell \bmod k}$). Let $\mathcal{C}_k$ denote the set of all chains of size k. We have that

$$\mathrm{Tr}(M^k) \le \mathrm{Tr}(M'^k) = \sum_{i_1,\dots,i_k} \sum_{\{a_\ell,b_\ell,c_\ell,d_\ell\}_{\ell\in[k]}\in\mathcal{C}_k} \prod_{\ell=1}^{k} \frac{1+o(1)}{n^2} \cdot r_\ell \cdot A_{i_\ell,a_\ell}A_{i_\ell,b_\ell}A_{i_\ell,c_\ell}A_{i_\ell,d_\ell},$$

where $r_\ell = \frac{1}{n-1}$ or $\frac{2}{n-1}$ depending on whether one or both of a_ℓ, b_ℓ are common with the following link in the chain. The quantity $\mathrm{Tr}(M^k)$ consists of a sum over cycles of k links. Each link is a star on 4 outer vertices $a_\ell, b_\ell, c_\ell, d_\ell$ with center vertex i_ℓ, and the non-central vertices of the link must have at least one vertex in common with the next link. So, each link has 4 edges and the cycle is a connected graph. See the figure below for an illustration (dashed lines indicate vertex equality, and are not edges).



Each term in the product has a factor of at most $\frac{2(1+o(1))}{n^3}$, due to the scaling of the entries of P_{01} and P_W. Thus we have

$$\mathbb{E}[\mathrm{Tr}(M'^k)] \le \left(\frac{3}{n^3}\right)^k \sum_{i_1,\dots,i_k} \sum_{\{a_\ell,b_\ell,c_\ell,d_\ell\}_{\ell\in[k]}\in\mathcal{C}_n} \mathbb{E}\left[\prod_{\ell=1}^{k} A_{i_\ell,a_\ell}A_{i_\ell,b_\ell}A_{i_\ell,c_\ell}A_{i_\ell,d_\ell}\right].$$

The only contributing terms correspond to those for which every edge variable in the product has even multiplicity. Each contributing term is a connected graph and has $4k$ edges and at most $5k$ vertices where every labeled edge appears twice, so we may apply Proposition 4.5.12 to conclude that there are at most $2k+1$ labels in any such cycle. We thus have that

$$\mathbb{E}[\mathrm{Tr}(M'^k)] \le \left(\frac{3}{n^3}\right)^k \cdot n^{2k+1} \cdot (5k)!,$$

and applying Fact 4.5.1, we conclude that $\|M\| \lesssim \frac{\log^2 n}{n}$ with probability $1 - O(n^{-\omega(\log n)})$, as desired. □

The following lemma allows us to approximate the projector to $V_0 \cup V_1$by a matrix that is easy to describe; we will use this matrix as an approximation to the projector in later proofs.

Lemma 4.5.13. *Let Π_{01} be the projection to the vector space $V_0 \cup V_1$. Let $P_{01} \in \mathbb{R}^{\binom{n}{2},\binom{n}{2}}$ be a matrix defined as follows:*

$$P_{01}(ab,cd) = \begin{cases} \frac{2}{n-1} & |\{a,b,c,d\}| = 2 \\ \frac{1}{n-1} & |\{a,b,c,d\}| = 3 \\ 0 & |\{a,b,c,d\}| = 4. \end{cases}$$

Then

$$2\Pi_{01} \succeq P_{01} \succeq (\tfrac{n-2}{n-1}) \cdot \Pi_{01},$$

Proof. We will write down a basis for $V_0 \cup V_1$, take a summation over its outer products, and then argue that this summation approximates Π_{01}. The vectors $v_1, \ldots, v_n \in \mathbb{R}^{\binom{n}{2}}$ are a basis for $V_1 \cup V_0$:

$$v_i(a,b) = \begin{cases} \frac{1}{\sqrt{n-1}} & \{a,b\} = \{i,\cdot\} \\ 0 & \text{otherwise.} \end{cases}$$

For any two v_i, v_j with $i \neq j$, we have $\langle v_i, v_j \rangle = \frac{1}{n-1}$. Let $U \in \mathbb{R}^{n^2 \times n}$ be the matrix whose ith column is given by v_i. Notice that the eigenvalues of $\sum_i v_i v_i^T = UU^T$ are equal to the eigenvalues of U^TU, and that $U^TU = \frac{1}{n-1} J_n + \frac{n-2}{n-1} \mathrm{Id}_n$. Therefore, as both Π_{01} and $UU^\top$ have the same column and row spaces,

$$2\Pi_{01} \succeq \sum_i v_i v_i^T \succeq \tfrac{n-2}{n-1} \Pi_{01},$$

Now, let $P_{01} = \sum_i v_i v_i^T$; we can explicitly calculate the entries of P_{01},

$$P_{01}(ab,cd) = \begin{cases} \frac{2}{n-1} & |\{a,b,c,d\}| = 2 \\ \frac{1}{n-1} & |\{a,b,c,d\}| = 3 \\ 0 & |\{a,b,c,d\}| = 4. \end{cases}$$

The conclusion follows. □

Lemma (Restatement of Lemma 4.4.4)**.** *With probability $1 - O(n^{-5})$,*

$$\sum_w D_w \Pi_2 \Pi_W \Pi_2 D_w \preceq O(n) \cdot \Pi_0 + O(\log^2 n) \cdot \mathrm{Id}_n \,.$$

Proof. We begin by replacing Π_2 with $(1 - \Pi_{01})$, as by Lemma 4.4.3, Π_{01} can be replaced by P_{01} which has a convenient form. For any vector $x \in \mathbb{R}^{\binom{n}{2}}$,

$$x^T \left(\sum_i D_i \Pi_2 \Pi_W \Pi_2 D_i \right) x = x^T \left(\sum_i D_i \Pi_W D_i \right) x - 2x^T \left(\sum_i D_i \Pi_{01} \Pi_W D_i \right) x + x^T \left(\sum_i D_i \Pi_{01} \Pi_W \Pi_{01} D_i \right) x$$

$$\leq \sum_i \left(\|\Pi_W D_i x\|^2 + 2\|\Pi_W \Pi_{01} D_i x\| \cdot \|\Pi_W D_i x\| + \|\Pi_W \Pi_{01} D_i x\|^2 \right)$$
$$\leq 2x^T \left(\sum_i D_i \Pi_W D_i \right) x + 2 \left(\sum_i (D_i x)^T \Pi_{01} \Pi_W \Pi_{01} D_i x \right)$$
$$\leq 2x^T \left(\sum_i D_i \Pi_W D_i \right) x + 2n \, \|\Pi_{01} \Pi_W \Pi_{01}\| \cdot \|x\|^2,$$

where to obtain the second line we have applied Cauchy-Schwarz, to obtain the third line we have used the fact that $a^2 + b^2 \geq 2ab$, and to obtain the final line we have used the fact that $\|D_i x\| = \|x\|$.

Now, the second term is $O(\log^2 n) \cdot \|x\|^2$ with overwhelming probability by Lemma 4.4.3. It remains to bound the first term. To this end, we apply Lemma 4.4.2 to replace Π_W with $\frac{1+o(1)}{n^2} \cdot P_W = \frac{1+o(1)}{n^2} \cdot \sum_i a_i a_i^T$. (For convenience we have written $a_i = r_i^{\otimes 2}$ here.) Let $M = \frac{1}{n^2} \cdot \sum_i D_i P_W D_i$. An entry of M has the form

$$M(ab, cd) = \frac{1}{n^2} \left(n + \sum_{i \neq j} A_{ia} A_{ib} A_{ic} A_{id} A_{ja} A_{jb} A_{jc} A_{jb} \right).$$

Thus we can see that $M = \frac{1}{n} J_{\binom{n}{2}} + \frac{1+o(1)}{n^2} BB^\top$, where $J_{\binom{n}{2}}$ is the all-ones matrix in $\mathbb{R}^{\binom{n}{2} \times \binom{n}{2}}$ and B is the matrix whose entries have the form

$$B(ab, ij) = A_{ia} A_{ib} A_{ja} A_{jb}.$$

The matrix B corresponds to a patterned matrix with a 2-matching, meeting the conditions of Lemma 4.5.7. Thus we have $\|B\| \lesssim n \log^3 n$ with probability $1 - O(n^{-5})$. We can thus conclude that with probability $1 - O(n^{-5})$, $\|M - \frac{1}{n} J_{\binom{n}{2}}\| \leq \frac{1+o(1)}{n^2} \|B\|^2 \leq \tilde{O}(1)$, and so $x^T M x \leq \frac{1+o(1)}{n} \langle x, \mathbb{1}_{\binom{n}{2}} \rangle^2 + x^T (M - n^{-1} J) x \leq O(n) \cdot \|\Pi_0 x\|^2 + \tilde{O}(1) \cdot \|x\|^2$, which gives the desired result. □

Chapter 5

Fast Spectral Algorithms from Sum-of-Squares Analyses

5.1 Introduction

A sequence of recent works applies the sum-of-squares method to basic problems that arise in unsupervised machine learning: in particular, recovering sparse vectors in linear subspaces and decomposing tensors in a robust way [BKS14, BKS15, HSS15, BM16, GM15]. For a wide range of parameters of these problems, SoS achieves significantly stronger guarantees than other methods, in polynomial or quasi-polynomial time. Unfortunately, even when the SDP has size polynomial in the input (when $d = O(1)$), solving the underlying semidefinite programs is prohibitively slow for large instances.

In this chapter, we introduce spectral algorithms for planted sparse vector, tensor decomposition, and tensor principal components analysis (PCA) that exploit the same high-degree information as the corresponding sum-of-squares algorithms without relying on semidefinite programming, and achieve the same (or close to the same) guarantees. The resulting algorithms are quite simple (a couple of lines of MATLAB code) and have considerably faster running times—quasi-linear or close to linear in the input size.

A surprising implication of our work is that for some problems, spectral algorithms can exploit information from larger values of the degree parameter d without spending time $n^{O(d)}$. For example, our algorithm for the planted sparse vector problem runs in nearly-linear time in the input size, even though it uses properties that the sum-of-squares method can only use for degree parameter $d \geq 4$. (In particular, the guarantees that the algorithm achieves are strictly stronger than the guarantees that SoS achieves for values of $d < 4$.)

The initial successes of SoS in the machine learning setting gave hope that techniques developed in the theory of approximation algorithms, specifically the techniques of hierarchies of convex relaxations and rounding convex relaxations, could broadly impact the practice of machine learning. This hope was dampened by the fact that in general, algorithms that rely on solving large semidefinite programs are too slow to be practical for the large-scale problems that arise in machine learning. Our work brings this hope back into focus by demonstrating for the first time that with some care SoS algorithms can be made practical for large-scale problems.

In the following subsections we describe each of the problems that we consider, the prior best-known guarantee via the SoS hierarchy, and our results.

Planted Sparse Vector in Random Linear Subspace

The problem of finding a sparse vector planted in a random linear subspace was introduced by Spielman, Wang, and Wright as a way of learning sparse dictionaries [SWW12]. Subsequent works have found further applications and begun studying the problem in its own right [DH14, BKS14, QSW14]. In this problem, we are given a basis for a d-dimensional linear subspace of $\mathbb{R}^n$ that is random except for one planted sparse direction, and the goal is to recover this sparse direction. The computational challenge is to solve this problem even when the planted vector is only mildly sparse (a constant fraction of non-zero coordinates) and the subspace dimension is large compared to the ambient dimension ($d \geq n^{\Omega(1)}$).

Several kinds of algorithms have been proposed for this problem based on linear programming (LP), basic semidefinite programming (SDP), sum-of-squares, and non-convex gradient descent (alternating directions method).

An inherent limitation of simpler convex methods (LP and basic SDP) [SWW12, dGJL04] is that they require the relative sparsity of the planted vector to be polynomial in the subspace dimension (less than $n/\sqrt{d}$ non-zero coordinates).

Sum-of-squares and non-convex methods do not share this limitation. They can recover planted vectors with constant relative sparsity even if the subspace has polynomial dimension (up to dimension $O(n^{1/2})$ for sum-of-squares [BKS14] and up to $O(n^{1/4})$ for non-convex methods [QSW14]).

We state the problem formally:

Problem 5.1.1 (Planted sparse vector problem with ambient dimension $n \in \mathbb{N}$, subspace dimension $d \leq n$, sparsity $\varepsilon > 0$, and accuracy $\eta > 0$)**.** Given an arbitrary orthogonal basis of a subspace spanned by vectors $v_0, v_1, \ldots, v_{d-1} \in \mathbb{R}^n$, where v_0 is a vector with at most εn non-zero entries and $v_1, \ldots, v_{d-1}$ are vectors sampled independently at random from the standard Gaussian distribution on $\mathbb{R}^n$, output a unit vector $v \in \mathbb{R}^n$ that has correlation $\langle v, v_0 \rangle^2 \geq 1 - \eta$ with the sparse vector v_0.

Our Results. Our algorithm runs in nearly linear time in the input size, and matches the best-known guarantees up to a polylogarithmic factor in the subspace dimension [BKS14].

Theorem 5.1.2 (Planted sparse vector in nearly-linear time)**.** *There exists an algorithm that, for every sparsity $\varepsilon > 0$, ambient dimension n, and subspace dimension d with $d \leq \sqrt{n}/(\log n)^{O(1)}$, solves the planted sparse vector problem with high probability for some accuracy $\eta \leq O(\varepsilon^{1/4}) + o_{n\to\infty}(1)$. The running time of the algorithm is $\widetilde{O}(nd)$.*

We give a technical overview of the proof in Section 5.2, and a full proof in Section 5.3.

Previous work also showed how to recover the planted sparse vector exactly. The task of going from an approximate solution to an exact one is a special case of standard compressed sensing (see e.g. [BKS14]).

Table 5.1: Comparison of algorithms for the planted sparse vector problem with ambient dimension n, subspace dimension d, and relative sparsity ε.

Reference	Technique	Runtime	Largest d	Largest ε
[DH14]	linear programming	poly	any	$\Omega(1/\sqrt{d})$
[BKS14]	SoS, general SDP	poly	$\Omega(\sqrt{n})$	$\Omega(1)$
[QSW14]	alternating minimization	$\widetilde{O}(n^2d^5)$	$\Omega(n^{1/4})$	$\Omega(1)$
this work	SoS, partial traces	$\widetilde{O}(nd)$	$\widetilde{\Omega}(\sqrt{n})$	$\Omega(1)$

Overcomplete Tensor Decomposition

Tensors naturally represent multilinear relationships in data. Algorithms for tensor decompositions have long been studied as a tool for data analysis across a wide-range of disciplines (see the early work of Harshman [Har70] and the survey [KB09]). While the problem is NP-hard in the worst-case [Hås90, HL13], algorithms for special cases of tensor decomposition have recently led to new provable algorithmic results for several unsupervised learning problems [AGH+14, BCMV14, GVX14, AGHK14] including independent component analysis, learning mixtures of Gaussians [GHK15], Latent Dirichlet topic modeling [AFH+15] and dictionary learning [BKS15]. Some previous learning algorithms can also be reinterpreted in terms of tensor decomposition [Cha96, MR06, NR09].

A key algorithmic challenge for tensor decompositions is *overcompleteness*, when the number of components is larger than their dimension (i.e., the components are linearly dependent). Most algorithms that work in this regime require tensors of order 4 or higher [LCC07, BCMV14]. For example, the FOOBI algorithm of [LCC07] can recover up to $\Omega(d^2)$ components given an order-4 tensor in dimension d under mild algebraic independence assumptions for the components—satisfied with high probability by random components. For overcomplete 3-tensors, which arise in many applications of tensor decompositions, such a result remains elusive.

Researchers have therefore turned to investigate average-case versions of the problem, when the components of the overcomplete 3-tensor are random: Given a 3-tensor $T \in \mathbb{R}^{d^3}$ of the form

$$T = \sum_{i=1}^{n} a_i \otimes a_i \otimes a_i \,,$$

where $a_1, \ldots, a_n$ are random unit or Gaussian vectors, the goal is to approximately recover the components $a_1, \ldots, a_n$.

Algorithms based on *tensor power iteration*—a gradient-descent approach for tensor decomposition—solve this problem in polynomial time when $n \leq C \cdot d$ for any constant $C \geq 1$ (the running time is exponential in C) [AGJ15]. Tensor power iteration also admits local convergence analyses for up to $n \leq \tilde{\Omega}(d^{1.5})$ components [AGJ15, AGJ14]. Unfortunately, these analyses do not give polynomial-time algorithms because it is not known how to efficiently obtain the kind of initializations assumed by the analyses.

Recently, Ge and Ma [GM15] were able to show that a tensor-decomposition algorithm [BKS15] based on sum-of-squares solves the above problem for $n \leq \tilde{\Omega}(d^{1.5})$ in quasi-polynomial time $n^{O(\log n)}$. The key ingredient of their elegant analysis is a subtle spectral

Table 5.2: Comparison of decomposition algorithms for overcomplete 3-tensors with n components in dimension d.

Reference	Technique	Runtime	Largest n	Components
[AGJ15][a]	tensor power iteration	poly	$C \cdot d$	incoherent
[GM15]	SoS, general SDP	$n^{O(\log n)}$	$\widetilde{\Omega}(d^{3/2})$	$\mathcal{N}(0, \frac{1}{d}\,\mathrm{Id}_d)$
this work[b]	SoS, partial traces	$\widetilde{O}(nd^{1+\omega})$	$\widetilde{\Omega}(d^{4/3})$	$\mathcal{N}(0, \frac{1}{d}\,\mathrm{Id}_d)$

[a] The analysis shows that for every constant $C \geq 1$, the running time is polynomial for $n \leq C \cdot d$ components, assuming that the components also satisfy other random-like properties besides incoherence. [b] Here, $\omega \leq 2.3729$ is the constant so that $d \times d$ matrices can be multiplied in $O(d^\omega)$ arithmetic operations.

concentration bound for a particular degree-4 matrix-valued polynomial associated with the decomposition problem of random overcomplete 3-tensors.

We state the problem formally:

Problem 5.1.3 (Random tensor decomposition with dimension d, rank n, and accuracy η)**.** Let $a_1, \ldots, a_n \in \mathbb{R}^d$ be independently sampled vectors from the Gaussian distribution $\mathcal{N}(0, \frac{1}{d}\,\mathrm{Id}_d)$, and let $\mathbf{T} \in (\mathbb{R}^d)^{\otimes 3}$ be the 3-tensor $\mathbf{T} = \sum_{i=1}^n a_i^{\otimes 3}$.

Single component: Given $\mathbf{T}$ sampled as above, find a unit vector b that has correlation $\max_i \langle a_i, b\rangle \geq 1 - \eta$ with one of the vectors a_i.

All components: Given $\mathbf{T}$ sampled as above, find a set of unit vectors $\{b_1, \ldots, b_n\}$ such that $\langle a_i, b_i\rangle \geq 1 - \eta$ for every $i \in [n]$.

Our Results. We give the first polynomial-time algorithm for decomposing random overcomplete 3-tensors with up to $\omega(d)$ components. Our algorithms works as long as the number of components satisfies $n \leq \tilde{\Omega}(d^{4/3})$, which comes close to the bound $\tilde{\Omega}(d^{1.5})$ achieved by the aforementioned quasi-polynomial algorithm of Ge and Ma. For the single-component version of the problem, our algorithm runs in time close to linear in the input size.

Theorem 5.1.4 (Fast random tensor decomposition)**.** *There exist randomized algorithms that, for every dimension d and rank n with $d \leq n \leq d^{4/3}/(\log n)^{O(1)}$, solve the random tensor decomposition problem with probability $1 - o(1)$ for some accuracy $\eta \leq \widetilde{O}(n^3/d^4)^{1/2}$. The running time for the single-component version of the problem is $\widetilde{O}(d^{1+\omega})$, where $d^\omega \leq d^{2.3279}$ is the time to multiply two d-by-d matrices. The running time for the all-components version of the problem is $\widetilde{O}(n \cdot d^{1+\omega})$.*

We give a technical overview of the proof in Section 5.2, and a full proof in Section 5.4.

We remark that the above algorithm only requires access to the input tensor with some fixed inverse polynomial accuracy because each of its four steps amplifies errors by at most a polynomial factor (see Algorithm 5.4.17). In this sense, the algorithm is robust.

Table 5.3: Comparison of algorithms for principal component analysis of 3-tensors in dimension d and with signal-to-noise ratio τ.

Reference	Technique	Runtime	Smallest τ
[RM14]	spectral	$\widetilde{O}(d^3)$	n
[HSS15]	SoS, spectral	$\widetilde{O}(d^3)$	$O(n^{3/4})$
this work	SoS, partial traces	$O(d^3)$	$\widetilde{O}(n^{3/4})$

Tensor Principal Component Analysis

The problem of tensor principal component analysis is similar to the tensor decomposition problem. However, here the focus is not on the number of components in the tensor, but about recovery in the presence of a large amount of random noise. We are given as input a tensor $\tau \cdot v^{\otimes 3} + \mathbf{A}$, where $v \in \mathbb{R}^n$ is a unit vector and the entries of $\mathbf{A}$ are chosen iid from $\mathcal{N}(0,1)$. This *spiked tensor* model was introduced by Montanari and Richard [RM14], who also obtained the first algorithms to solve the model with provable statistical guarantees. The spiked tensor model was subsequently addressed by a subset of the present authors [HSS15], who applied the SoS approach to improve the signal-to-noise ratio required for recovery from odd-order tensors.

We state the problem formally:

Problem 5.1.5 (Tensor principal components analysis with signal-to-noise ratio τ and accuracy η)**.** Let $\mathbf{T} \in (\mathbb{R}^d)^{\otimes 3}$ be a tensor so that $\mathbf{T} = \tau \cdot v^{\otimes 3} + \mathbf{A}$, where $\mathbf{A}$ is a tensor with independent standard Gaussian entries and $v \in \mathbb{R}^d$ is a unit vector. Given $\mathbf{T}$, recover a unit vector $v' \in \mathbb{R}^d$ such that $\langle v', v\rangle \geq 1 - \eta$.

Our results. For this problem, our improvements over the previous results are more modest—we achieve signal-to-noise guarantees matching [HSS15], but with an algorithm that runs in linear time rather than near-linear time (time $O(d^3)$ rather than $O(d^3 \operatorname{polylog} d)$, for an input of size d^3).

Theorem 5.1.6 (Tensor principal component analysis in linear time)**.** *There is an algorithm which solves the tensor principal component analysis problem with accuracy $\eta > 0$ whenever the signal-to-noise ratio satisfies $\tau \geq O(n^{3/4} \cdot \eta^{-1} \cdot \log^{1/2} n)$. Furthermore, the algorithm runs in time $O(d^3)$.*

Though for tensor PCA our improvement over previous work is modest, we include the results here as this problem is a pedagogically poignant illustration of our techniques. We give a technical overview of the proof in Section 5.2, and a full proof in Section 5.5.

Related Work

Foremost, this work builds upon the SoS algorithms of [BKS14, BKS15, GM15, HSS15]. In each of these previous works, a machine learning decision problem is solved using an SDP

relaxation for SoS. In these works, the SDP value is large in the YES case and small in the NO case, and the SDP value can be bounded using the spectrum of a specific matrix. This was implicit in [BKS14, BKS15], and in [HSS15] it was used to obtain a fast algorithm as well. In our work, we design spectral algorithms which use smaller matrices, inspired by the SoS certificates in previous works, to solve these machine-learning problems much faster, with almost matching guarantees.

A key idea in our work is that given a large matrix with information encoded in the matrix's spectral gap, one can often efficiently "compress" the matrix to a much smaller one without losing that information. This is particularly true for problems with planted solutions. In this way, we are able to improve running time by replacing an $n^{O(d)}$-sized SDP with an eigenvector computation for an $n^k \times n^k$ matrix, for some $k < d$.

The idea of speeding up LP and SDP hierarchies for specific problems has been investigated in a series of previous works [dlVK07, BRS11, GS12], which shows that with respect to local analyses of the sum-of-squares algorithm it is sometimes possible to improve the running time from $n^{O(d)}$ to $2^{O(d)} \cdot n^{O(1)}$. However, the scopes and strategies of these works are completely different from ours. First, the notion of local analysis from these works does not apply to the problems considered here. Second, these works employ the ellipsoid method with a separation oracle inspired by rounding algorithms, whereas we reduce the problem to an ordinary eigenvector computation.

Since the publication of our results, our methods have been extended to speed up some of the other recent successful applications of SoS to machine-learning type problems, such as [BM16] in the work of [MS16], and the application of [MSS16] to tensor decomposition with components that are orthogonal in the presence of noise in [SS17].

Finally, we would be remiss not to mention that SoS lower bounds exist for several of these problems, specifically for tensor principal components analysis, tensor prediction, and sparse PCA [HSS15, BM16, MW15]. The lower bounds in the SoS framework are a good indication that we cannot expect spectral algorithms achieving better guarantees.

5.2 Techniques

Sum-of-Squares Method (for Polynomial Optimization over the Sphere). The problems we consider are connected to optimization problems of the following form: Given a homogeneous n-variate real polynomial f of constant degree, find a unit vector $x \in \mathbb{R}^n$ so as to maximize $f(x)$. The sum-of-squares method allows us to efficiently compute upper bounds on the maximum value of such a polynomial f over the unit sphere.

For the case that $k = \deg(f)$ is even, the most basic upper bound of this kind is the largest eigenvalue of a *matrix representation* of f. A *matrix representation* of a polynomial f is a symmetric matrix M with rows and columns indexed by monomials of degree $k/2$ so that $f(x)$ can be written as the quadratic form $f(x) = \langle x^{\otimes k/2}, Mx^{\otimes k/2} \rangle$, where $x^{\otimes k/2}$ is the $k/2$-fold tensor power of x. The largest eigenvalue of a matrix representation M is an upper bound on the value of $f(x)$ over all unit vectors $x \in \mathbb{R}^n$ because

$$f(x) = \langle x^{\otimes k/2}, Mx^{\otimes k/2} \rangle \leq \lambda_{\max}(M) \cdot \|x^{\otimes k/2}\|_2^2 = \lambda_{\max}(M) \,.$$

As explained in Chapter 2, the sum-of-squares methods improves on this basic spectral bound systematically by associating a large family of polynomials (potentially of degree

higher than $\deg(f)$) with the input polynomial f and computing the best possible spectral bound within this family of polynomials.

Our approach for faster algorithms based on SoS algorithms is to construct specific matrices (polynomials) in the family, then compute their top eigenvectors. By designing our matrices carefully, we ensure that our algorithms have access to the same higher degree information that the sum-of-squares algorithm can access, and this information affords an advantage over the basic spectral methods for these problems. At the same time, our algorithms avoid searching for the best polynomial and matrix representation, which gives us faster running times since we avoid semidefinite programming. This approach is well suited to average-case problems where we avoid the problem of adversarial choice of input; in particular it is applicable to machine learning problems where noise and inputs are assumed to be random.

Compressing Matrices with Partial Traces. A serious limitation of the above approach is that the representation of a degree-d, n-variate polynomial requires size roughly n^d. Hence, even avoiding the use of semidefinite programming, improving upon running time $O(n^d)$ requires additional ideas.

In each of the problems that we consider, we have a large matrix (suggested by a SoS algorithm) with a "signal" planted in some amount of "noise". We show that in some situations, this large matrix can be compressed significantly without loss in the signal by applying *partial trace* operations. In these situations, the partial trace yields a smaller matrix with the same signal-to-noise ratio as the large matrix suggested by the SoS algorithm, even in situations when lower degree sum-of-squares approaches are known to fail (as for the planted sparse vector and tensor PCA problems).[1]

The partial trace $\mathrm{Tr}_{\mathbb{R}^d} \colon \mathbb{R}^{d^2 \times d^2} \to \mathbb{R}^{d \times d}$ is the linear operator that satisfies $\mathrm{Tr}_{\mathbb{R}^d} A \otimes B = (\mathrm{Tr}\, A) \cdot B$ for all $A, B \in \mathbb{R}^{d \times d}$. To see how the partial trace can be used to compress large matrices to smaller ones with little loss, consider the following problem: Given a matrix $M \in \mathbb{R}^{d^2 \times d^2}$ of the form $M = \tau \cdot (v \otimes v)(v \otimes v)^\top + A \otimes B$ for some unit vector $v \in \mathbb{R}^d$ and matrices $A, B \in \mathbb{R}^{d \times d}$, we wish to recover the vector v. (This is a simplified version of the planted problems that we consider in this paper, where $\tau \cdot (v \otimes v)(v \otimes v)^\top$ is the signal and $A \otimes B$ plays the role of noise.)

It is straightforward to see that the matrix $A \otimes B$ has spectral norm $\|A \otimes B\| = \|A\| \cdot \|B\|$, and so when $\tau \gg \|A\|\|B\|$, the matrix M has a noticeable spectral gap, and the top eigenvector of M will be close to $v \otimes v$. If $|\mathrm{Tr}\, A| \approx \|A\|$, the matrix $\mathrm{Tr}_{\mathbb{R}^d} M = \tau \cdot vv^\top + \mathrm{Tr}(A) \cdot B$ has a matching spectral gap, and we can still recover v, but now we only need to compute the top eigenvector of a $d \times d$ (as opposed to $d^2 \times d^2$) matrix.[2]

If A is a Wigner matrix (e.g. a symmetric matrix with iid ± 1 entries), then both $\mathrm{Tr}(A), \|A\| \approx \sqrt{n}$, and the above condition is indeed met. In our average case/machine

[1] For both problems we use matrices with dimensions corresponding to degree-4 SoS programs. An argument of Spielman et al. ([SWW12], Theorem 9) shows that degree-2 sum-of-squares can only find sparse vectors with sparsity $k \leq \widetilde{O}(\sqrt{n})$, whereas we achieve sparsity as large as $k = \Theta(n)$. For tensor PCA, the degree-2 SoS program cannot even express the objective function.

[2] In some of our applications, the matrix M is only represented implicitly and has size super-linear in the size of the input, but nevertheless we can compute the top eigenvector of the partial trace $\mathrm{Tr}_{\mathbb{R}^d} M$ in nearly-linear time.

learning settings the "noise" component is not as simple as $A \otimes B$ with A a Wigner matrix. Nonetheless, we are able to ensure that the noise displays a similar behavior under partial trace operations. In some cases, this requires additional algorithmic steps, such as random projection in the case of tensor decomposition, or centering the matrix eigenvalue distribution in the case of the planted sparse vector.

It is an interesting question if there are general theorems describing the behavior of spectral norms under partial trace operations. In the current work, we compute the partial traces explicitly and estimate their norms directly. Indeed, our analyses boil down to concentrations bounds for special matrix polynomials. A general theory for the concentration of matrix polynomials is a notorious open problem (see [MW13]).

Partial trace operations have previously been applied for rounding SoS relaxations. Specifically, the operation of *reweighing* and *conditioning*, used in rounding algorithms for sum-of-squares such as [BRS11, RT12, BKS14, BKS15, LR15], corresponds to applying a partial trace operation to the moments matrix returned by the sum-of-squares relaxation.

We now give a technical overview of our algorithmic approach for each problem, and some broad strokes of the analysis for each case. Our most substantial improvements in runtime are for the planted sparse vector and overcomplete tensor decomposition problems (Section 5.2 and Section 5.2 respectively). Our algorithm for tensor PCA is the simplest application of our techniques, and it may be instructive to skip ahead and read about tensor PCA first (Section 5.2).

Planted Sparse Vector in Random Linear Subspace

Recall that in this problem we are given a linear subspace U (represented by some basis) that is spanned by a k-sparse unit vector $v_0 \in \mathbb{R}^d$ and random unit vectors $v_1, \ldots, v_{d-1} \in \mathbb{R}^d$. The goal is to recover the vector v_0 approximately.

Background and SoS Analysis. Let $A \in \mathbb{R}^{n \times d}$ be a matrix whose columns form an orthonormal basis for U. Our starting point is the polynomial $f(x) = \|Ax\|_4^4 = \sum_{i=1}^n (Ax)_i^4$. Previous work showed that for $d \ll \sqrt{n}$ the maximizer of this polynomial over the sphere corresponds to a vector close to v_0 and that degree-4 sum-of-squares is able to capture this fact [BBH+12, BKS14]. Indeed, typical random vectors v in $\mathbb{R}^n$ satisfy $\|v\|_4^4 \approx 1/n$ whereas our planted vector satisfies $\|v_0\|_4^4 \geq 1/k \gg 1/n$, and this degree-4 information is leveraged by the SoS algorithms.

The polynomial f has a convenient matrix representation $M = \sum_{i=1}^n (a_i a_i^\top)^{\otimes 2}$, where $a_1, \ldots, a_n$ are the rows of the generator matrix A. It turns out that the eigenvalues of this matrix indeed give information about the planted sparse vector v_0. In particular, the vector $x_0 \in \mathbb{R}^d$ with $Ax_0 = v_0$ witnesses that M has an eigenvalue of at least $1/k$ because M's quadratic form with the vector $x_0^{\otimes 2}$ satisfies $\langle x_0^{\otimes 2}, M x_0^{\otimes 2} \rangle = \|v_0\|_4^4 \geq 1/k$. If we let M' be the corresponding matrix for the subspace U without the planted sparse vector, M' turns out to have only eigenvalues of at most $O(1/n)$ up to a single spurious eigenvalue with eigenvector far from any vector of the form $x \otimes x$ [BBH+12].

It follows that in order to distinguish between a random subspace with a planted sparse vector (YES case) and a completely random subspace (NO case), it is enough to compute the second-largest eigenvalue of a d^2-by-d^2 matrix (representing the 4-norm polynomial over

the subspace as above). This decision version of the problem, while strictly speaking easier than the search version above, is at the heart of the matter: one can show that the large eigenvalue for the YES case corresponds to an eigenvector which encodes the coefficients of the sparse planted vector in the basis.

Improvements. The best running time we can hope for with this basic approach is $O(d^4)$ (the size of the matrix). Since we are interested in $d \leq O(\sqrt{n})$, the resulting running time $O(nd^2)$ would be subquadratic but still super-linear in the input size $n \cdot d$ (for representing a d-dimensional subspace of $\mathbb{R}^n$). To speed things up, we use the partial trace approach outlined above. We will begin by applying the partial trace approach naively, obtaining reasonable bounds, and then show that a small modification to the matrix before the partial trace operation allows us to achieve even smaller signal-to-noise ratios.

In the planted case, we may approximate $M \approx \frac{1}{k}(x_0 x_0^\top)^{\otimes 2} + Z$, where x_0 is the vector of coefficients of v_0 in the basis representation given by A (so that $Ax_0 = v_0$), and Z is the noise matrix. Since $\|x_0\| = 1$, the partial trace operation preserves the projector $(x_0 x_0^\top)^{\otimes 2}$ in the sense that $\mathrm{Tr}_{\mathbb{R}^d}(x_0 x_0^\top)^{\otimes 2} = x_0 x_0^\top$. Hence, with our heuristic approximation for M above, we could show that the top eigenvector of $\mathrm{Tr}_{\mathbb{R}^d} M$ is close to x_0 by showing that the spectral norm bound $\|\mathrm{Tr}_{\mathbb{R}^d} Z\| \leq o(1/k)$.

The partial trace of our matrix $M = \sum_{i=1}^n (a_i a_i^\top) \otimes (a_i a_i^\top)$ is easy to compute directly,

$$N = \mathrm{Tr}_{\mathbb{R}^d} M = \sum_{i=1}^n \|a_i\|_2^2 \cdot a_i a_i^\top .$$

In the YES case (random subspace with planted sparse vector), a direct computation shows that

$$\lambda_{\text{YES}} \geq \langle x_0, N x_0 \rangle \approx \tfrac{d}{n} \cdot \left(1 + \tfrac{n}{d}\|v_0\|_4^4\right) \geq \tfrac{d}{n}\left(1 + \tfrac{n}{dk}\right) .$$

Hence, a natural approach to distinguish between the YES case and NO case (completely random subspace) is to upper bound the spectral norm of N in the NO case.

In order to simplify the bound on the spectral norm of N in the NO case, suppose that the columns of A are iid samples from the Gaussian distribution $\mathcal{N}(0, \frac{1}{d}\,\mathrm{Id})$ (rather than an orthogonal basis for the random subspace)–Lemma 5.3.6 establishes that this simplification is legitimate. In this simplified setup, the matrix N in the NO case is the sum of n iid matrices $\{\|a_i\|^2 \cdot a_i a_i^\top\}$, and we can upper bound its spectral norm λ_{NO} by $d/n \cdot (1 + O(\sqrt{d/n}))$ using standard matrix concentration bounds. Hence, using the spectral norm of N, we will be able to distinguish between the YES case and the NO case as long as

$$\sqrt{d/n} \ll n/(dk) \implies \lambda_{\text{NO}} \ll \lambda_{\text{YES}} .$$

For linear sparsity $k = \varepsilon \cdot n$, this inequality is true so long as $d \ll (n/\varepsilon^2)^{1/3}$, which is somewhat worse than the bound $\sqrt{n}$ bound on the dimension that we are aiming for.

Recall that $\mathrm{Tr}\, B = \sum_i \lambda_i(B)$ for a symmetric matrix B. As discussed above, the partial trace approach works best when the noise behaves as the tensor of two Wigner matrices, in that there are cancellations when the eigenvalues of the noise are summed. In our case, the noise terms $(a_i a_i^\top) \otimes (a_i a_i^\top)$ do not have this property, as in fact $\mathrm{Tr}\, a_i a_i^\top = \|a_i\|^2 \approx d/n$.

Thus, in order to improve the dimension bound, we will center the eigenvalue distribution of the noise part of the matrix. This will cause it to behave more like a Wigner matrix, in that the spectral norm of the noise will not increase after a partial trace. Consider the partial trace of a matrix of the form

$$M - \alpha \cdot \mathrm{Id} \otimes \sum_i a_i a_i^\top ,$$

for some constant $\alpha > 0$. The partial trace of this matrix is

$$N' = \sum_{i=1}^{n} (\|a_i\|_2^2 - \alpha) \cdot a_i a_i^\top .$$

We choose the constant $\alpha \approx d/n$ such that our matrix N' has expectation 0 in the NO case, when the subspace is completely random. In the YES case, the Rayleigh quotient of N' at x_0 simply shifts as compared to N, and we have $\lambda_{\text{YES}} \geq \langle x_0, N' x_0 \rangle \approx \|v_0\|_4^4 \geq 1/k$ (see Lemma 5.3.5 and sublemmas). On the other hand, in the NO case, this centering operation causes significant cancellations in the eigenvalues of the partial trace matrix (instead of just shifting the eigenvalues). In the NO case, N' has spectral norm $\lambda_{\text{NO}} \leq O(d/n^{3/2})$ for $d \ll \sqrt{n}$ (using standard matrix concentration bounds; again see Lemma 5.3.5 and sublemmas). Therefore, the spectral norm of the matrix N' allows us to distinguish between the YES and NO case as long as $d/n^{3/2} \ll 1/k$, which is satisfied as long as $k \ll n$ and $d \ll \sqrt{n}$. We give the full formal argument in Section 5.3.

Overcomplete Tensor Decomposition

Recall that in this problem we are given a 3-tensor $\mathbf{T}$ of the form $\mathbf{T} = \sum_{i=1}^n a_i^{\otimes 3} \in \mathbb{R}^{d^3}$, where $a_1, \ldots, a_n \in \mathbb{R}^d$ are independent random vectors from $\mathcal{N}(0, \frac{1}{d}\,\mathrm{Id})$. The goal is to find a unit vector $a \in \mathbb{R}^d$ that is highly correlated with one[3] of the vectors $a_1, \ldots, a_n$.

Background. The starting point of our algorithm is the polynomial $f(x) = \sum_{i=1}^n \langle a_i, x \rangle^3$. It turns out that for $n \ll d^{1.5}$ the (approximate) maximizers of this polynomial are close to the components $a_1, \ldots, a_n$, in the sense that $f(x) \approx 1$ if and only if $\max_{i \in [n]} \langle a_i, x \rangle^2 \approx 1$. Indeed, Ge and Ma [GM15] show that the sum-of-squares method already captures this fact at degree 12, which implies a quasipolynomial time algorithm for this tensor decomposition problem via a general rounding result of Barak, Kelner, and Steurer [BKS15].

The simplest approach to this problem is to consider the tensor representation of the polynomial $\mathbf{T} = \sum_{i \in [n]} a_i^{\otimes 3}$, and flatten it, hoping the singular vectors of the flattening are correlated with the a_i. However, this approach is doomed to failure for two reasons: firstly, the simple flattenings of $\mathbf{T}$ are $d^2 \times d$ matrices, and since $n \gg d$ the $a_i^{\otimes 2}$ collide in the column space, so that it is impossible to determine $\mathrm{Span}\{a_i^{\otimes 2}\}$. Secondly, even for $n \leq d$, because the a_i are random vectors, their norms concentrate very closely about 1. This makes it difficult to distinguish any one particular a_i even when the span is computable.

[3] We can then approximately recover all the components $a_1, \ldots, a_n$ by running independent trials of our randomized algorithm repeatedly on the same input.

Improvements. We will try to circumvent both of these issues by going to higher dimensions. Suppose, for example, that we had access to $\sum_{i\in[n]} a_i^{\otimes 4}$.[4] The eigenvectors of the flattenings of this matrix are all within $\mathrm{Span}_{i\in[n]}\{a_i^{\otimes 2}\}$, addressing our first issue, leaving us only with the trouble of extracting individual $a_i^{\otimes 2}$ from their span. If furthermore we had access to $\sum_{i\in[n]} a_i^{\otimes 6}$, we could perform a partial random projection $(\Phi\otimes \mathrm{Id}\otimes \mathrm{Id})\sum_{i\in[n]} a_i^{\otimes 6}$ where $\Phi\in\mathbb{R}^{d\times d}$ is a matrix with independent Gaussian entries, and then taking a partial trace, we end up with

$$\mathrm{Tr}_{\mathbb{R}^d}\left((\Phi\otimes \mathrm{Id}\otimes \mathrm{Id})\sum_{i\in[n]} a_i^{\otimes 6}\right) = \sum_{i\in[n]}\langle \Phi, a_i^{\otimes 2}\rangle a_i^{\otimes 4}\,.$$

With reasonable probability (for exposition's sake, say with probability $1/n^{10}$), Φ is closer to some $a_i^{\otimes 2}$ than to all of the others so that $\langle \Phi, a_i^{\otimes 2}\rangle \geq 100\langle \Phi, a_j^{\otimes 2}\rangle$ for all $j\in[n]$, and then $a_i^{\otimes 2}$ is distinguishable from the other vectors in the span of our matrix, taking care of the second issue . As we show, a much smaller gap is sufficient to distinguish the top a_i from the other a_j, and so the higher-probability event that Φ is only slightly closer to a_i suffices (allowing us to recover all vectors at an additional runtime cost of a factor of $\widetilde{O}(n)$). This discussion ignores the presence of a single spurious large eigenvector, which we address in the technical sections.

Of course, we do not have access to the higher-order tensor $\sum_{i\in[n]} a_i^{\otimes 6}$. Instead, we can obtain a noisy version of this tensor. Our approach considers the following matrix representation of the polynomial f^2,

$$M = \sum_{i,j} a_i a_j^\top \otimes (a_i a_i^\top)\otimes(a_j a_j^\top)\in\mathbb{R}^{d^3\times d^3}\,.$$

Alternatively, we can view this matrix as a particular flattening of the Kronecker-squared tensor $\mathbf{T}^{\otimes 2}$. It is instructive to decompose $M = M_{\mathrm{diag}} + M_{\mathrm{cross}}$ into its diagonal terms $M_{\mathrm{diag}} = \sum_i (a_i a_i^\top)^{\otimes 3}$ and its cross terms $M_{\mathrm{cross}} = \sum_{i\neq j} a_i a_j^\top \otimes (a_i a_i^\top)\otimes(a_j a_j^\top)$. The algorithm described above is already successful for M_{diag}; we need only control the eigenvalues of the partial trace of the "noise" component, M_{cross}. The main technical work will be to show that $\|\mathrm{Tr}_{\mathbb{R}^d} M_{\mathrm{diag}}\|$ is small. In fact, we will have to choose Φ from a somewhat different distribution—observing that $\mathrm{Tr}_{\mathbb{R}^d}(\Phi\otimes \mathrm{Id}\otimes \mathrm{Id}) = \sum_{i,j}\langle a_i, \Phi a_j\rangle\cdot(a_i\otimes a_j)(a_i\otimes a_j)^\top$, we will sample Φ so that $\langle a_i, \Phi a_i\rangle \gg \langle a_i, \Phi a_j\rangle$. We give a more detailed overview of this algorithm in the beginning of Section 5.4, explaining in more detail our choice of Φ and justifying heuristically the boundedness of the spectral norm of the noise.

Connection to SoS Analysis. To explain how the above algorithm is a speedup of SoS, we give an overview of the SoS algorithm of [GM15, BKS15]. There, the degree-t SoS

[4] As the problem is defined, we assume that we do not have access to this input, and in many machine learning applications this is a valid assumption, as gathering the data necessary to generate the 4th order input tensor requires a prohibitively large number of samples.

SDP program is used to obtain an order-t tensor χ_t (or a *pseudodistribution*). Informally speaking, we can understand χ_t as a proxy for $\sum_{i\in[n]} a_i^{\otimes t}$, so that $\chi_t = \sum_{i\in[n]} a_i^{\otimes t} + N$, where N is a noise tensor. While the precise form of N is unclear, we know that N must obey a set of constraints imposed by the SoS hierarchy at degree t. For a formal discussion of pseudodistributions, see [BKS15].

To extract a single component a_i from the tensor $\sum_{i\in[n]} a_i^{\otimes t}$, there are many algorithms which would work (for example, the algorithm we described for M_{diag} above). However, any algorithm extracting an a_i from χ_t must be robust to the noise tensor N. For this it turns out the following algorithm will do: suppose we have the tensor $\sum_{i\in[n]} a_i^{\otimes t}$, taking $t = O(\log n)$. Sample $g_1, \dots, g_{\log(n)-2}$ random unit vectors, and compute the matrix $M = \sum_i (\prod_{1\le j\le \log(n)-2} \langle g_j, a_i\rangle) \cdot a_i a_i^\top$. If we are lucky enough, there is some a_i so that every g_j is a bit closer to a_i than any other $a_{i'}$, and $M = a_i a_i^\top + E$ for some $\|E\| \ll 1$. The proof that $\|E\|$ is small can be made so simple that it applies also to the SDP-produced proxy tensor $\chi_{\log n}$, and so this algorithm is robust to the noise N. This last step is very general and can handle tensors whose components a_i are less well-behaved than the random vectors we consider, and also more overcomplete, handling tensors of rank up to $n = \tilde{\Omega}(d^{1.5})$.[5]

Our subquadratic-time algorithm can be viewed as a low-degree, spectral analogue of the [BKS15] SoS algorithm. However, rather than relying on an SDP to produce an object close to $\sum_{i\in[n]} a_i^{\otimes t}$, we manufacture one ourselves by taking the Kronecker square of our input tensor. We explicitly know the form of the deviation of $\mathbf{T}^{\otimes 2}$ from $\sum_{i\in[n]} a_i^{\otimes 6}$, unlike in [BKS15], where the deviation of the SDP certificate χ_t from $\sum_{i\in[n]} a_i^{\otimes t}$ is poorly understood. We are thus able to control this deviation (or "noise") in a less computationally intensive way, by cleverly designing a partial trace operation which decreases the spectral norm of the deviation. Since the tensor handled by the algorithm is much smaller—order 6 rather than order $\log n$—this provides the desired speedup.

Tensor Principal Component Analysis

Recall that in this problem we are given a tensor $\mathbf{T} = \tau \cdot v^{\otimes 3} + \mathbf{A}$, where $v \in \mathbb{R}^d$ is a unit vector, $\mathbf{A}$ has iid entries from $\mathcal{N}(0,1)$, and $\tau > 0$ is the signal-to-noise ratio. The aim is to recover v approximately.

Background and SoS Analysis. A previous application of SoS techniques to this problem discussed several SoS or spectral algorithms, including one that runs in quasi-linear time [HSS15]. Here we apply the partial trace method to a subquadratic spectral SoS algorithm discussed in [HSS15] to achieve nearly the same signal-to-noise guarantee in only linear time.

Our starting point is the polynomial $\mathbf{T}(x) = \tau \cdot \langle v, x\rangle^3 + \langle x^{\otimes 3}, \mathbf{A}\rangle$. The maximizer of $\mathbf{T}(x)$ over the sphere is close to the vector v so long as $\tau \gg \sqrt{n}$ [RM14]. In [HSS15], it was shown that degree-4 SoS maximizing this polynomial can recover v with a signal-to-noise ratio of at least $\widetilde{\Omega}(n^{3/4})$, since there exists a suitable SoS bound on the noise term $\langle x^{\otimes 3}, \mathbf{A}\rangle$.

[5] It is an interesting open question whether taking $t = O(\log n)$ is really necessary, or whether this heavy computational requirement is simply an artifact of the SoS proof.

Specifically, let A_i be the ith slice of $\mathbf{A}$, so that $\langle x, A_i x\rangle$ is the quadratic form $\sum_{j,k} \mathbf{A}_{ijk} x_j x_k$. Then there is a SoS proof that $\mathbf{T}(x)$ is bounded by $|\mathbf{T}(x) - \tau \cdot \langle v, x\rangle^3| \leq f(x)^{1/2} \cdot \|x\|$, where $f(x)$ is the degree-4 polynomial $f(x) = \sum_i \langle x, A_i x\rangle^2$. The polynomial f has a convenient matrix representation: $f(x) = \langle x^{\otimes 2}, (\sum_i A_i \otimes A_i) x^{\otimes 2}\rangle$: since this matrix is a sum of iid random matrices $A_i \otimes A_i$, a matrix Chernoff bound shows that this matrix spectrally concentrates to its expectation. So with high probability one can show that the eigenvalues of $\sum_i A_i \otimes A_i$ are at most $\approx d^{3/2} \log(d)^{1/2}$ (except for a single spurious eigenvector), and it follows that degree-4 SoS solves tensor PCA so long as $\tau \gg d^{3/4} \log(d)^{1/4}$.

This leads the authors to consider a slight modification of $f(x)$, given by $g(x) = \sum_i \langle x, T_i x\rangle^2$, where T_i is the ith slice of $\mathbf{T}$. Like $\mathbf{T}$, the function g also contains information about v, and the SoS bound on the noise term in $\mathbf{T}$ carries over as an analogous bound on the noise in g. In particular, expanding $T_i \otimes T_i$ and ignoring some negligible cross-terms yields

$$\sum_i T_i \otimes T_i \approx \tau^2 \cdot (v \otimes v)(v \otimes v)^\top + \sum_i A_i \otimes A_i \,.$$

Using $v \otimes v$ as a test vector, the quadratic form of the latter matrix can be made at least $\tau^2 - O(d^{3/2} \log(d)^{1/2})$. Together with the boundedness of the eigenvalues of $\sum_i A_i \otimes A_i$ this shows that when $\tau \gg d^{3/4} \log(d)^{1/4}$ there is a spectral algorithm to recover v. Since the matrix $\sum_i T_i \otimes T_i$ is $d^2 \times d^2$, computing the top eigenvector requires $\widetilde{O}(d^4 \log n)$ time, and by comparison to the input size d^3 the algorithm runs in subquadratic time.

Improvements. In this work we speed this up to a linear time algorithm via the partial trace approach. As we have seen, the heart of the matter is to show that taking the partial trace of $\tau^2 \cdot (v \otimes v)(v \otimes v)^\top + \sum_i A_i \otimes A_i$ does not increase the spectral noise. That is, we require that

$$\left\| \mathrm{Tr}_{\mathbb{R}^d} \sum_i A_i \otimes A_i \right\| = \left\| \sum_i \mathrm{Tr}(A_i) \cdot A_i \right\| \leq O(d^{3/2} \log(d)^{1/2}) \,.$$

The A_i have iid Gaussian entries, and so as in the case of Wigner matrices, it is roughly true that $|\mathrm{Tr}(A_i)| \approx \|A_i\|$. Thus the situation is very similar to our toy example of the application of partial traces in Section 5.2.

Heuristically, because $\sum_{i \in [n]} A_i \otimes A_i$ and $\sum_{i \in [n]} \mathrm{Tr}(A_i) \cdot A_i$ are random matrices, we expect that their eigenvalues are all of roughly the same magnitude. This means that their spectral norm should be close to their Frobenius norm divided by the square root of the dimension, since for a matrix M with eigenvalues $\lambda_1, \ldots, \lambda_n$, $\|M\|_F = \sqrt{\sum_{i \in [n]} \lambda_i^2}$. By estimating the sum of the squared entries, we expect that the Frobenius norm of $\sum_i \mathrm{Tr}(A_i) \cdot A_i$ is less than that of $\sum_i A_i \otimes A_i$ by a factor of $\sqrt{d}$ after the partial trace, while the dimension decreases by a factor of d, and so assuming that the eigenvalues are all of the same order, a typical eigenvalue should remain unchanged. We formalize these heuristic calculations using standard matrix concentration arguments in Section 5.5.

5.3 Planted Sparse Vector in Random Linear Subspace

In this section we give a nearly-linear-time algorithm to recover a sparse vector planted in a random subspace.

Problem 5.3.1. Let $v_0 \in \mathbb{R}^n$ be a unit vector such that $\|v_0\|_4^4 \geq \frac{1}{\varepsilon n}$. Let $v_1, \ldots, v_{d-1} \in \mathbb{R}^n$ be iid from $\mathcal{N}(0, \frac{1}{n}\mathrm{Id}_n)$. Let $w_0, \ldots, w_{d-1}$ be an orthogonal basis for $\mathrm{Span}\{v_0, \ldots, v_{d-1}\}$. **Given:** $w_0, \ldots, w_{d-1}$ **Find:** a vector v such that $\langle v, v_0\rangle^2 \geq 1 - o(1)$.

Sparse Vector Recovery in Nearly-Linear Time

Algorithm 5.3.2. Input: $w_0, \ldots, w_{d-1}$ as in Problem 5.3.1. Goal: Find v with $\langle \hat{v}, v_0\rangle^2 \geq 1 - o(1)$.

- Compute leverage scores $\|a_1\|^2, \ldots, \|a_n\|^2$, where a_i is the ith row of the $n \times d$ matrix $S := \begin{pmatrix} w_0 & \cdots & w_{d-1} \end{pmatrix}$.
- Compute the top eigenvector u of the matrix

$$A \stackrel{\text{def}}{=} \sum_{i\in[n]} (\|a_i\|_2^2 - \tfrac{d}{n}) \cdot a_i a_i^\top .$$

- Output Su.

Remark 5.3.3 (Implementation of Algorithm 5.3.2 in nearly-linear time). The leverage scores $\|a_1\|^2, \ldots, \|a_n\|^2$ are clearly computable in time $O(nd)$. In the course of proving correctness of the algorithm we will show that A has constant spectral gap, so by a standard analysis $O(\log d)$ matrix-vector multiplies suffice to recover its top eigenvector. A single matrix-vector multiply Ax requires computing $c_i := (\|a_i\|^2 - \frac{d}{n})\langle a_i, x\rangle$ for each i (in time $O(nd)$) and summing $\sum_{i\in[n]} c_i x_i$ (in time $O(nd)$). Finally, computing Su requires summing d vectors of dimension n, again taking time $O(nd)$.

The following theorem expresses correctness of the algorithm.

Theorem 5.3.4. *Let $v_0 \in \mathbb{R}^n$ be a unit vector with $\|v_0\|_4^4 \geq \frac{1}{\varepsilon n}$. Let $v_1, \ldots, v_{d-1} \in \mathbb{R}^n$ be iid from $\mathcal{N}(0, \frac{1}{n}\mathrm{Id}_n)$. Let $w_0, \ldots, w_{d-1}$ be an orthogonal basis for $\mathrm{Span}\{v_0, \ldots, v_{d-1}\}$. Let a_i be the i-th row of the $n \times d$ matrix $S := \begin{pmatrix} w_0 & \cdots & w_{d-1} \end{pmatrix}$.*

When $d \leq n^{1/2}/\operatorname{polylog}(n)$, for any sparsity $\varepsilon > 0$, w.ov.p. the top eigenvector u of $\sum_{i=1}^n (\|a_i\|^2 - \frac{d}{n}) \cdot a_i a_i^\top$ has $\langle Su, v_0\rangle^2 \geq 1 - O(\varepsilon^{1/4}) - o(1)$.

We have little control over the basis vectors the algorithm is given. However, there is a particularly nice (albeit non-orthogonal) basis for the subspace which exposes the underlying randomness. Suppose that we are given the basis vectors $v_0, \ldots, v_d$, where v_0 is the sparse vector normalized so that $\|v_0\| = 1$, and $v_1, \ldots, v_{d-1}$ are iid samples from $\mathcal{N}(0, \frac{1}{n}\mathrm{Id}_n)$. The

following lemma shows that if the algorithm had been handed this good representation of the basis rather than an arbitrary orthogonal one, its output would be the correlated to the vector of coefficients giving of the planted sparse vector (in this case the standard basis vector e_1).

Lemma 5.3.5. *Let $v_0 \in \mathbb{R}^n$ be a unit vector. Let $v_1, \ldots, v_{d-1} \in \mathbb{R}^n$ be iid from $\mathcal{N}(0, \frac{1}{n}\operatorname{Id})$. Let a_i be the ith row of the $n \times d$ matrix $S := \begin{pmatrix} v_0 & \cdots & v_{d-1} \end{pmatrix}$. Then there is a universal constant $\varepsilon^* > 0$ so that for any $\varepsilon \leq \varepsilon^*$, so long as $d \leq n^{1/2}/\operatorname{polylog}(n)$, w.ov.p.*

$$\sum_{i=1}^n (\|a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top = \|v_0\|_4^4 \cdot e_1 e_1^\top + M\,,$$

where e_1 is the first standard basis vector and $\|M\| \leq O(\|v_0\|_4^3 \cdot n^{-1/4} + \|v_0\|_4^2 \cdot n^{-1/2} + \|v_0\|_4 \cdot n^{-3/4} + n^{-1})$.

The second ingredient we need is that the algorithm is robust to exchanging this good basis for an arbitrary orthogonal basis.

Lemma 5.3.6. *Let $v_0 \in \mathbb{R}^n$ have $\|v_0\|_4^4 \geq \frac{1}{\varepsilon n}$. Let $v_1, \ldots, v_{d-1} \in \mathbb{R}^n$ be iid from $\mathcal{N}(0, \frac{1}{n}\operatorname{Id}_n)$. Let $w_0, \ldots, w_{d-1}$ be an orthogonal basis for $\operatorname{Span}\{v_0, \ldots, v_{d-1}\}$. Let a_i be the ith row of the $n \times d$ matrix $S := \begin{pmatrix} v_0 & \cdots & v_{d-1} \end{pmatrix}$. Let a'_i be the ith row of the $n \times d$ matrix $S' := \begin{pmatrix} w_0 & \cdots & w_{d-1} \end{pmatrix}$. Let $A := \sum_i a_i a_i^\top$. Let $Q \in \mathbb{R}^{d \times d}$ be the orthogonal matrix so that $SA^{-1/2} = S'Q$, which exists since $SA^{-1/2}$ is orthogonal, and which has the effect that $a'_i = QA^{-1/2}a_i$. Then when $d \leq n^{1/2}/\operatorname{polylog}(n)$, w.ov.p.*

$$\left\| \sum_{i=1}^n (\|a'_i\|^2 - \tfrac{d}{n}) \cdot a'_i a_i'^\top - Q \left(\sum_{i=1}^n (\|a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right) Q^\top \right\| \leq O\left(\frac{1}{n}\right) + o(\|v\|_4^4)$$

Last, we will need the following fact, which follows from standard concentration. The proof is in Section 5.6.

Lemma 5.3.7. *Let $v \in \mathbb{R}^n$ be a unit vector. Let $b_1, \ldots, b_n \in \mathbb{R}^{d-1}$ be iid from $\mathcal{N}(0, \frac{1}{n}\operatorname{Id}_{d-1})$. Let $a_i \in \mathbb{R}^d$ be given by $a_i := (v(i)\ b_i)$. Then w.ov.p. $\|\sum_{i=1}^n a_i a_i^\top - \operatorname{Id}_d\| \leq \widetilde{O}(d/n)^{1/2}$. In particular, when $d = o(n)$, this implies that w.ov.p. $\|(\sum_{i=1}^n a_i a_i^\top)^{-1} - \operatorname{Id}_d\| \leq \widetilde{O}(d/n)^{1/2}$ and $\|(\sum_{i=1}^n a_i a_i^\top)^{-1/2} - \operatorname{Id}_d\| \leq \widetilde{O}(d/n)^{1/2}$.*

We are ready to prove Theorem 5.3.4.

Proof of Theorem 5.3.4. Let $b_1, \ldots, b_n$ be the rows of the matrix $S' := \begin{pmatrix} v_0 & \cdots & v_{d-1} \end{pmatrix}$. Let $B = \sum_i b_i b_i^\top$. Note that $S'B^{-1/2}$ has columns which are an orthogonal basis for $\operatorname{Span}\{w_0, \ldots, w_{d-1}\}$. Let $Q \in \mathbb{R}^{d \times d}$ be the rotation so that $S'B^{-1/2} = SQ$.

By Lemma 5.3.5 and Lemma 5.3.6, we can write the matrix $A = \sum_{i=1}^n (\|a_i\|_2^2 - \frac{d}{n}) \cdot a_i a_i^\top$ as

$$A = \|v_0\|_4^4 \cdot Q e_1 e_1^\top Q^\top + M\,,$$

where w.ov.p.

$$\|M\| \le O(\|v_0\|_4^3 \cdot n^{-1/4} + \|v_0\|_4^2 \cdot n^{-1/2} + \|v_0\|_4 \cdot n^{-3/4} + n^{-1}) + o(\|v\|_4^4)\,.$$

We have assumed that $\|v_0\|_4^4 \ge (\varepsilon n)^{-1}$, and so since A is an almost-rank-one matrix (Lemma A.1.3), the top eigenvector u of A has $\langle u, Qe_1\rangle^2 \ge 1 - O(\varepsilon^{1/4})$, so that $\langle Su, SQe_1\rangle^2 \ge 1 - O(\varepsilon^{1/4})$ by column-orthogonality of S.

At the same time, $SQe_1 = S'B^{-1/2}e_1$, and by Lemma 5.3.7, $\|B^{-1/2} - \mathrm{Id}\,\| \le \widetilde{O}(d/n)^{1/2}$ w.ov.p., so that $\langle Su, S'e_1\rangle^2 \ge \langle Su, SQe_1\rangle^2 - o(1)$. Finally, $S'e_1 = v_0$ by definition, so $\langle Su, v_0\rangle^2 \ge 1 - O(\varepsilon^{1/4}) - o(1)$. □

Algorithm Succeeds on Good Basis

We now prove Lemma 5.3.5. We decompose the matrix in question into a contribution from $\|v_0\|_4^4$ and the rest: explicitly, the decomposition is $\sum(\|a_i\|_2^2 - \frac{d}{n}) \cdot a_i a_i^\top = \sum v(i)^2 \cdot a_i a_i^\top + \sum(\|b_i\|_2^2 - \frac{d}{n} \cdot a_i a_i^\top)$. This first lemma handles the contribution from $\|v_0\|_4^4$.

Lemma 5.3.8. *Let $v \in \mathbb{R}^n$ be a unit vector. Let $b_1, \dots, b_n \in \mathbb{R}^{d-1}$ be random vectors iid from $\mathcal{N}(0, \frac{1}{n} \cdot \mathrm{Id}_{d-1})$. Let $a_i = (v(i)\ \ b_i) \in \mathbb{R}^d$. Suppose $d \le n^{1/2}/\operatorname{polylog}(n)$. Then*

$$\sum_{i=1}^n v(i)^2 \cdot a_i a_i^\top = \|v\|_4^4 \cdot e_1 e_1^\top + M'\,,$$

where $\|M'\| \le O(\|v\|_4^3 n^{-1/4} + \|v\|_4^2 n^{-1/2})$ w.ov.p..

Proof of Lemma 5.3.8. We first show an operator-norm bound on the principal submatrix $\sum_{i=1}^n v(i)^2 \cdot b_i b_i^\top$ using the truncated matrix Bernstein inequality Proposition A.3.3. First, the expected operator norm of each summand is bounded:

$$\mathbb{E}\, v(i)^2 \|b_i\|_2^2 \le (\max_j v(j)^2) \cdot O\left(\frac{d}{n}\right) \le \|v\|_4^2 \cdot O\left(\frac{d}{n}\right)\,.$$

The operator norms are bounded by constant-degree polynomials in Gaussian variables, so Lemma A.3.4 applies to truncate their tails in preparation for application of a Bernstein bound. We just have to calculate the variance of the sum, which is at most

$$\left\| \mathbb{E} \sum_{i=1}^n v(i)^4 \|b_i\|_2^2 \cdot b_i b_i^\top \right\| = \|v\|_4^4 \cdot O\left(\frac{d}{n^2}\right)\,.$$

The expectation $\mathbb{E}\sum_{i=1}^n v(i)^2 \cdot b_i b_i^\top$ is $\frac{\|v\|^2}{n} \cdot \mathrm{Id}$. Applying a matrix Bernstein bound (Proposition A.3.3) to the deviation from expectation, we get that w.ov.p.,

$$\left\| \left(\sum_{i=1}^n v(i)^2 \cdot b_i b_i^\top \right) - \frac{1}{n} \cdot \mathrm{Id} \right\| \le \|v\|_4^2 \cdot \widetilde{O}\left(\frac{d}{n}\right) \le O(\|v\|_4^2 n^{-1/2})$$

for appropriate choice of $d \le n^{-1/2}/\operatorname{polylog}(n)$. Hence, by triangle inequality, $\|\sum_{i=1}^n v(i)^2 \cdot b_i b_i^\top\| \le \|v\|_4^2 n^{-1/2}$ w.ov.p..

Using a Cauchy-Schwarz-style inequality (Lemma A.1.1) we now show that the bound on this principal submatrix is essentially enough to obtain the lemma. Let $p_i, q_i \in \mathbb{R}^d$ be given by

$$p_i \stackrel{\text{def}}{=} v_0(i) \cdot \begin{pmatrix} v_0(i) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad q_i \stackrel{\text{def}}{=} v_0(i) \cdot \begin{pmatrix} 0 \\ b_i \end{pmatrix}.$$

Then

$$\sum_{i=1}^n v(i)^2 \cdot b_i b_i^\top = \|v\|_4^4 + \sum_{i=1}^n p_i q_i^\top + q_i p_i^\top + q_i q_i^\top .$$

We have already bounded $\sum_{i=1}^n q_i q_i^\top = \sum_{i=1}^n v(i)^2 \cdot b_i b_i^\top$. At the same time, $\|\sum_{i=1}^n p_i p_i^\top\| = \|v\|_4^4$. By Lemma A.1.1, then,

$$\left\| \sum_{i=1}^n p_i q_i^\top + q_i p_i^\top \right\| \le O(\|v\|_4^3 n^{-1/4})$$

w.ov.p.. A final application of triangle inequality gives the lemma. □

Our second lemma controls the contribution from the random part of the leverage scores.

Lemma 5.3.9. *Let $v \in \mathbb{R}^n$ be a unit vector. Let $b_1, \dots, b_n \in \mathbb{R}^{d-1}$ be random vectors iid from $\mathcal{N}(0, \frac{1}{n} \cdot \mathrm{Id}_{d-1})$. Let $a_i = (v(i) \;\; b_i) \in \mathbb{R}^d$. Suppose $d \le n^{1/2}/\operatorname{polylog}(n)$. Then w.ov.p.*

$$\left\| \sum_{i=1}^n (\|b_i\|_2^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right\| \le \|v\|_4^2 \cdot O(n^{-3/4}) + \|v\|_4 \cdot O(n^{-1}) + O(n^{-1}) .$$

Proof. Like in the proof of Lemma 5.3.8, $\sum_{i=1}^n (\|b_i\|_2^2 - \frac{d}{n}) \cdot a_i a_i^\top$ decomposes into a convenient block structure; we will bound each block separately.

$$\sum_{i=1}^n (\|b_i\|_2^2 - \tfrac{d}{n}) \cdot a_i a_i^\top = \sum_{i=1}^n (\|b_i\|_2^2 - \tfrac{d}{n}) \cdot \begin{pmatrix} v(i)^2 & v(i) \cdot b_i^\top \\ v(i) \cdot b_i & b_i b_i^\top \end{pmatrix} . \tag{5.3.1}$$

In each block we can apply a (truncated) Bernstein inequality. For the large block $\sum_{i=1}^n (\|b_i\|_2^2 - \frac{d}{n}) b_i b_i^\top$, the choice $\frac{d}{n}$ ensures that $\mathbb{E}(\|b_i\|_2^2 - \frac{d}{n}) b_i b_i^\top = O(\frac{1}{n^2}) \cdot \mathrm{Id}$. The expected operator norm of each summand is small:

$$\begin{aligned}
\mathbb{E}\,\|(\|b_i\|_2^2 - \tfrac{d}{n}) b_i b_i^\top\| &= \mathbb{E}\,|(\|b_i\|_2^2 - \tfrac{d}{n})| \|b_i\|_2^2 \\
&\le (\mathbb{E}(\|b_i\|_2^2 - \tfrac{d}{n})^2)^{1/2} (\mathbb{E}\,\|b_i\|_2^4)^{1/2} \quad \text{by Cauchy-Schwarz} \\
&\le O\left(\frac{d^{1/2}}{n}\right) \cdot O\left(\frac{d}{n}\right) \quad \text{variance of } \chi^2 \text{ with } k \text{ degrees of freedom is } O(k)
\end{aligned}$$

$$= O\left(\frac{d^{3/2}}{n^2}\right).$$

The term-wise operator norms are bounded by constant-degree polynomials in Gaussian variables, so Lemma A.3.4 applies to truncate the tails of the summands in preparation for a Bernstein bound. We just have to compute the variance of the sum, which is small because we have centered the coefficients:

$$\left\|\sum_i \mathbb{E}(\|b_i\|_2^2 - \tfrac{d}{n})^2 \|b_i\|_2^2 \cdot b_i b_i^\top\right\| \le O\left(\frac{d^2}{n^3}\right)$$

by direct computation of $\mathbb{E}(\|b_i\|_2^2 - \frac{d}{n})^2\|b_i\|_2^2 b_i b_i^\top$ using Fact A.2.3. These facts together are enough to apply the matrix Bernstein inequality (Proposition A.3.3) and conclude that w.ov.p.

$$\left\|\sum_{i=1}^n (\|b_i\|_2^2 - \tfrac{d}{n}) \cdot b_i b_i^\top\right\| \le \widetilde{O}\left(\frac{d}{n^{3/2}}\right) \le O\left(\frac{1}{n}\right)$$

for appropriate choice of $d \le n/\operatorname{polylog}(n)$.

We turn to the other blocks from (5.3.1). The upper-left block contains just the scalar $\sum_{i=1}^n (\|b_i\|_2^2 - \frac{d}{n})v(i)^2$. By standard concentration each term is bounded: w.ov.p.,

$$(\|b_i\|_2^2 - \tfrac{d}{n})v(i)^2 \le (\max_i v(i)^2) \cdot \widetilde{O}\left(\frac{d^{1/2}}{n}\right) \le \|v\|_4^2 \cdot \widetilde{O}\left(\frac{d^{1/2}}{n}\right).$$

The sum has variance at most $\sum_{i=1}^n v(i)^4\,\mathbb{E}(\|b_i\|_2^2 - \frac{d}{n})^2 \le \|v\|_4^4 \cdot O(d/n^2)$. Again using Lemma A.3.4 and Proposition A.3.3, we get that w.ov.p.

$$\left|\sum_{i=1}^n (\|b_i\|_2^2 - \tfrac{d}{n})v(i)^2\right| \le \|v\|_4^2 \cdot \widetilde{O}\left(\frac{d^{1/2}}{n}\right).$$

It remains just to address the block $\sum_{i=1}^n (\|b_i\|_2^2 - \frac{d}{n})v(i) \cdot b_i$. Each term in the sum has expected operator norm at most

$$(\max_i v(i)^2)^{1/2} \cdot O\left(\frac{d}{n^{3/2}}\right) \le \|v\|_4 \cdot O\left(\frac{d}{n^{3/2}}\right) \cdot,$$

and once again the since the summands' operator norms are bounded by constant-degree polynomials of Gaussian variables Lemma A.3.4 applies to truncate their tails in preparation to apply a Bernstein bound. The variance of the sum is at most $\|v\|_2^2 \cdot O(d^2/n^3)$, again by Fact A.2.3. Finally, Lemma A.3.4 and Proposition A.3.3 apply to give that w.ov.p.

$$\left\|\sum_{i=1}^n (\|b_i\|_2^2 - \tfrac{d}{n})v(i) \cdot b_i\right\| \le \|v\|_4 \cdot \widetilde{O}\left(\frac{d}{n^{3/2}}\right) + \widetilde{O}\left(\frac{d}{n^{3/2}}\right) = \|v\|_4 \cdot n^{-1} + n^{-1}$$

for appropriate choice of $d \le n^{1/2}/\operatorname{polylog}(n)$. Putting it all together gives the lemma. □

We are now ready to prove Lemma 5.3.5

Proof of Lemma 5.3.5. We decompose $\|a_i\|_2^2 = v_0(i)^2 + \|b_i\|_2^2$ and use Lemma 5.3.8 and Lemma 5.3.9.

$$\sum_{i=1}^n (\|a_i\|_2^2 - \tfrac{d}{n}) \cdot a_i a_i^\top = \left(\sum_{i=1}^n v_0(i)^2 \cdot a_i a_i^\top\right) + \left(\sum_{i=1}^n (\|b_i\|_2^2 - \tfrac{d}{n}) \cdot a_i a_i^\top\right)$$
$$= \|v_0\|_4^4 \cdot e_1 e_1^\top + M\,,$$

where

$$\|M\| \le O(\|v_0\|_4^3 \cdot n^{-1/4} + \|v_0\|_4^2 \cdot n^{-1/2}) + O(\|v_0\|_4 \cdot n^{-1} + n^{-1})\,.$$

Since $\|v_0\|_4^4 \ge (\varepsilon n)^{-1}$, we get $\|v_0\|_4^4/\|M\| \ge \frac{1}{\varepsilon^{1/4}}$, completing the proof. □

Closeness of Input Basis and Good Basis

We turn now to the proof of Lemma 5.3.6. We recall the setting. We have two matrices: M, which the algorithm computes, and M', which is induced by a basis for the subspace which reveals the underlying randomness and which we prefer for the analysis. M' differs from M by a rotation and a basis orthogonalization step (the good basis is only almost orthogonal). The rotation is easily handled. The following lemma gives the critical fact about the orthogonalization step: orthogonalizing does not change the leverage scores too much. [6]

Lemma 5.3.10 (Restatement of Lemma 5.6.4)**.** *Let $v \in \mathbb{R}^n$ be a unit vector and let $b_1, \dots, b_n \in \mathbb{R}^{d-1}$ be iid from $\mathcal{N}(0, \frac{1}{n}\operatorname{Id}_{d-1})$. Let $a_i \in \mathbb{R}^d$ be given by $a_i := (v(i)\ b_i)$. Let $A := \sum_i a_i a_i^\top$. Let $c \in \mathbb{R}^{d-1}$ be given by $c := \sum_i v(i) b_i$. Then for every index $i \in [n]$, w.ov.p.,*

$$\left|\|A^{-1/2} a_i\|^2 - \|a_i\|^2\right| \le \widetilde{O}\left(\frac{d+\sqrt{n}}{n}\right) \cdot \|a_i\|^2\,.$$

The proof again uses standard concentration and matrix inversion formulas, and can be found in Section 5.6. We are ready to prove Lemma 5.3.6.

Proof of Lemma 5.3.6. The statement we want to show is

$$\left\|\sum_{i=1}^n (\|a_i'\|^2 - \tfrac{d}{n}) \cdot a_i' a_i'^\top - Q\left(\sum_{i=1}^n (\|a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top\right) Q^\top\right\| \le O\left(\frac{1}{n}\right) + o(\|v\|_4^4)\,.$$

Conjugating by Q and multiplying by -1 does not change the operator norm, so that this is equivalent to

$$\left\|\sum_{i=1}^n (\|a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top - Q^\top\left(\sum_{i=1}^n (\|a_i'\|^2 - \tfrac{d}{n}) \cdot a_i' a_i'^\top\right) Q\right\| \le O\left(\frac{1}{n}\right) + o(\|v\|_4^4)\,.$$

[6] Strictly speaking the good basis does not have leverage scores since it is not orthogonal, but we can still talk about the norms of the rows of the matrix whose columns are the basis vectors.

Finally, substituting $a_i' = QA^{-1/2}a_i$, and using the fact that Q is a rotation, it will be enough to show

$$\left\| \left(\sum_{i=1}^n (\|a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right) - A^{-1/2} \left(\sum_{i=1}^n (\|A^{-1/2}a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right) A^{-1/2} \right\| \leq O\left(\frac{1}{n}\right) + o(\|v\|_4^4)\,. \tag{5.3.2}$$

We write the right-hand matrix as

$$\begin{aligned} &A^{-1/2} \left(\sum_{i=1}^n (\|A^{-1/2}a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right) A^{-1/2} \\ &= A^{-1/2} \left(\sum_{i=1}^n (\|A^{-1/2}a_i\|^2 - \|a_i\|^2) \cdot a_i a_i^\top \right) A^{-1/2} + A^{-1/2} \left(\sum_{i=1}^n (\|a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right) A^{-1/2}\,. \end{aligned}$$

The first of these we observe has bounded operator norm w.ov.p.:

$$\begin{aligned} &\left\| A^{-1/2} \left(\sum_{i=1}^n (\|A^{-1/2}a_i\|^2 - \|a_i\|^2) \cdot a_i a_i^\top \right) A^{-1/2} \right\| \\ &\qquad \leq \left\| A^{-1/2} \left(\sum_{i=1}^n |\|A^{-1/2}a_i\|^2 - \|a_i\|^2| \cdot a_i a_i^\top \right) A^{-1/2} \right\| \\ &\qquad \leq \widetilde{O}\left(\frac{d+\sqrt{n}}{n}\right) \cdot \left\| \sum_{i=1}^n \|a_i\|^2 \cdot a_i a_i^\top \right\| \end{aligned}$$

where we have used Lemma 5.3.7 to find that $A^{1/2}$ is close to identity, and Lemma 5.3.10 to simplify the summands

$$\begin{aligned} &= \widetilde{O}\left(\frac{d+\sqrt{n}}{n}\right) \cdot \left(\left\| \sum_{i=1}^n v_0(i)^2 \cdot a_i a_i^\top \right\| + \left\| \sum_{i=1}^n \|b_i\|_2^2 \cdot a_i a_i^\top \right\| \right) \\ &\leq \widetilde{O}\left(\frac{d+\sqrt{n}}{n}\right) \cdot \left(O(\|v\|_4^4) + \widetilde{O}\left(\frac{d}{n}\right) \right)\,, \end{aligned}$$

using in the last step Lemma 5.3.8 and standard concentration to bound $\sum_{i=1}^n \|b_i\|_2^2 \cdot a_i a_i^\top$ (Lemma 5.3.7). Thus, by triangle inequality applied to (5.3.2), we get

$$\begin{aligned} &\left\| \left(\sum_{i=1}^n (\|a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right) - A^{-1/2} \left(\sum_{i=1}^n (\|A^{-1/2}a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right) A^{-1/2} \right\| \\ &\qquad \leq \widetilde{O}\left(\frac{d+\sqrt{n}}{n}\right) \cdot \left(O(\|v\|_4^4) + \widetilde{O}\left(\frac{d}{n}\right) \right) \end{aligned}$$

$$+ \left\| \left(\sum_{i=1}^{n} (\|a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right) - A^{-1/2} \left(\sum_{i=1}^{n} (\|a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right) A^{-1/2} \right\| .$$

Finally, since w.ov.p. $\|A^{-1/2} - \mathrm{Id}\| = \widetilde{O}(d/n)^{1/2}$, we get

$$\begin{aligned}
&\left\| \left(\sum_{i=1}^{n} (\|a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right) - A^{-1/2} \left(\sum_{i=1}^{n} (\|A^{-1/2} a_i\|^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right) A^{-1/2} \right\| \\
&\leq \widetilde{O}\left(\frac{d+\sqrt{n}}{n} \right) \cdot \left(O(\|v\|_4^4) + \widetilde{O}\left(\frac{d}{n} \right) \right) + \widetilde{O}\left(\frac{d}{n} \right)^{1/2} \cdot \left\| \sum_{i=1}^{n} (\|a_i\|_2^2 - \tfrac{d}{n}) \cdot a_i a_i^\top \right\| \\
&\leq \widetilde{O}\left(\frac{d+\sqrt{n}}{n} \right) \cdot \left(O(\|v\|_4^4) + \widetilde{O}\left(\frac{d}{n} \right) \right) + \widetilde{O}\left(\frac{d}{n} \right)^{1/2} \cdot O(\|v\|_4^4) .
\end{aligned}$$

using Lemma 5.3.5 in the last step. For appropriate choice of $d \leq n^{-1/2}/\operatorname{polylog}(n)$, this is at most $O(n^{-1}) + o(\|v\|_4^4)$. □

5.4 Overcomplete Tensor Decomposition

In this section, we give a polynomial-time algorithm for the following problem when $n \leq d^{4/3}/(\operatorname{polylog} d)$:

Problem 5.4.1. Given an order-3 tensor $\mathbf{T} = \sum_{i=1}^{n} a_i \otimes a_i \otimes a_i$, where $a_1, \ldots, a_n \in \mathbb{R}^d$ are iid vectors sampled from $\mathcal{N}(0, \frac{1}{d}\,\mathrm{Id})$, find vectors $b_1, \ldots, b_n \in \mathbb{R}^n$ such that for all $i \in [n]$,

$$\langle a_i, b_i \rangle \geq 1 - o(1) .$$

We give an algorithm that solves this problem, so long as the overcompleteness of the input tensor is bounded such that $n \ll d^{4/3}/\operatorname{polylog} d$.

Theorem 5.4.2. *Given as input the tensor $\mathbf{T} = \sum_{i=1}^{n} a_i \otimes a_i \otimes a_i$ where $a_i \sim \mathcal{N}(0, \frac{1}{d}\,\mathrm{Id}_d)$ with $d \leq n \leq d^{4/3}/\operatorname{polylog} d$,[7] there is an algorithm which may run in time $\widetilde{O}(nd^{1+\omega})$ or $\widetilde{O}(nd^{3.257})$, where d^ω is the time to multiply two $d \times d$ matrices, which with probability $1 - o(1)$ over the input $\mathbf{T}$ and the randomness of the algorithm finds unit vectors $b_1, \ldots, b_n \in \mathbb{R}^d$ such that for all $i \in [n]$,*

$$\langle a_i, b_i \rangle \geq 1 - \widetilde{O}\left(\frac{n^{3/2}}{d^2} \right) .$$

[7] The lower bound $d \leq n$ on n, is a matter of technical convenience, avoiding separate concentration analyses and arithmetic in the undercomplete ($n < d$) and overcomplete ($n \geq d$) settings. Indeed, our algorithm still works in the undercomplete setting (tensor decomposition is easier in the undercomplete setting than the overcomplete one), but here other algorithms based on local search also work [AGJ15].

We remark that this accuracy can be improved from $1 - \widetilde{O}(n^{3/2}/d^2)$ to an arbitrarily good precision using existing local search methods with local convergence guarantees—see Corollary 5.4.23.

As discussed in Section 5.2, to decompose the tensor $\sum_i a_i^{\otimes 6}$ (note we do not actually have access to this input!) there is a very simple tensor decomposition algorithm: sample a random $g \in \mathbb{R}^{d^2}$ and compute the matrix $\sum_i \langle g, a_i^{\otimes 2}\rangle (a_i a_i^\top)^{\otimes 2}$. With probability roughly $n^{-O(\varepsilon)}$ this matrix has (up to scaling) the form $(a_i a_i^\top)^{\otimes 2} + E$ for some $\|E\| \le 1 - \varepsilon$, and this is enough to recover a_i.

However, instead of $\sum_i a_i^{\otimes 6}$, we have only $\sum_{i,j}(a_i \otimes a_j)^{\otimes 3}$. Unfortunately, running the same algorithm on the latter input will not succeed. To see why, consider the extra terms $E' := \sum_{i \neq j} \langle g, a_i \otimes a_j\rangle (a_i \otimes a_j)^{\otimes 2}$. Since $|\langle g, a_i \otimes a_j\rangle| \approx 1$, it is straightforward to see that $\|E'\|_F \approx n$. Since the rank of E' is clearly d^2, even if we are lucky and all the eigenvalues have similar magnitudes, still a typical eigenvalue will be $\approx n/d \gg 1$, swallowing the $\sum_i a_i^{\otimes 6}$ term.

A convenient feature separating the signal terms $\sum_i (a_i \otimes a_i)^{\otimes 3}$ from the crossterms $\sum_{i \neq j}(a_i \otimes a_j)^{\otimes 3}$ is that the crossterms are not within the span of the $a_i \otimes a_i$. Although we cannot algorithmically access $\mathrm{Span}\{a_i \otimes a_i\}$, we have access to something almost as good: the unfolded input tensor, $T = \sum_{i \in [n]} a_i (a_i \otimes a_i)^\top$. The rows of this matrix lie in $\mathrm{Span}\{a_i \otimes a_i\}$, and so for $i \neq j$, $\|T(a_i \otimes a_i)\| \gg \|T(a_i \otimes a_j)\|$. In fact, careful computation reveals that $\|T(a_i \otimes a_i)\| \ge \widetilde{\Omega}(\sqrt{n}/d)\|T(a_i \otimes a_j)\|$.

The idea now is to replace $\sum_{i,j}\langle g, a_i \otimes a_j\rangle (a_i \otimes a_j)^{\otimes 2}$ with $\sum_{i,j}\langle g, T(a_i \otimes a_j)\rangle (a_i \otimes a_j)^{\otimes 2}$, now with $g \sim \mathcal{N}(0, \mathrm{Id}_d)$. As before, we are hoping that there is i_0 so that $\langle g, T(a_{i_0} \otimes a_{i_0})\rangle \gg \max_{j \neq i_0}\langle g, T(a_j \otimes a_j)\rangle$. But now we also require $\|\sum_{i \neq j}\langle g, T(a_i \otimes a_j)\rangle (a_i \otimes a_j)(a_i \otimes a_j)^\top\| \ll \langle g, T(a_{i_0} \otimes a_{i_0})\rangle \approx \|T(a_i \otimes a_i)\|$. If we are lucky and all the eigenvalues of this cross-term matrix have roughly the same magnitude (indeed, we will be lucky in this way), then we can estimate heuristically that

$$\begin{aligned}
&\left\|\sum_{i \neq j}\langle g, T(a_i \otimes a_j)\rangle (a_i \otimes a_j)(a_i \otimes a_j)^\top\right\| \\
&\qquad \approx \tfrac{1}{d}\left\|\sum_{i \neq j}\langle g, T(a_i \otimes a_j)\rangle (a_i \otimes a_j)(a_i \otimes a_j)^\top\right\|_F \\
&\qquad \le \tfrac{1}{d} \cdot \tfrac{\sqrt{n}}{d}|\langle g, T(a_{i_0} \otimes a_{i_0})\rangle| \left\|\sum_{i \neq j}(a_i \otimes a_j)(a_i \otimes a_j)^\top\right\|_F \\
&\qquad \le \tfrac{n^{3/2}}{d^2}|\langle g, T(a_{i_0} \otimes a_{i_0})\rangle|\,,
\end{aligned}$$

suggesting our algorithm will succeed when $n^{3/2} \ll d^2$, which is to say $n \ll d^{4/3}$.

The following theorem, which formalizes the intuition above, is at the heart of our tensor decomposition algorithm.

Theorem 5.4.3. *Let $a_1, \dots, a_n$ be independent random vectors from $\mathcal{N}(0, \frac{1}{d}\mathrm{Id}_d)$ with $d \le n \le d^{4/3}/(\mathrm{polylog}\, d)$ and let g be a random vector from $\mathcal{N}(0, \mathrm{Id}_d)$. Let $\Sigma :=$*

$\mathbb{E}_{x\sim\mathcal{N}(0,\mathrm{Id}_d)}(xx^\top)^{\otimes 2}$ *and let* $R := \sqrt{2}\cdot(\Sigma^+)^{1/2}$. *Let* $T = \sum_{i\in[n]} a_i(a_i\otimes a_i)^\top$. *Define the matrix* $M\in\mathbb{R}^{d^2\times d^2}$,

$$M = \sum_{i,j\in[n]} \langle g, T(a_i\otimes a_j)\rangle\cdot(a_i\otimes a_j)(a_i\otimes a_j)^\top.$$

With probability $1-o(1)$ *over the choice of* $a_1,\dots,a_n$, *for every* $\operatorname{polylog} d/\sqrt{d} < \varepsilon < 1$, *the spectral gap of RMR is at least* $\lambda_2/\lambda_1 \le 1-O(\varepsilon)$ *and the top eigenvector* $u\in\mathbb{R}^{d^2}$ *of RMR satisfies, with probability* $\tilde\Omega(1/n^{O(\varepsilon)})$ *over the choice of* g,

$$\max_{i\in[n]} \langle Ru, a_i\otimes a_i\rangle^2/\left(\|u\|^2\cdot\|a_i\|^4\right) \ge 1-\widetilde{O}\left(\frac{n^{3/2}}{\varepsilon d^2}\right).$$

Moreover, with probability $1-o(1)$ *over the choice of* $a_1,\dots,a_n$, *for every* $\operatorname{polylog} d/\sqrt{d} < \varepsilon < 1$ *there are events* $E_1,\dots,E_n$ *so that* $\mathbb{P}_g E_i \ge \widetilde{\Omega}(1/n^{1+O(\varepsilon)})$ *for all* $i\in[n]$ *and when* E_i *occurs,* $\langle Ru, a_i\otimes a_i\rangle^2/\|u\|^2\cdot\|a_i\|^4 \ge 1-\widetilde{O}\left(\frac{n^{3/2}}{\varepsilon d^2}\right)$.

We will eventually set $\varepsilon = 1/\log n$, which gives us a spectral algorithm for recovering a vector $(1-\widetilde{O}(n/d^{3/2}))$-correlated to some $a_i^{\otimes 2}$. Once we have a vector correlated with each $a_i^{\otimes 2}$, obtaining vectors close to the a_i is straightforward. We will begin by proving this theorem, and defer the algorithmic details to Section 5.4.

The rest of this section is organized as follows. In Section 5.4 we prove Theorem 5.4.3 using two core facts: the Gaussian vector g is closer to some a_i than to any other with good probability, and the noise term $\sum_{i\neq j}\langle g, T(a_i\otimes a_j)\rangle(a_i\otimes a_j)(a_i\otimes a_j)^\top$ is bounded in spectral norm. In Section 5.4 we prove the first of these two facts, and in Section 5.4 we prove the second. In Section 5.4, we give the full details of our tensor decomposition algorithm, then prove Theorem 5.4.2 using Theorem 5.4.3. Finally, Section 5.7 contains proofs of elementary or long-winded lemmas we use along the way.

Proof of Theorem 5.4.3

The strategy to prove Theorem 5.4.3 is to decompose the matrix M into two parts $M = M_{\text{diag}} + M_{\text{cross}}$, one formed by diagonal terms $M_{\text{diag}} = \sum_{i\in[n]}\langle g, T(a_i\otimes a_i)\rangle\cdot(a_i\otimes a_i)(a_i\otimes a_i)^\top$ and one formed by cross terms $M_{\text{cross}} = \sum_{i\neq j}\langle g, T(a_i\otimes a_j)\rangle\cdot(a_i\otimes a_j)(a_i\otimes a_j)^\top$. We will use the fact that the top eigenvector M_{diag} is likely to be correlated with one of the vectors $a_j^{\otimes 2}$, and also the fact that the spectral gap of M_{diag} is noticeable.

The following two propositions capture the relevant facts about the spectra of M_{diag} and M_{cross}, and will be proven in Section 5.4 and Section 5.4.

Proposition 5.4.4 (Spectral gap of diagonal terms)**.** *Let* $R = \sqrt{2}\cdot((\mathbb{E}(xx^\top)^{\otimes 2})^+)^{1/2}$ *for* $x\sim\mathcal{N}(0,\mathrm{Id}_d)$. *Let* $a_1,\dots,a_n$ *be independent random vectors from* $\mathcal{N}(0,\frac{1}{d}\mathrm{Id}_d)$ *with* $d\le n\le d^{2-\Omega(1)}$ *and let* $g\sim\mathcal{N}(0,\mathrm{Id}_d)$ *be independent of all the others. Let* $T := \sum_{i\in[n]} a_i(a_i\otimes a_i)^\top$. *Suppose* $M_{\text{diag}} = \sum_{i\in[n]}\langle g, Ta_i^{\otimes 2}\rangle\cdot(a_ia_i^\top)^{\otimes 2}$. *Let also* v_j *be such that* $v_jv_j^\top = \langle g, Ta_j^{\otimes 2}\rangle\cdot$

$(a_j a_j^\top)^{\otimes 2}$. *Then, with probability* $1 - o(1)$ *over* $a_1, \dots, a_n$, *for each* $\varepsilon > \operatorname{polylog} d/\sqrt{d}$ *and each* $j \in [n]$, *the event*

$$E_{j,\varepsilon} \stackrel{\text{def}}{=} \left\{ \left\| RM_{\text{diag}}R - \varepsilon \cdot Rv_j v_j^\top R \right\| \le \|RM_{\text{diag}}R\| - \left(\varepsilon - \widetilde{O}\left(\sqrt{n}/d\right)\right) \cdot \left\| Rv_j v_j^\top R \right\| \right\}$$

has probability at least $\widetilde{\Omega}(1/n^{1+O(\varepsilon)})$ *over the choice of* g.

Second, we show that when $n \ll d^{4/3}$ the spectral norm of M_{cross} is negligible compared to this spectral gap.

Proposition 5.4.5 (Bound on crossterms)**.** *Let* $a_1, \dots, a_n$ *be independent random vectors from* $\mathcal{N}(0, \frac{1}{d}\operatorname{Id}_d)$, *and let* g *be a random vector from* $\mathcal{N}(0, \operatorname{Id}_d)$. *Let* $T := \sum_{i\in[n]} a_i (a_i \otimes a_i)^\top$. *Let* $M_{\text{cross}} := \sum_{i\neq j\in[n]} \langle g, T(a_i \otimes a_j)\rangle a_i a_i^\top \otimes a_j a_j^\top$. *Suppose* $n \ge d$. *Then with w.ov.p.,*

$$\|M_{\text{cross}}\| \le \widetilde{O}\left(\frac{n^3}{d^4}\right)^{1/2}.$$

Using these two propositions we will conclude that the top eigenvector of RMR is likely to be correlated with one of the vectors $a_j^{\otimes 2}$. We also need two simple concentration bounds; we defer the proof to the appendix.

Lemma 5.4.6. *Let* $a_1, \dots, a_n$ *be independently sampled vectors from* $\mathcal{N}(0, \frac{1}{d}\operatorname{Id}_d)$, *and let* g *be sampled from* $\mathcal{N}(0, \operatorname{Id}_d)$. *Let* $T = \sum_i a_i (a_i \otimes a_i)^\top$. *Then with overwhelming probability, for every* $j \in [n]$,

$$\left| \langle g, T(a_j \otimes a_j)\rangle - \langle g, a_j\rangle \|a_j\|^4 \right| \le \widetilde{O}\left(\frac{\sqrt{n}}{d}\right).$$

Fact 5.4.7 (Simple version of Fact 5.7.1)**.** *Let* $x, y \sim \mathcal{N}(0, \frac{1}{d}\operatorname{Id})$. *With overwhelming probability,* $|1 - \|x\|^2| \le \widetilde{O}(1/\sqrt{d})$ *and* $\langle x, y\rangle^2 = \widetilde{O}(1/d)$.

As a last technical tool we will need a simple claim about the fourth moment matrix of the multivariate Gaussian:

Fact 5.4.8 (simple version of Fact 5.7.4)**.** *Let* $\Sigma = \mathbb{E}_{x\sim\mathcal{N}(0,\operatorname{Id}_d)} (xx^\top)^{\otimes 2}$ *and let* $R = \sqrt{2}\,(\Sigma^+)^{1/2}$. *Then* $\|R\| = 1$, *and for any* $v \in \mathbb{R}^d$,

$$\|R(v \otimes v)\|_2^2 = \left(1 - \tfrac{1}{d+2}\right) \cdot \|v\|^4.$$

We are prepared prove Theorem 5.4.3.

Proof of Theorem 5.4.3. Let $d \le n \le d^{4/3}/(\operatorname{polylog} d)$ for some $\operatorname{polylog} d$ to be chosen later. Let $a_1, \dots, a_n$ be independent random vectors from $\mathcal{N}(0, \frac{1}{d}\operatorname{Id}_d)$ and let $g \sim \mathcal{N}(0, \operatorname{Id}_d)$ be independent of the others. Let

$$M_{\text{diag}} := \sum_{i\in[n]} \langle g, T(a_i \otimes a_i)\rangle \cdot (a_i a_i^\top)^{\otimes 2} \quad \text{and} \quad M_{\text{cross}} := \sum_{i\neq j\in[n]} \langle g, T(a_i \otimes a_j)\rangle \cdot a_i a_i^\top \otimes a_j a_j^\top.$$

Note that $M := M_{\text{diag}} + M_{\text{cross}}$.

Proposition 5.4.5 implies that

$$\mathbb{P}\{\, \|M_{\text{cross}}\| \le \tilde{O}(n^{3/2}/d^2)\} \ge 1 - d^{-\omega(1)}\,. \tag{5.4.1}$$

Recall that $\Sigma = \mathbb{E}_{x\sim\mathcal{N}(0,\mathrm{Id}_d)}(xx^\top)^{\otimes 2}$ and $R = \sqrt{2}\cdot(\Sigma^+)^{1/2}$. By Proposition 5.4.4, with probability $1-o(1)$ over the choice of $a_1,\dots,a_n$, each of the following events $E_{j,\varepsilon}$ for $j\in[n]$ and $\varepsilon > \mathrm{polylog}(d)/\sqrt{d}$ has probability at least $\widetilde{\Omega}(1/n^{1+O(\varepsilon)})$ over the choice of g:

$$\begin{aligned} E^0_{j,\varepsilon}\colon \quad & \left\| R\left(M_{\text{diag}} - \varepsilon\langle g, Ta_j^{\otimes 2}\rangle (a_j a_j^\top)^{\otimes 2}\right) R\right\| \\ & \le \|RM_{\text{diag}}R\| - \left(\varepsilon - \widetilde{O}\left(\tfrac{n^{1/2}}{d}\right)\right)\cdot |\langle g, Ta_j{}^{\otimes 2}\rangle| \cdot \left\|Ra_j{}^{\otimes 2}\right\|^2. \end{aligned}$$

Together with (5.4.1), with probability $1-o(1)$ over the choice of $a_1,\dots,a_n$, each of the following events $E^*{}_{j,\varepsilon}$ has probability at least $\widetilde{\Omega}(1/n^{1+O(\varepsilon)}) - d^{-\omega(1)} \ge \widetilde{\Omega}(1/n^{1+O(\varepsilon)})$ over the choice of g,

$$\begin{aligned} E^*_{j,\varepsilon}\colon \quad & \left\| R\left(M - \varepsilon\langle g, Ta_j^{\otimes 2}\rangle (a_j a_j^\top)^{\otimes 2}\right) R\right\| \\ & \le \|R\cdot M\cdot R\| - \left(\varepsilon - \widetilde{O}\left(\tfrac{n^{1/2}}{d}\right)\right)\cdot |\langle g, Ta_j{}^{\otimes 2}\rangle| \cdot \left\|Ra_j{}^{\otimes 2}\right\|^2 + \widetilde{O}(n^{3/2}/d^2) \end{aligned}$$

Here, we used that $M = M_{\text{diag}} + M_{\text{cross}}$ and that $\|R\cdot M_{\text{cross}}\cdot R\| \le \|M_{\text{cross}}\|$ as $\|R\|\le 1$ (Fact 5.4.8).

By standard reasoning about the top eigenvector of a matrix with a spectral gap (recorded in Lemma A.1.3), the event $E^*{}_{j,\varepsilon}$ implies that the top eigenvector $u\in\mathbb{R}^{d^2}$ of $R\cdot M\cdot R$ satisfies

$$\left\langle u, \frac{Ra_j^{\otimes 2}}{\|Ra_j^{\otimes 2}\|}\right\rangle^2 \ge 1 - \frac{\widetilde{O}(\sqrt{n}/d)}{\varepsilon\|Ra_j^{\otimes 2}\|^2} - \frac{\widetilde{O}(n^{3/2}/d^2)}{\varepsilon\|Ra_j^{\otimes 2}\|^2|\langle g, Ta_j^{\otimes 2}\rangle|}\,.$$

Since $\|Ra_j{}^{\otimes 2}\|^2 \ge \Omega(\|a_j\|^4)$ (by Fact 5.4.8), and since $\|a_j\| \ge 1 - \widetilde{O}(1/\sqrt{d})$ (by Fact 5.4.7),

$$\ge 1 - \widetilde{O}\left(\frac{\sqrt{n}}{\varepsilon d}\right) - \frac{\widetilde{O}(n^{3/2}/d^2)}{\varepsilon\cdot|\langle g, Ta_j^{\otimes 2}\rangle|}$$

Now, by Lemma 5.4.6 we have that for all $j\in[n]$, $|\langle g, Ta_j^{\otimes 2}\rangle - \langle g, a_j\rangle\|a_j\|^4| \le \widetilde{O}(\sqrt{n}/d)$ with probability $1 - n^{-\omega(1)}$. By standard concentration (see Fact 5.7.1 for a proof) $|\langle g, a_j\rangle\|a_j\|^4 - 1| \le \widetilde{O}(1/\sqrt{d})$ for all $j\in[n]$ with probability $1-n^{-\omega(1)}$. Therefore with overwhelming probability, the final term is bounded by $\widetilde{O}(n^{3/2}/\varepsilon d^2)$. A union bound now gives the desired conclusion.

Finally, we give a bound on the spectral gap. We note that the second eigenvector w has $\langle u, w\rangle = 0$, and therefore

$$\left\langle w, \frac{Ra_j^{\otimes 2}}{\|Ra_j^{\otimes 2}\|}\right\rangle = \left\langle w, \frac{Ra_j^{\otimes 2}}{\|Ra_j^{\otimes 2}\|} - u\right\rangle \le \left\| \frac{Ra_j^{\otimes 2}}{\|Ra_j^{\otimes 2}\|} - u\right\| \le \widetilde{O}(n^{3/2}/\varepsilon d^2).$$

Thus, using our above bound on $\|R(M - \varepsilon\langle g, Ta_j^{\otimes 2}\rangle (a_j a_j^\top)^{\otimes 2})R\|$ and the concentration bounds we have already applied for $\|a_j\|$, $\langle g, Ta_j^{\otimes 2}\rangle$, and $\|Ra_j^{\otimes 2}\|$, we have that

$$\begin{aligned}\lambda_2(RMR) &= w^\top RMRw \\ &= w^\top R\left(M - \varepsilon\langle g, Ta_j^{\otimes 2}\rangle \cdot (a_j a_j^\top)^{\otimes 2}\right) Rw + \varepsilon\langle g, Ta_j^{\otimes 2}\rangle \langle w, Ra_j^{\otimes 2}\rangle^2 \\ &\le \left\|R\left(M - \varepsilon\langle g, Ta_j^{\otimes 2}\rangle \cdot (a_j a_j^\top)^{\otimes 2}\right) R\right\| + \widetilde{O}(n^{3/2}/\varepsilon d^2) \\ &\le 1 - \widetilde{O}(\varepsilon) + \widetilde{O}(n^{3/2}/\varepsilon d^2)\,.\end{aligned}$$

We conclude that the above events also imply that $\lambda_2(RMR)/\lambda_1(RMR) \le 1 - O(\varepsilon)$. □

Spectral Gap for Diagonal Terms: Proof of Proposition 5.4.4

We now prove that the signal matrix, when preconditioned by R, has a noticeable spectral gap:

Proposition (Restatement of Proposition 5.4.4)**.** *Let $R = \sqrt{2}\cdot((\mathbb{E}(xx^\top)^{\otimes 2})^+)^{1/2}$ for $x \sim \mathcal{N}(0, \mathrm{Id}_d)$. Let $a_1, \dots, a_n$ be independent random vectors from $\mathcal{N}(0, \frac{1}{d}\mathrm{Id}_d)$ with $d \le n \le d^{2-\Omega(1)}$ and let $g \sim \mathcal{N}(0, \mathrm{Id}_d)$ be independent of all the others. Let $T := \sum_{i\in[n]} a_i(a_i \otimes a_i)^\top$. Suppose $M_{\mathrm{diag}} = \sum_{i\in[n]}\langle g, Ta_i^{\otimes 2}\rangle \cdot (a_i a_i^\top)^{\otimes 2}$. Let also v_j be such that $v_j v_j^\top = \langle g, Ta_j^{\otimes 2}\rangle \cdot (a_j a_j^\top)^{\otimes 2}$. Then, with probability $1 - o(1)$ over $a_1, \dots, a_n$, for each $\varepsilon > \mathrm{polylog}\, d/\sqrt{d}$ and each $j \in [n]$, the event*

$$E_{j,\varepsilon} \stackrel{\text{def}}{=} \left\{\left\|RM_{\mathrm{diag}}R - \varepsilon \cdot Rv_j v_j^\top R\right\| \le \|RM_{\mathrm{diag}}R\| - \left(\varepsilon - \widetilde{O}\left(\sqrt{n}/d\right)\right) \cdot \left\|Rv_j v_j^\top R\right\|\right\}$$

has probability at least $\widetilde{\Omega}(1/n^{1+O(\varepsilon)})$ over the choice of g.

The proof has two parts. First we show that for $a_1, \dots, a_n \sim \mathcal{N}(0, \mathrm{Id}_d)$ the matrix $P := \sum_{i\in[n]}(a_i a_i^\top)^{\otimes 2}$ has tightly bounded spectral norm when preconditioned with R: more precisely, that $\|RPR\| \le 1 + \widetilde{O}(n/d^{3/2})$.

Lemma 5.4.9. *Let $a_1, \dots, a_n \sim \mathcal{N}(0, \frac{1}{d}\mathrm{Id}_d)$ be independent random vectors with $d \le n$. Let $R := \sqrt{2}\cdot((\mathbb{E}(aa^\top)^{\otimes 2})^+)^{1/2}$ for $a \sim \mathcal{N}(0, \mathrm{Id}_d)$. For $S \subseteq [n]$, let $P_S = \sum_{i\in S}(a_i {a_i}^\top)^{\otimes 2}$ and let Π_S be the projector into the subspace spanned by $\{Ra_i^{\otimes 2} \mid i \in S\}$. Then, with probability $1 - o(1)$ over the choice of $a_1, \dots, a_n$,*

$$\forall S \subseteq [n]. \quad \left(1 - \widetilde{O}(n/d^{3/2})\right) \cdot \Pi_S \preceq RP_SR \preceq \left(1 + \widetilde{O}(n/d^{3/2})\right) \cdot \Pi_S\,.$$

Remark 5.4.10. In [GM15, Lemma 5] a similar lemma to this one is proved in the context of the SoS proof system. However, since Ge and Ma leverage the full power of the SoS algorithm their proof goes via a spectral bound on a different (but related) matrix. Since our algorithm avoids solving an SDP we need a bound on this matrix in particular.

The proof of Lemma 5.4.9 proceeds by standard spectral concentration for tall matrices with independent columns (here the columns are $Ra_i^{\otimes 2}$). The arc of the proof is straightforward but it involves some bookkeeping; we have deferred it to Section 5.7.

We also need the following lemma on the concentration of some scalar random variables involving R; the proof is straightforward by finding the eigenbasis of R and applying standard concentration, and it is deferred to the appendix.

Lemma 5.4.11. *Let $a_1, \dots, a_n \sim \mathcal{N}(0, \frac{1}{d}\mathrm{Id}_d)$. Let Σ, R be as in Fact 5.4.8. Let $u_i = a_i \otimes a_i$. With overwhelming probability, every $j \in [n]$ satisfies $\sum_{i \neq j} \langle u_j, R^2 u_i\rangle^2 = \widetilde{O}(n/d^2)$ and $|1 - \|Ru_j\|^2| \leq \widetilde{O}(1/\sqrt{d})$.*

The next lemma is the linchpin of the proof of Proposition 5.4.4: one of the inner products $\langle g, Ta_j{}^{\otimes 2}\rangle$, is likely to be a $\approx (1 + 1/\log(n))$-factor larger than the maximum of the inner products $\langle g, Ta_i{}^{\otimes 2}\rangle$ over $i \neq j$. Together with standard linear algebra these imply that the matrix $M_{\mathrm{diag}} = \sum_{i \in [n]} \langle g, Ta_i^{\otimes 2}\rangle (a_i a_i^\top)^{\otimes 2}$ has top eigenvector highly correlated or anti-correlated with some a_i.

Lemma 5.4.12. *Let $a_1, \dots, a_n \in \mathbb{R}^d$ be independent random vectors from $\mathcal{N}(0, \frac{1}{d}\mathrm{Id}_d)$, and let g be a random vector from $\mathcal{N}(0, \mathrm{Id}_d)$. Let $T = \sum_{i \in [n]} a_i (a_i \otimes a_i)^\top$. Let $\varepsilon > 0$ and $j \in [n]$. Then with overwhelming probability over $a_1, \dots, a_n$, the following event $\hat{E}_{j,\varepsilon}$ has probability $1/n^{1+O(\varepsilon)+\widetilde{O}(1/\sqrt{d})}$ over the choice of g,*

$$\hat{E}_{j,\varepsilon} = \left\{ \langle g, Ta_j^{\otimes 2}\rangle \geq (1+\varepsilon)(1 - \widetilde{O}(1/\sqrt{d})) \cdot \max_{i \neq j} \left|\langle g, Ta_i^{\otimes 2}\rangle\right| \right\}.$$

Now we can prove Proposition 5.4.4.

Proof of Proposition 5.4.4. Let $u_i := a_i^{\otimes 2}$. Fix $j \in [n]$. We begin by showing a lower bound on the spectral norm $\|RM_{\mathrm{diag}}R\|$.

$$\begin{aligned}
\|RM_{\mathrm{diag}}R\| &= \max_{\|v\|=1} |\langle v, RM_{\mathrm{diag}}Rv\rangle| \\
&\geq \frac{\langle Ru_j, (RM_{\mathrm{diag}}R)Ru_j\rangle}{\|Ru_j\|^2} \\
&= \frac{1}{\|Ru_j\|^2}\left(\langle g, Tu_j\rangle \|Ru_j\|^4 + \langle Ru_j, \sum_{i \neq j} \langle g, T_i\rangle Ru_i u_i^\top R \cdot Ru_j\rangle\right)
\end{aligned}$$

From Lemma 5.4.12, the random vector g is closer to Tu_j than to all Tu_i for $i \neq j, i \in [n]$ with reasonable probability. More concretely there is some polylog d so that as long as $\varepsilon > \mathrm{polylog}\, d/\sqrt{d}$ there is some $\alpha = \Theta(\varepsilon)$ with $1 - \varepsilon = 1/[(1+\alpha)(1 - \widetilde{O}(d^{-1/2}))]$ so that with w.ov.p. over $a_1, \dots, a_n$ the following event (a direct consequence of $\hat{E}_{j,\varepsilon}$) has probability $\widetilde{\Omega}(1/n^{1+O(\alpha)+\widetilde{O}(d^{-1/2})}) = \widetilde{\Omega}(1/n^{1+O(\varepsilon)})$ over g:

$$-(1-\varepsilon)|\langle g, Tu_j\rangle| \cdot \left(\sum_{i \neq j} Ru_i u_i^\top R\right) \preceq \sum_{i \neq j} \langle g, Tu_i\rangle \cdot Ru_i u_i^\top R \preceq (1-\varepsilon)|\langle g, Tu_j\rangle| \cdot \left(\sum_{i \neq j} Ru_i u_i^\top R\right). \tag{5.4.2}$$

When (5.4.2) occurs,

$$\begin{aligned}
\|RM_{\mathrm{diag}}R\| &\geq \frac{1}{\|Ru_j\|^2}\left(|\langle g, Tu_j\rangle|\|Ru_j\|^4 - (1-\varepsilon)|\langle g, Tu_j\rangle|\langle Ru_j, \sum_{i\neq j} Ru_iu_i^\top R\cdot Ru_j\rangle\right)\\
&= \frac{|\langle g, Tu_j\rangle|}{\|Ru_j\|^2}\left(\|Ru_j\|^4 - (1-\varepsilon)\sum_{i\neq j}\langle u_j, R^2u_i\rangle^2\right)\\
&\geq \frac{|\langle g, Tu_j\rangle|\left(1-\widetilde{O}(1/\sqrt{d}) - (1-\varepsilon)\widetilde{O}(n/d^2)\right)}{1+\widetilde{O}(1/\sqrt{d})} \quad \text{w.ov.p. over } a_1,\dots,a_n \text{ (Lemma 5.4.11)}\\
&\geq |\langle g, Tu_j\rangle|\cdot(1-\eta_{\mathrm{norm}})\,, \qquad (5.4.3)
\end{aligned}$$

where we have chosen some $0 \leq \eta_{\mathrm{norm}} \leq \widetilde{O}(1/\sqrt{d}) + \widetilde{O}(n/d^2)$ (since for any $x \in \mathbb{R}$, $(1+x)(1-x) \leq 1$).

Next we exhibit an upper bound on $\|RM_{\mathrm{diag}}R - \varepsilon\langle g, Tu_j\rangle Ru_ju_j^\top R\|$. Again when (5.4.2) occurs,

$$\begin{aligned}
&\|RM_{\mathrm{diag}}R - \varepsilon\langle g, Tu_j\rangle Ru_ju_j^\top R\| \qquad (5.4.4)\\
&= \left\|(1-\varepsilon)\langle g, Tu_j\rangle Ru_ju_j^\top R + \sum_{i\neq j}\langle g, Tu_i\rangle Ru_iu_i^\top R\right\|\\
&\leq (1-\varepsilon)|\langle g, Tu_j\rangle|\left\|\sum_{i\in[n]} Ru_iu_i^\top R\right\| \quad \text{when (5.4.2) occurs}\\
&\leq (1-\varepsilon)|\langle g, Tu_j\rangle|(1+\widetilde{O}(n/d^{1.5})) \quad \text{w.p. } 1-o(1) \text{ over } a_1,\dots,a_n \text{ by Lemma 5.4.9}\\
&\leq (1-\varepsilon)|\langle g, Tu_j\rangle|(1+\eta_{\mathrm{gap}}) \qquad (5.4.5)
\end{aligned}$$

where we have chosen some $0 \leq \eta_{\mathrm{gap}} \leq \widetilde{O}(n/d^{1.5})$.

Putting together (5.4.3) and (5.4.5) with our bounds on η_{norm} and η_{gap} and recalling the conditions on (5.4.2), we have shown that

$$\begin{aligned}
\mathop{\mathbb{P}}_{a_1,\dots,a_n}\Big\{\mathop{\mathbb{P}}_{g}\{\|RM_{\mathrm{diag}}R - \varepsilon\langle g, Tu_j\rangle Ru_ju_j^\top R\| \leq \|RM_{\mathrm{diag}}R\| - (\varepsilon - \widetilde{O}(\sqrt{n}/d))\cdot|\langle g, Tu_j\rangle|\cdot\|Ru_j\|^2\}\\
\geq \widetilde{\Omega}(1/n^{1+O(\varepsilon)})\Big\} \geq 1-o(1)\,.
\end{aligned}$$

This concludes the argument. □

We now turn to proving that with reasonable probability, g is closer to some $Ta_j^{\otimes 2}$ than all others.

Proof of Lemma 5.4.12. To avoid proliferation of indices, without loss of generality fix $j = 1$. We begin by expanding the expression $\langle g, Ta_i^{\otimes 2}\rangle$,

$$\langle g, Ta_i^{\otimes 2}\rangle = \sum_{\ell\in[n]}\langle g, a_\ell\rangle\langle a_\ell, a_i\rangle^2 = \|a_i\|^4\langle g, a_i\rangle + \sum_{\ell\neq i}\langle g, a_\ell\rangle\langle a_\ell, a_i\rangle^2.$$

The latter sum is bounded by

$$\left|\sum_{\ell\neq i}\langle g,a_\ell\rangle\langle a_\ell,a_i\rangle^2\right| \le \widetilde{O}\left(\frac{\sqrt{n}}{d}\right),$$

with overwhelming probability for all i and choices of g; this follows from a Bernstein bound, given in Lemma 5.4.6.

For ease of notation, let $\hat{a}_i \stackrel{\text{def}}{=} a_i/\|a_i\|_2$. We conclude from Fact 5.4.7 that with overwhelming probability, $1-\widetilde{O}(1/\sqrt{d}) \le \|a_i\|_2 \le 1+\widetilde{O}(1/\sqrt{d})$ for all $i\in[n]$. Thus $\|a_i\|_2$ is roughly equal for all i, and we may direct our attention to $\langle g,\hat{a}_i\rangle$.

Let $\mathcal{G}_1$ be the event that $\sqrt{2\alpha}\log^{1/2} n \le |\langle g,\hat{a}_1\rangle| \le d^{1/4}$ for some $\alpha \le d^{1/2-\Omega(1)}$ to be chosen later. We note that $\langle g,\hat{a}_1\rangle$ is distributed as a standard Gaussian, and that g is independent of $a_1,\dots,a_n$. Thus, we can use standard tail estimates on univariate Gaussians (Lemma A.2.1) to conclude that

$$\mathbb{P}\left(|\langle g,\hat{a}_1\rangle| \ge \sqrt{2\alpha}\log^{1/2} n\right) = \widetilde{\Theta}(n^{-\alpha}) \quad\text{and}\quad \mathbb{P}\left(|\langle g,\hat{a}_1\rangle| \ge d^{1/4}\right) = \Theta\left(\frac{\exp(-\sqrt{d}/2)}{d^{1/4}}\right).$$

So by a union bound, $\mathbb{P}(\mathcal{G}_1) \ge \widetilde{\Omega}(n^{-\alpha}) - O(e^{-d^{1/2}/3}) = \widetilde{\Omega}(n^{-\alpha})$.

Now, we must obtain an estimate for the probability that all other inner products with g are small. Let $\mathcal{G}_{i>1}$ be the event that $|\langle g,\hat{a}_i\rangle| \le \sqrt{(2+\rho)}\log^{1/2} n$ for all $i\in[n], i>1$ and for some ρ to be chosen later. We will show that conditioned on $\mathcal{G}_1$, $\mathcal{G}_{i>1}$ occurs with probability $1-O(n^{1-(2+\rho)/2})$. Define $g_1 := \langle g,\hat{a}_1\rangle\hat{a}_1$ to be the component of g parallel to a_1, let $g_\perp := g-g_1$ be the component of g orthogonal to $\hat{a}_1$, and similarly let $\hat{a}_2^\perp,\dots,\hat{a}_n^\perp$ be the components of $\hat{a}_2,\dots,\hat{a}_n$ orthogonal to a_1. Because $g_\perp$ is independent of g_1, even conditioned on $\mathcal{G}_1$ we may apply the standard tail bound for univariate Gaussians (Lemma A.2.1), concluding that for all $i>1$,

$$\mathbb{P}\left(|\langle g_\perp,\hat{a}_i\rangle| \ge \sqrt{(2+\rho)}\log^{1/2} n \,\middle|\, \mathcal{G}_1\right) = \widetilde{\Theta}(n^{-(2+\rho)/2}).$$

Thus, a union bound over $i\neq 1$ allows us to conclude that conditioned on $\mathcal{G}_1$, with probability $1-\widetilde{O}(n^{-\rho/2})$ every $i\in[n]$ with $i>1$ has $|\langle g_\perp,\hat{a}_i^\perp\rangle| \le \sqrt{(2+\rho)}\log^{1/2} n$.

On the other hand, let $\hat{a}_2^\parallel,\dots,\hat{a}_n^\parallel$ be the components of the $\hat{a}_i$ parallel to $\hat{a}_1$. We compute the projection of $\hat{a}_i$ onto $\hat{a}_1$. With overwhelming probability,

$$\begin{aligned}
\langle \hat{a}_1,\hat{a}_i\rangle &= \frac{\langle a_1,a_i\rangle}{\|a_1\|_2\cdot\|a_i\|_2}\\
&= (1\pm\widetilde{O}(1/\sqrt{d}))\cdot\langle a_1,a_i\rangle \quad \text{w.ov.p. by } \|a_i\|,\|a_1\| = 1\pm\widetilde{O}(1/\sqrt{d}) \text{ (Fact 5.4.7)}\\
&= (1\pm\widetilde{O}(1/\sqrt{d}))\cdot\widetilde{O}(1/\sqrt{d}) \quad \text{w.ov.p. by } \langle a_1,a_i\rangle = \widetilde{O}(1/\sqrt{d}) \text{ (Fact 5.4.7)},
\end{aligned}$$

Thus w.ov.p.,

$$\langle g_1,\hat{a}_i^\parallel\rangle = \langle g,\hat{a}_1\rangle\cdot\langle \hat{a}_1,\hat{a}_i\rangle \;\le\; \langle g,\hat{a}_1\rangle\cdot\widetilde{O}(1/\sqrt{d}),$$

for all $i \in [n]$. Now we can analyze $\mathcal{G}_{i>1}$. Taking a union bound over the overwhelmingly probable events (including $\|a_i\| \leq 1 + \widetilde{O}(1/\sqrt{d})$) and the event that $\langle g_\perp, a_i \rangle$ is small for all i, we have that with probability $1 - \widetilde{O}(n^{-\rho/2})$, for every $i \in [n]$ with $i > 1$,

$$\begin{aligned} |\langle g, \hat{a}_i \rangle| &\leq |\langle g_\perp, \hat{a}_i \rangle| + |\langle g_1, \hat{a}_i \rangle| \\ &\sqrt{(2+\rho)} \log^{1/2} n + \widetilde{O}(1/\sqrt{d}) \cdot \langle g, \hat{a}_1 \rangle \\ &\leq \sqrt{(2+\rho)} \log^{1/2} n + \widetilde{O}(1/d^{1/4}). \end{aligned}$$

We conclude that

$$\begin{aligned} \mathbb{P}\left(\mathcal{G}_1, \mathcal{G}_{i>1}\right) &= \mathbb{P}\left(\mathcal{G}_{i>1} | \mathcal{G}_1\right) \cdot \mathbb{P}\left(\mathcal{G}_1\right) \\ &\geq (1 - O(n^{-\rho/2})) \cdot \widetilde{\Omega}(n^{-\alpha}) \end{aligned}$$

Setting $\rho = 2\frac{\log\log n}{\log n}$ and $\alpha = (1+\varepsilon)^2(1 + \log\log n/\log n + \widetilde{O}(1/\sqrt{d}))$, the conclusion follows. □

Bound for Cross Terms: Proof of Proposition 5.4.5

We proceed to the bound on the cross terms M_{cross}.

Proposition (Restatement of Proposition 5.4.5)**.** *Let $a_1, \ldots, a_n$ be independent random vectors from $\mathcal{N}(0, \frac{1}{d}\operatorname{Id}_d)$, and let g be a random vector from $\mathcal{N}(0, \operatorname{Id}_d)$. Let $T := \sum_{i\in[n]} a_i(a_i \otimes a_i)^\top$. Let $M_{\text{cross}} := \sum_{i\neq j\in[n]} \langle g, T(a_i \otimes a_j)\rangle a_i a_i^\top \otimes a_j a_j^\top$. Suppose $n \geq d$. Then with w.ov.p.,*

$$\|M_{\text{cross}}\| \leq \widetilde{O}\left(\frac{n^3}{d^4}\right)^{1/2}.$$

The proof will use two iterations of Matrix Rademacher bounds. The first step will be to employ a classical decoupling inequality that has previously been used in a tensor decomposition context [GM15].

Theorem 5.4.13 (Special Case of Theorem 1 in [dlPMS95])**.** *Let $\{s_i\}, \{t_i\}$ be independent iid sequences of random signs. Let $\{M_{ij}\}$ be a family of matrices. There is a universal constant C so that for every $t > 0$,*

$$\mathbb{P}\left(\left\|\sum_{i\neq j} s_i s_j M_{ij}\right\|_{op} > t\right) \leq C \cdot \mathbb{P}\left(C\left\|\sum_{i\neq j} s_i t_j M_{ij}\right\|_{op} > t\right).$$

Once the simplified cross terms are decoupled, we can use a matrix Rademacher bound on one set of signs.

Theorem 5.4.14 (Adapted from Theorem 4.1.1 in [Tro12][8]). *Consider a finite sequence $\{M_i\}$ of fixed $m \times m$ Hermitian matrices. Let s_i be a sequence of independent sign variables. Let $\sigma^2 := \|\sum_i M_i^2\|$. Then for every $t \geq 0$,*

$$\mathbb{P}\left(\left\|\sum_i s_i M_i\right\|_{op} \geq t\right) \leq 2m \cdot e^{-t^2/2\sigma^2}\,.$$

Also,

$$\mathbb{E}\left\|\sum_i s_i M_i\right\| \leq \sqrt{8\sigma^2 \log d}\,.$$

Corollary 5.4.15. *Let $s_1, \ldots, s_n$ be independent signs in $\{-1, 1\}$. Let $A_1, \ldots, A_n$ and $B_1, \ldots, B_n$ be Hermetian matrices. Then w.ov.p.,*

$$\left\|\sum_i s_i \cdot A_i \otimes B_i\right\| \leq \widetilde{O}\left(\max_i \|B_i\| \cdot \left\|\sum_i A_i^2\right\|^{1/2}\right)\,.$$

Proof. We use a matrix Rademacher bound and standard manipulations:

$$\begin{aligned}
\left\|\sum_i s_i \cdot A_i \otimes B_i\right\| &\overset{\text{w.ov.p.}}{\leq} \widetilde{O}\left(\left\|\sum_i A_i^2 \otimes B_i^2\right\|\right)^{1/2} \\
&\leq \widetilde{O}\left(\left\|\sum_i \|B_i\|^2 \cdot (A_i^2 \otimes \mathrm{Id})\right\|\right)^{1/2} \qquad \text{since } A_i^2 \text{ is PSD for all } i \\
&\leq \widetilde{O}\left(\max_i \|B_i\|^2 \cdot \left\|\sum_i A_i^2\right\|\right)^{1/2} \qquad \text{since } A_i^2 \otimes \mathrm{Id} \text{ is PSD for all } i\,. \square
\end{aligned}$$

We also need a few further concentration bounds on matrices which will come up as parts of M_{cross}. These can be proved by standard inequalities for sums of independent matrices.

Lemma 5.4.16 (Restatement of Fact 5.7.2 and Lemma 5.7.3)**.** *Let $a_1, \ldots, a_n$ be independent from $\mathcal{N}(0, \frac{1}{d}\mathrm{Id}_d)$ with $n \geq d\,\mathrm{polylog}(d)$. With overwhelming probability, $\widetilde{\Omega}(n/d) \cdot \mathrm{Id} \preceq \sum_{i\in[n]} a_i a_i^\top \preceq \widetilde{O}(n/d) \cdot \mathrm{Id}$. Additionally, if $g \sim \mathcal{N}(0, \mathrm{Id}_d)$ is independent of the rest, for every $j \in [n]$ w.ov.p.*

$$\left\|\sum_{\substack{i\in[n]\\ i\neq j}} \langle g, a_i\rangle \|a_i\|^2 \langle a_i, a_j\rangle \cdot a_i a_i^\top\right\| \leq \widetilde{O}(n/d^2)^{1/2}\,.$$

[8] We remark that Tropp's bound is phrased in terms of $\lambda_{max}\sum_i s_i M_i$. Since $\lambda_{max}\sum_i s_i M_i = \lambda_{min}\sum_i -s_i M_i$, and the distribution of $s_i M_i$ is negation-invariant, the result we state here follows from an easy union bound.

Proof of Proposition 5.4.5. We expand M_{cross}:

$$\begin{aligned} M_{\text{cross}} &= \sum_{i\neq j} \langle g, T(a_i \otimes a_j)\rangle \cdot a_i a_i^\top \otimes a_j a_j^\top \\ &= \sum_{i\neq j} \left(\sum_{\ell\in[n]} \langle a_\ell, a_i\rangle \langle a_\ell, a_j\rangle \langle g, a_\ell\rangle \right) \cdot a_i a_i^\top \otimes a_j a_j^\top . \end{aligned}$$

Since the joint distribution of $(a_1, \dots, a_n)$ is identical to that of $(s_1 a_1, \dots, s_n a_n)$, this is distributed identically to

$$M'_{\text{cross}} = \sum_{\ell\in[n]} \sum_{i\neq j} s_i s_j s_\ell \langle g, a_\ell\rangle \langle a_\ell, a_i\rangle \langle a_\ell, a_j\rangle \cdot a_i a_i^\top \otimes a_j a_j^\top ,$$

(where we have also swapped the sums over ℓ and $i \neq j$). We split M'_{cross} into M_{diff}, for which $i \neq \ell$ and $j \neq \ell$, and M_{same}, for which $\ell = i$ or $\ell = j$, and bound the norm of each of these sums separately. We begin with M_{same}.

$$M_{\text{same}} \overset{\text{def}}{=} \sum_{i\neq j} s_i^2 s_j \langle g, a_i\rangle \langle a_i, a_i\rangle \langle a_i, a_j\rangle \cdot a_i a_i^\top \otimes a_j a_j^\top + \sum_{i\neq j} s_j^2 s_i \langle g, a_j\rangle \langle a_j, a_j\rangle \langle a_i, a_j\rangle \cdot a_i a_i^\top \otimes a_j a_j^\top .$$

By a union bound and an application of the triangle inequality it will be enough to show that just one of these two sums is $\widetilde{O}(n^3/d^4)^{1/2}$ w.ov.p.. We rewrite the left-hand one:

$$\sum_{i\neq j} s_i^2 s_j \langle g, a_i\rangle \langle a_i, a_i\rangle \langle a_i, a_j\rangle \cdot a_i a_i^\top \otimes a_j a_j^\top = \sum_{j\in[n]} s_j a_j a_j^\top \otimes \left(\sum_{i\neq j} \langle g, a_i\rangle \|a_i\|^2 \langle a_i, a_j\rangle \cdot a_i a_i^\top \right) .$$

Define

$$M_j \overset{\text{def}}{=} \sum_{i\neq j} \langle g, a_i\rangle \|a_i\|^2 \langle a_i, a_j\rangle \cdot a_i a_i^\top$$

so that now we need to bound $\sum_{j\in[n]} s_j a_j a_j^\top \otimes M_j$. By Corollary 5.4.15,

$$\begin{aligned} \left\| \sum_{j\in[n]} s_j a_j a_j^\top \otimes M_j \right\| &\overset{\text{w.ov.p.}}{\leq} \widetilde{O}(\max_j \|M_j\|) \cdot \widetilde{O}\left(\left\| \sum_{j\in[n]} \|a_j\|^2 a_j a_j^\top \right\|^{1/2} \right) \\ &\leq \widetilde{O}(\max_j \|M_j\|) \cdot \max_j \|a_j\| \cdot \widetilde{O}\left(\left\| \sum_{j\in[n]} a_j a_j^\top \right\|^{1/2} \right) \end{aligned}$$

In Lemma 5.4.16, we bound $\max_j \|M_j\| \leq \widetilde{O}(n/d^2)^{1/2}$ w.ov.p. using a matrix Bernstein inequality. Combining this bound with the concentration of $\|a_j\|$ around 1 (Fact 5.4.7), we obtain

$$\overset{\text{w.ov.p.}}{\leq} \widetilde{O}(n/d^2)^{1/2} \cdot \widetilde{O}(n/d)^{1/2}$$

$$= \widetilde{O}(n/d^{1.5}) .$$

Having finished with M_{same}, we turn to M_{diff}.

$$\begin{aligned}\|M_{\text{diff}}\| &= \left\| \sum_{\ell \neq i \neq j} s_i s_j s_\ell \langle g, a_\ell \rangle \langle a_\ell, a_i \rangle \langle a_\ell, a_j \rangle \cdot a_i a_i^\top \otimes a_j a_j^\top \right\| \\ &= \left\| \sum_{\ell} s_\ell \langle g, a_\ell \rangle \left(\sum_{i \neq \ell} s_i \langle a_\ell, a_i \rangle a_i a_i^\top \otimes \left(\sum_{j \neq \ell, i} s_j \langle a_\ell, a_j \rangle a_j a_j^\top \right) \right) \right\| .\end{aligned}$$

Letting $t_1, \dots, t_n$ and $r_1, \dots, r_n$ be independent uniformly random signs, by Theorem 5.4.13, it will be enough to bound the spectral norm after replacing the second and third occurrences of s_i for t_i and r_i. To this end, we define

$$M'_{\text{diff}} \stackrel{\text{def}}{=} \sum_{\ell} s_\ell \langle g, a_\ell \rangle \left(\sum_{i \neq \ell} t_i \langle a_\ell, a_i \rangle a_i a_i^\top \otimes \left(\sum_{j \neq \ell, i} r_j \langle a_\ell, a_j \rangle a_j a_j^\top \right) \right) .$$

Let

$$N_\ell \stackrel{\text{def}}{=} \sum_{i \neq \ell} t_i \langle a_\ell, a_i \rangle a_i a_i^\top \otimes \left(\sum_{j \neq \ell, i} r_j \langle a_\ell, a_j \rangle a_j a_j^\top \right)$$

so that we are to bound $\left\| \sum_{\ell \in [n]} s_\ell \langle g, a_\ell \rangle \cdot N_\ell \right\|$. By a matrix Rademacher bound and elementary manipulations,

$$\begin{aligned}\left\| \sum_{\ell \in [n]} s_\ell \langle g, a_\ell \rangle \cdot N_\ell \right\| &\stackrel{\text{w.ov.p.}}{\leq} \widetilde{O}\left(\left\| \sum_{\ell \in [n]} \langle g, a_\ell \rangle^2 \cdot N_\ell^2 \right\| \right)^{1/2} \\ &\leq \widetilde{O}(\sqrt{n}) \cdot \max_{\ell \in [n]} |\langle g, a_\ell \rangle| \cdot \max_{\ell \in [n]} \|N_\ell\| \\ &\stackrel{\text{w.ov.p.}}{\leq} \widetilde{O}(\sqrt{n}) \cdot \max_{\ell \in [n]} \|N_\ell\| \quad \text{since } |\langle g, a_i \rangle| \leq \widetilde{O}(1) \text{ (Fact 5.4.7)} .\end{aligned}$$

The rest of the proof is devoted to bounding $\|N_\ell\|$.

We start with Corollary 5.4.15 to get

$$\|N_\ell\| \stackrel{\text{w.ov.p.}}{\leq} \widetilde{O}\left(\left(\max_i \left\| \sum_{j \neq \ell, i} r_j \langle a_\ell, a_j \rangle \cdot a_j a_j^\top \right\| \right) \cdot \left\| \sum_{i \neq \ell} \langle a_\ell, a_i \rangle^2 \|a_i\|^2 \cdot a_i a_i^\top \right\|^{1/2} \right)$$

We use a matrix Rademacher bound for the left-hand matrix,

$$\left\| \sum_{j \neq \ell, i} r_j \langle a_\ell, a_j \rangle \cdot a_j a_j^\top \right\| \stackrel{\text{w.ov.p.}}{\leq} \widetilde{O}\left(\left\| \sum_{j \neq \ell, i} \langle a_\ell, a_j \rangle^2 \|a_j\|^2 \cdot a_j a_j^\top \right\| \right)^{1/2}$$

$$\leq \widetilde{O}\left(\max_{j\neq\ell}\langle a_\ell, a_j\rangle^2\|a_j\|^2\left\|\sum_j a_j a_j^\top\right\|\right)^{1/2}$$
$$\overset{\text{w.ov.p.}}{\leq} \widetilde{O}\left(\frac{\sqrt{n}}{d}\right),$$

where we have used that $\langle a_\ell, a_i\rangle^2$ concentrates around $\frac{1}{d}$ (Fact 5.4.7), that $\|a_i\|^2$ concentrates around 1 (Fact 5.4.7), and that $\left\|\sum_i a_i a_i^\top\right\|$ concentrates around $\frac{n}{d}$ (Lemma 5.4.16) within logarithmic factors all with overwhelming probability.

For the right-hand matrix, we use the fact that the summands are PSD to conclude that

$$\left\|\sum_{i\neq\ell}\langle a_\ell, a_i\rangle^2\|a_i\|^2\cdot a_i a_i^\top\right\| \leq \max_{i\neq\ell}\langle a_\ell, a_i\rangle^2\|a_i\|^2\cdot\left\|\sum_{i\neq\ell} a_i a_i^\top\right\|$$
$$\overset{\text{w.ov.p.}}{\leq} \widetilde{O}(1/d)\cdot\widetilde{O}(n/d),$$

using the same concentration facts as earlier.

Putting these together, w.ov.p.

$$\|N_\ell\| \leq \widetilde{O}(\sqrt{n}/d)\cdot\widetilde{O}(\sqrt{n}/d) = \widetilde{O}(n/d^2).$$

Now we are ready to make the final bound on M'_{diff}. With overwhelming probability,

$$\|M'_{\text{diff}}\| \leq \widetilde{O}(\sqrt{n})\cdot\max_{\ell\in[n]}\|N_\ell\| \leq \widetilde{O}(n^3/d^4)^{1/2}$$

and hence by Theorem 5.4.13, $\|M_{\text{diff}}\| \leq \widetilde{O}(\sqrt{n})\cdot\max_{\ell\in[n]}\|N_\ell\| \leq \widetilde{O}(n^3/d^4)^{1/2}$ w.ov.p..

Finally, by triangle inequality and all our bounds thus far, w.ov.p.

$$\|M_{\text{cross}}\| \leq \|M_{\text{same}}\| + \|M_{\text{diff}}\| \leq \widetilde{O}(n/d^{1.5}) + \widetilde{O}(n^3/d^4)^{1/2} \leq \widetilde{O}(n^3/d^4)^{1/2}. \qquad \square$$

Full Algorithm and Proof of Theorem 5.4.2

In this subsection we give the full details of our tensor decomposition algorithm. As discussed above, the algorithm proceeds by constructing a random matrix from the input tensor, then computing and post-processing its top eigenvector.

Spectral Tensor Decomposition (One Attempt)

This is the main subroutine of our algorithm—we will run it $\widetilde{O}(n)$ times to recover all of the components $a_1, \dots, a_n$.

Algorithm 5.4.17. Input: $\mathbf{T} = \sum_{i=1}^n a_i \otimes a_i \otimes a_i$. Goal: Recover a_i for some $i \in [n]$.

- Compute the matrix unfolding $T \in \mathbb{R}^{d^2 \times d}$ of $\mathbf{T}$. Then compute a 3-tensor $\mathbf{S} \in \mathbb{R}^{d^2 \times d^2 \times d^2}$ by starting with the 6-tensor $\mathbf{T} \otimes \mathbf{T}$, permuting indices, and flattening to a 3-tensor. Apply T in one mode of $\mathbf{S}$ to obtain $\mathbf{M} \in \mathbb{R}^{d \otimes d^2 \otimes d^2}$, so that:

$$T = \sum_{i \in [n]} a_i (a_i \otimes a_i)^\top, \qquad \mathbf{S} = \mathbf{T}^{\otimes 2} = \sum_{i,j=1}^n (a_i \otimes a_j)^{\otimes 3},$$

$$\mathbf{M} = \mathbf{S}(T, \mathrm{Id}_{d^2}, \mathrm{Id}_{d^2}) = \sum_{i,j \in [n]} T(a_i \otimes a_j) \otimes (a_i \otimes a_j) \otimes (a_i \otimes a_j).$$

- Sample a vector $g \in \mathbb{R}^d$ with iid standard gaussian entries. Evaluate $\mathbf{M}$ in its first mode in the direction of g to obtain $M \in \mathbb{R}^{d^2 \times d^2}$:

$$M := \mathbf{M}(g, \mathrm{Id}_{d^2}, \mathrm{Id}_{d^2}) = \sum_{i,j \in [n]} \langle g, T(a_i \otimes a_j) \rangle \cdot (a_i \otimes a_j)(a_i \otimes a_j)^\top.$$

- Let $\Sigma \stackrel{\text{def}}{=} \mathbb{E}[(aa^\top)^{\otimes 2}]$ for $a \sim \mathcal{N}(0, \mathrm{Id}_d)$. Let $R \stackrel{\text{def}}{=} \sqrt{2} \cdot (\Sigma^+)^{1/2}$. Compute the top eigenvector $u \in \mathbb{R}^{d^2}$ of RMR, and reshape Ru to a matrix $U \in \mathbb{R}^{d \times d}$.
- For each of the signings of the top 2 unit left (or right) singular vectors $\pm u_1, \pm u_2$ of U, check if $\sum_{i \in [n]} \langle a_i, \pm u_j \rangle^3 \geq 1 - c(n,d)$ where $c(n,d) = \Theta(n/d^{3/2})$ is an appropriate threshold. If so, output $\pm u_j$. Otherwise output nothing.

Theorem 5.4.3 gets us most of the way to the correctness of Algorithm 5.4.17, proving that the top eigenvector of the matrix RMR is correlated with some $a_i^{\otimes 2}$ with reasonable probability. We need a few more ingredients to prove Theorem 5.4.2. First, we need to show a bound on the runtime of Algorithm 5.4.17.

Lemma 5.4.18. *Algorithm 5.4.17 can be implemented in time $\widetilde{O}(d^{1+\omega}) \leq \widetilde{O}(d^{3.3729})$, where d^ω is the runtime for multiplying two $d \times d$ matrices.*

Proof. To run the algorithm, we only require access to power iteration using the matrix RMR. We first give a fast implementation for power iteration with the matrix M, and handle the multiplications with R separately.

Consider a vector $v \in \mathbb{R}^{d^2}$, and a random vector $g \sim \mathcal{N}(0, \mathrm{Id}_d)$, and let $V, G \in \mathbb{R}^{d \times d}$ be the reshapings of v and gT respectively into matrices. Call $\mathbf{T}_v = \mathbf{T}(\mathrm{Id}_d, V, G)$, where we

have applied V and G in the second and third modes of $\mathbf{T}$, and call T_v the reshaping of $\mathbf{T}_v$ into a $d \times d^2$ matrix. We have that

$$T_v = \sum_{i \in [n]} a_i (Va_i \otimes Ga_i)^\top .$$

We show that the matrix-vector multiply Mv can be computed as a flattening of the following product:

$$\begin{aligned} T_v T^\top &= \left(\sum_{i \in [n]} a_i (Va_i \otimes Ga_i)^\top \right) \left(\sum_{j \in [n]} (a_j \otimes a_j) a_j^\top \right) \\ &= \sum_{i,j \in [n]} \langle a_j, Va_i \rangle \cdot \langle a_j, Ga_i \rangle \cdot a_i a_j^\top \\ &= \sum_{i,j \in [n]} \langle a_i \otimes a_j, v \rangle \cdot \langle gT, a_i \otimes a_j \rangle \cdot a_i a_j^\top . \end{aligned}$$

Flattening $T_v T^\top$ from a $d \times d$ matrix to a vector $v_{TT} \in \mathbb{R}^{d^2}$, we have that

$$v_{TT} = \sum_{i,j \in [n]} \langle gT, a_i \otimes a_j \rangle \cdot \langle a_i \otimes a_j, v \rangle \cdot a_i \otimes a_j = Mv .$$

So we have that Mv is a flattening of the product $T_v T^\top$, which we will compute as a proxy for computing Mv via direct multiplication.

Computing $T_v = \mathbf{T}(\mathrm{Id}, V, G)$ can be done with two matrix multiplication operations, both times multiplying a $d^2 \times d$ matrix with a $d \times d$ matrix. Computing $T_v T^\top$ is a multiplication of a $d \times d^2$ matrix by a $d^2 \times d$ matrix. Both these steps may be done in time $O(d^{1+\omega})$, by regarding the $d \times d^2$ matrices as block matrices with blocks of size $d \times d$. The asymptotically fastest known algorithm for matrix multiplication gives a time of $O(d^{3.3729})$ [Gal14].

Now, to compute the matrix-vector multiply $RMRu$ for any vector $u \in \mathbb{R}^{d^2}$, we may first compute $v = Ru$, perform the operation Mv in time $O(d^{1+\omega})$ as described above, and then again multiply by R. The matrix R is sparse: it has $O(d)$ entries per row (see Fact 5.7.4), so the multiplication Ru requires time $O(d^3)$.

Performing the update $RMRv$ a total of $O(\log^2 n)$ times is sufficient for convergence, as we have that with reasonable probability, the spectral gap $\lambda_2(RMR)/\lambda_1(RMR) \leq 1 - O(\frac{1}{\log n})$, as a result of applying Theorem 5.4.3 with the choice of $\varepsilon = O(\frac{1}{\log n})$.

Finally, checking the value of $\sum_i \langle a_i, x \rangle^3$ requires $O(d^3)$ operations, and we do so a constant number of times, once for each of the signings of the top 2 left (or right) singular vectors of U. □

Next, we need to show that given u with $\langle Ru, a_i \otimes a_i \rangle^2 \geq (1 - \widetilde{O}(n^{3/2}/\varepsilon d^2)) \cdot \|u\|^2 \cdot \|a_i\|^4$ we can actually recover the tensor component a_i. Here Algorithm 5.4.17 reshapes Ru to a $d \times d$ matrix and checks the top two left- or right-singular vectors; the next lemma shows one of these singular vectors must be highly correlated with a_i. (The proof is deferred to Section A.1.)

Lemma 5.4.19. *Let $M \in \mathbb{R}^{d^2 \times d^2}$ be a symmetric matrix with $\|M\| \le 1$, and let $v \in \mathbb{R}^d$ and $u \in \mathbb{R}^{d^2}$ be vectors. Furthermore, let U be the reshaping of the vector $Mu \in \mathbb{R}^{d^2}$ to a matrix in $\mathbb{R}^{d \times d}$. Fix $c > 0$, and suppose that $\langle Mu, v \otimes v\rangle^2 \ge c^2 \cdot \|u\|^2 \cdot \|v\|^4$. Then U has some left singular vector a and some right singular vector b such that*

$$|\langle a, v\rangle|, |\langle b, v\rangle| \ge c \cdot \|v\| .$$

Furthermore, for any $0 < \alpha < 1$, there are a', b' among the top $\lfloor \frac{1}{\alpha c^2} \rfloor$ singular vectors of U with

$$|\langle a', v\rangle|, |\langle b', v\rangle| \ge \sqrt{1-\alpha} \cdot c \cdot \|v\| .$$

If $c \ge \sqrt{\frac{1}{2}(1+\eta)}$ for some $\eta > 0$, then a, b are amongst the top $\lfloor \frac{(1+\eta)}{\eta c^2} \rfloor$ singular vectors.

Since here $c^2 = 1 - o(1)$, we can choose $\eta = 1 - o(1)$ and check only the top 2 singular vectors.

Next, we must show how to choose the threshold $c(n, d)$ so that a big enough value $\sum_{i \in [n]} \langle a_i, u_j\rangle^3$ is ensures that u_j is close to a tensor component. The proof is at the end of this section. (A very similar fact appears in [GM15]. We need a somewhat different parameterization here, but we reuse many of their results in the proof.)

Lemma 5.4.20. *Let $T = \sum_{i \in [n]} a_i \otimes a_i \otimes a_i$ for normally distributed vectors $a_i \sim \mathcal{N}(0, \frac{1}{d} \mathrm{Id}_d)$. For all $0 < \gamma, \gamma' < 1$,*

1. *With overwhelming probability, for every $v \in \mathbb{R}^d$ such that $\sum_{i \in [n]} \langle a_i, v\rangle^3 \ge 1 - \gamma$,*

$$\max_{i \in [n]} |\langle a_i, v\rangle| \ge 1 - O(\gamma) - \widetilde{O}(n/d^{3/2}) .$$

2. *With overwhelming probability over $a_1, \dots, a_n$ if $v \in \mathbb{R}^d$ with $\|v\| = 1$ satisfies $\langle v, a_j\rangle \ge 1 - \gamma'$ for some j then $\sum_i \langle a_i, v\rangle^3 \ge 1 - O(\gamma') - \widetilde{O}(n/d^{3/2})$.*

We are now ready to prove Theorem 5.4.2.

Proof of Theorem 5.4.2. By Theorem 5.4.3, with probability $1 - o(1)$ over $a_1, \dots, a_n$ there are events $E_1, \dots, E_n$ so that $\mathbb{P}_g(E_i) \ge \widetilde{O}(1/n^{1+O(\varepsilon)})$ such that when event E_i occurs the top eigenvector u of RMR satisfies

$$\frac{\langle Ru, a_i \otimes a_i\rangle^2}{\|u\|^2 \cdot \|a_i\|^4} \ge 1 - \widetilde{O}\left(\frac{n^{3/2}}{\varepsilon d^2}\right) .$$

For a particular sample $g \sim \mathcal{N}(0, \mathrm{Id}_d)$, let u_g be this eigenvector.

The algorithm is as follows. Sample $g_1, \dots, g_r \sim \mathcal{N}(0, \mathrm{Id}_d)$ independently for some r to be chosen later. Compute $Ru_{g_1}, \dots, Ru_{g_r}$, reshape each to a $d \times d$ matrix, and compute its singular value decomposition. This gives a family of (right) singular vectors $v_1, \dots, v_{dr}$. For each, evaluate $\sum_i \langle a_i, v_j\rangle^3$. Let $c(n, d)$ be a threshold to be chosen later. Initialize $S \subset \mathbb{R}^d$ to

the empty set. Examining each $1 \le j \le dr$ in turn, add v_j to S if $\sum_i \langle a_i, v_j\rangle^3 \ge 1 - c(n,d)$ and for every v already in S, $\langle v, v_j\rangle^2 \le 1/2$. Output the set S.

Choose $\varepsilon = 1/\log n$. By Lemma 5.4.19, when E_i occurs for g_j one of $v \in \{\pm v_{jr}, \dots, \pm v_{(j+1)r}\}$ has $\langle v, a_i\rangle \ge (1 - \widetilde{O}(n^{3/2}/d^2))(\|u_j\|^2 \cdot \|a_j\|^4)$. Then by Lemma 5.4.20, when E_i occurs for g_j, this v we will have $\sum_i \langle a_i, \pm v\rangle^3 \ge 1 - \widetilde{O}(n/d^{3/2})$. Choose $c(n,d) = \tilde{\Theta}(n^{3/2}/d^2)$ so that when E_i occurs for g_j, so long as it has not previously occurred for some $j' < j$, the algorithm adds $\pm v$ to S.

The events $E_i^{(t)}$ and $E_i^{(t')}$ are independent for any two executions of the algorithm t and t' and have probability $\widetilde{\Omega}(1/n)$. Thus, after $r = \widetilde{O}(n)$ executions of the algorithm, with high probability for every $i \in [n]$ there is $j \in [r]$ so that E_i occurs for g_j. Finally, by Lemma 5.4.20, the algorithm can never add to S a vector which is not $(1 - \widetilde{O}(n/d^{3/2}))$-close to some a_i. □

It just remains to prove Lemma 5.4.20.

Proof of Lemma 5.4.20. We start with the first claim. By [GM15, Lemma 2, (proof of) Lemma 8, Theorem 4.2], the following inequalities all hold w.ov.p..

$$\sum_{i\in n} \langle a_i, x\rangle^4 \le 1 + \widetilde{O}(n/d^{3/2}) \quad \text{for all } \|x\| = 1\,, \tag{5.4.6}$$

$$\sum_{i\in[n]} \langle a_i, x\rangle^6 \ge 1 - O\left(\sum_{i\in[n]} \langle a_i, x\rangle^3 - 1\right) - \widetilde{O}(n/d^{3/2}) \quad \text{for all } \|x\| = 1\,, \tag{5.4.7}$$

$$\left|\sum_{i\in[n]} \langle a_i, x\rangle^3\right| \le 1 + \widetilde{O}(n/d^{3/2}) \quad \text{for all } \|x\| = 1\,. \tag{5.4.8}$$

To begin,

$$\sum_{i\in[n]} \langle a_i, v\rangle^6 \le \left(\max_{i\in[n]} \langle a_i, v\rangle^2\right) \cdot \left(\sum_{i\in[n]} \langle a_i, v\rangle^4\right).$$

By (5.4.6), this implies

$$\max_{i\in[n]} \langle a_i, v\rangle^2 \ge (1 - \widetilde{O}(n/d^{3/2})) \cdot \sum_{i\in[n]} \langle v, a_i\rangle^6. \tag{5.4.9}$$

Now combining (5.4.7) with (5.4.9) we have

$$\max_{i\in[n]} \langle a_i, v\rangle^2 \ge (1 - \widetilde{O}(n/d^{3/2})) \cdot (1 - O(1 - \sum_i \langle a_i, v\rangle^3) - \widetilde{O}(n/d^{3/2}))\,.$$

Together with (5.4.8) this concludes the of the first claim.

For the second claim, we note that by (5.4.8), and homogeneity, $|\sum_{i\neq j}\langle a_i,x\rangle^3| \le \|x\|^3(1+\widetilde{O}(n/d^{3/2})$ w.ov.p.. We write $v = \langle a_j,x\rangle a_j + x^\perp$, where $\langle x^\perp, a_j\rangle = 0$. Now we expand

$$\begin{aligned}\sum_i \langle a_i,v\rangle^3 &\ge (1-\gamma')^3 + \sum_{i\neq j}\langle\langle a_j,x\rangle a_j + x^\perp, a_i\rangle^3 \\ &= (1-\gamma')^3 + \sum_{i\neq j}\langle a_j,x\rangle^3\langle a_j,a_i\rangle^3 + 3\langle a_j,x\rangle^2\langle a_j,a_i\rangle^2\langle x^\perp,a_i\rangle \\ &\quad + 3\langle a_j,x\rangle\langle a_j,a_i\rangle\langle x^\perp a_i\rangle^2 + \langle x^\perp,a_i\rangle^3 .\end{aligned}$$

We estimate each term in the expansion:

$$\left|\sum_{i\neq j}\langle a_j,x\rangle^3\langle a_j,a_i\rangle^3\right| \le |\langle a_j,x\rangle^3|\sum_{i\neq j}|\langle a_j,a_i\rangle|^3 \le \widetilde{O}\left(\frac{n}{d^{3/2}}\right)$$

w.ov.p. by Cauchy-Schwarz and standard concentration.

$$\left|\sum_{i\neq j}\langle a_j,x\rangle^2\langle a_j,a_i\rangle^2\langle x^\perp,a_i\rangle\right| \le \left(\sum_{i\neq j}\langle a_j,x\rangle^4\langle a_j,a_i\rangle^4\right)^{1/2}\left(\sum_{i\neq j}\langle x^\perp,a_i\rangle^2\right)^{1/2}$$

by Cauchy-Schwarz

$$\le O(\sqrt{n})\cdot\max_{i\neq j}\langle a_j,a_i\rangle^2\cdot\widetilde{O}\left(\frac{n}{d}\right)^{1/2}$$

w.ov.p. by standard concentration.

$$\le \widetilde{O}\left(\frac{n}{d^{3/2}}\right)$$

w.ov.p. by standard concentration

$$\left|\sum_{i\neq j}\langle a_j,x\rangle\langle a_j,a_i\rangle\langle x^\perp,a_i\rangle^2\right| \le O(1)\cdot\max_{i\neq j}|\langle a_j,a_i\rangle|\cdot\sum_{i\neq j}\langle x^\perp,a_i\rangle^2$$

w.ov.p. by standard concentration

$$\le \widetilde{O}\left(\frac{1}{\sqrt{d}}\right)\cdot\widetilde{O}\left(\frac{n}{d}\right)$$

w.ov.p. by standard concentration

$$\le \widetilde{O}\left(\frac{n}{d^{3/2}}\right)$$

$$\left|\sum_{i\neq j}\langle x^\perp,a_i\rangle^3\right| \le \gamma' + \widetilde{O}\left(\frac{n}{d^{3/2}}\right) \quad \text{w.ov.p. by (5.4.8) and homogeneity}.$$

Now we estimate

$$\sum_i\langle a_i,v\rangle^3 \ge (1-\gamma')^3 + \sum_{i\neq j}\langle a_i,x\rangle^3 \ge (1-\gamma)^3 - \gamma' - \widetilde{O}(n/d^{3/2}) \ge 1 - O(\gamma') - \widetilde{O}(n/d^{3/2}).$$

since $\gamma' < 1$. □

Boosting Accuracy with Local Search

We remark that Algorithm 5.4.17 may be used in conjunction with a local search algorithm to obtain much greater guarantees on the accuracy of the recovered vectors. Previous progress on the tensor decomposition problem has produced iterative algorithms that provide local convergence guarantees given a good enough initialization, but which leave the question of how to initialize the procedure up to future work, or up to the specifics of an implementation. In this context, our contribution can be seen as a general method of obtaining good initializations for these local iterative procedures.

In particular, Anandkumar et al. [AGJ15] give an algorithm that combines tensor power iteration and a form of coordinate descent, which when initialized with the output of Algorithm 5.4.17, achieves a linear convergence rate to the true decomposition within polynomial time.

Theorem 5.4.21 (Adapted from Theorem 1 in [AGJ15])**.** *Given a rank-n tensor $\mathbf{T} = \sum_i a_i \otimes a_i \otimes a_i$ with random Gaussian components $a_i \sim \mathcal{N}(0, \frac{1}{d}\mathrm{Id}_d)$. There is a constant $c > 0$ so that if a set of unit vectors $\{x_i \in \mathbb{R}^d\}_i$ satisfies*

$$\langle x_i, a_i \rangle \geq 1 - c, \quad \forall i \in [n],$$

then there exists a procedure which with overwhelming probability over $\mathbf{T}$ and for any $\varepsilon > 0$, recovers a set of vectors $\{\hat{a}_i\}$ such that

$$\langle \hat{a}_i, a_i \rangle \geq 1 - \varepsilon, \quad \forall i \in [n],$$

in time $O(\mathrm{poly}(d) + nd^3 \log \varepsilon)$.

Remark 5.4.22. Theorem 1 of Anandkumar et al. is stated for random asymmetric tensors, but the adaptation to symmetric tensors is stated in equations (14) and (27) in the same paper.

The theorem of Anandkumar et al. allows for a perturbation tensor Φ, which is just the zero tensor in our setting. Additionally, the weight ratios specifying the weight of each rank-one component in the input tensor are $w_{max} = w_{min} = 1$. Lastly, the initialization conditions are given in terms of the distance between the intialization vectors and the true vectors $|x_i - a_i|$, which is related to our measure of closeness $\langle x_i, a_i \rangle$ by the equation $|x_i - a_i|^2 = |x_i|^2 + |a_i|^2 - 2\langle x_i, a_i \rangle$.

The linear convergence guarantee is stated in Lemma 12 of Anandkumar et al.

Corollary 5.4.23 (Corollary of Theorem 5.4.2)**.** *Given as input the tensor $\mathbf{T} = \sum_{i=1}^n a_i \otimes a_i \otimes a_i$ where $a_i \sim \mathcal{N}(0, \frac{1}{d}\mathrm{Id}_d)$ with $d \leq n \leq d^{4/3}/\operatorname{polylog} d$, there is a polynomial-time algorithm which with probability $1 - o(1)$ over the input $\mathbf{T}$ and the algorithm randomness finds unit vectors $\hat{a}_1, \ldots, \hat{a}_n \in \mathbb{R}^d$ such that for all $i \in [n]$,*

$$\langle \hat{a}_i, a_i \rangle \geq 1 - O\left(2^{-n}\right).$$

Proof. We repeatedly invoke Algorithm 5.4.17 until we obtain a full set of n vectors as characterized by Theorem 5.4.2. Apply Theorem 5.4.21 to the recovered set of vectors until the desired accuracy is obtained. □

5.5 Tensor Principal Component Analysis

The Tensor PCA problem in the spiked tensor model is similar to the setting of tensor decomposition, but here the goal is to recover a single large component with all smaller components of the tensor regarded as random noise.

Problem 5.5.1 (Tensor PCA in the Order-3 Spiked Tensor Model)**.** Given an input tensor $\mathbf{T} = \tau \cdot v^{\otimes 3} + \mathbf{A}$, where $v \in \mathbb{R}^n$ is an arbitrary unit vector, $\tau \geq 0$ is the signal-to-noise ratio, and $\mathbf{A}$ is a random noise tensor with iid standard Gaussian entries, recover the signal v approximately.

Using the partial trace method, we give the first linear-time algorithm for this problem that recovers v for signal-to-noise ratio $\tau = O(n^{3/4}/\operatorname{poly}\log n)$. In addition, the algorithm requires only $O(n^2)$ auxiliary space (compared to the input size of n^3) and uses only one non-adaptive pass over the input.

Spiked Tensor Model

This spiked tensor model (for general order-k tensors) was introduced by Montanari and Richard [RM14], who also obtained the first algorithms to solve the model with provable statistical guarantees. Subsequently, the SoS approach was applied to the model to improve the signal-to-noise ratio required for odd-order tensors [HSS15]; for 3-tensors reducing the requirement from $\tau = \Omega(n)$ to $\tau = \Omega(n^{3/4}\log(n)^{1/4})$.

Using the linear-algebraic objects involved in the analysis of the SoS relaxation, the previous work has also described algorithms with guarantees similar to those of the SoS SDP relaxation, while requiring only nearly subquadratic or linear time [HSS15].

The algorithm here improves on the previous results by use of the partial trace method, simplifying the analysis and improving the runtime by a factor of $\log n$.

Linear-Time Algorithm

Linear-Time Algorithm for Tensor PCA

Algorithm 5.5.2. Input: $\mathbf{T} = \tau \cdot v^{\otimes 3} + \mathbf{A}$. Goal: Recover v' with $\langle v, v' \rangle \geq 1 - o(1)$.

- Compute the partial trace $M := \operatorname{Tr}_{\mathbb{R}^n} \sum_i T_i \otimes T_i \in \mathbb{R}^{n \times n}$, where T_i are the first-mode slices of $\mathbf{T}$.
- Output the top eigenvector v' of M.

Theorem 5.5.3. *When $\mathbf{A}$ has iid standard Gaussian entries and $\tau \geq Cn^{3/4}\log(n)^{1/2}/\varepsilon$ for some constant C, Algorithm 5.5.2 recovers v' with $\langle v, v' \rangle \geq 1 - O(\varepsilon)$ with high probability over $\mathbf{A}$.*

Theorem 5.5.4. *Algorithm 5.5.2 can be implemented in linear time and sublinear space.*

These theorems are proved by routine matrix concentration results, showing that in the partial trace matrix, the signal dominates the noise.

To implement the algorithm in linear time it is enough to show that this (sublinear-sized) matrix has constant spectral gap; then a standard application of the matrix power method computes the top eigenvector.

Lemma 5.5.5. *For any v, with high probability over $\mathbf{A}$, the following occur:*

$$\left\|\sum_i \mathrm{Tr}(A_i)\cdot A_i\right\| \le O(n^{3/2}\log^2 n)$$
$$\left\|\sum_i v(i)\cdot A_i\right\| \le O(\sqrt{n}\log n)$$
$$\left\|\sum_i \mathrm{Tr}(A_i)v(i)\cdot vv^\top\right\| \le O(\sqrt{n}\log n)\,.$$

The proof may be found in Appendix 5.8.

Proof of Theorem 5.5.3. We expand the partial trace $\mathrm{Tr}_{\mathbb{R}^n}\sum_i T_i\otimes T_i$.

$$\begin{aligned}\mathrm{Tr}_{\mathbb{R}^n}\sum_i T_i\otimes T_i &= \sum_i \mathrm{Tr}(T_i)\cdot T_i\\ &= \sum_i \mathrm{Tr}(\tau\cdot v(i)vv^\top + A_i)\cdot(\tau\cdot v(i)vv^\top + A_i)\\ &= \sum_i (\tau v(i)\|v\|^2 + \mathrm{Tr}(A_i))\cdot(\tau\cdot v(i)vv^\top + A_i)\\ &= \tau^2 vv^\top + \tau\left(\sum_i v(i)\cdot A_i + \sum_i \mathrm{Tr}(A_i)v(i)vv^\top\right) + \sum_i \mathrm{Tr}(A_i)\cdot A_i\,.\end{aligned}$$

Applying Lemma 5.5.5 and the triangle inequality, we see that

$$\left\|\tau\left(\sum_i v(i)\cdot A_i + \sum_i \mathrm{Tr}(A_i)v(i)vv^\top\right) + \sum_i \mathrm{Tr}(A_i)\cdot A_i\right\| \le O(n^{3/2}\log n)$$

with high probability. Thus, for appropriate choice of $\tau = \Omega(n^{3/4}\sqrt{(\log n)/\varepsilon})$, the matrix $\mathrm{Tr}_{\mathbb{R}^n}\sum_i T_i\otimes T_i$ is close to rank one, and the result follows by standard manipulations. □

Proof of Theorem 5.5.4. Carrying over the expansion of the partial trace from above and setting $\tau = O(n^{3/4}\sqrt{(\log n)/\varepsilon})$, the matrix $\mathrm{Tr}_{\mathbb{R}^n}\sum_i T_i\otimes T_i$ has a spectral gap ratio equal to $\Omega(1/\varepsilon)$ and so the matrix power method finds the top eigenvector in $O(\log(n/\varepsilon))$ iterations. This matrix has dimension $n\times n$, so a single iteration takes $O(n^2)$ time, which is sublinear in the input size n^3. Finally, to construct $\mathrm{Tr}_{\mathbb{R}^n}\sum_i T_i\otimes T_i$ we use

$$\mathrm{Tr}_{\mathbb{R}^n}\sum_i T_i\otimes T_i = \sum_i \mathrm{Tr}(T_i)\cdot T_i$$

and note that to construct the right-hand side it is enough to examine each entry of $\mathbf{T}$ just $O(1)$ times and perform $O(n^3)$ additions. At no point do we need to store more than $O(n^2)$ matrix entries at the same time. □

5.6 Concentration Bounds for Planted Sparse Vector in Random Linear Subspace

Proof of Lemma 5.3.7. Let $c := \sum_{i=1}^n v(i)b_i$. The matrix in question has a nice block structure:

$$\sum_{i=1}^n a_i a_i^\top = \begin{pmatrix} \|v\|_2^2 & c^\top \\ c & \sum_{i=1}^n b_i b_i^\top \end{pmatrix}.$$

The vector c is distributed as $\mathcal{N}(0, \frac{1}{n}\operatorname{Id}_{d-1})$ so by standard concentration has $\|c\| \le \widetilde{O}(d/n)^{1/2}$ w.ov.p.. By assumption, $\|v\|_2^2 = 1$. Thus by triangle inequality w.ov.p.

$$\left\|\sum_{i=1}^n a_i a_i^\top - \operatorname{Id}_d\right\| \le \widetilde{O}\left(\frac{d}{n}\right)^{1/2} + \left\|\sum_{i=1}^n b_i b_i^\top - \operatorname{Id}_{d-1}\right\|.$$

By [Ver12, Corollary 5.50] applied to the subgaussian vectors nb_i, w.ov.p.

$$\left\|\sum_{i=1}^n b_i b_i^\top - \operatorname{Id}_{d-1}\right\| \le O\left(\frac{d}{n}\right)^{1/2}$$

and hence $\|\sum_{i=1}^n a_i a_i^\top - \operatorname{Id}_d\| \le \widetilde{O}(d/n)^{1/2}$ w.ov.p.. This implies $\|(\sum_{i=1}^n a_i a_i^\top)^{-1} - \operatorname{Id}_d\| \le \widetilde{O}(d/n)^{1/2}$ and $\|(\sum_{i=1}^n a_i a_i^\top)^{-1/2} - \operatorname{Id}_d\| \le \widetilde{O}(d/n)^{1/2}$ when $d = o(n)$ by the following facts applied to the eigenvalues of $\sum_{i=1}^n a_i a_i^\top$. For $0 \le \varepsilon < 1$,

$$\begin{aligned} (1+\varepsilon)^{-1} &= 1 - O(\varepsilon) \quad \text{and} \quad (1-\varepsilon)^{-1} = 1 + O(\varepsilon), \\ (1+\varepsilon)^{-1/2} &= 1 - O(\varepsilon) \quad \text{and} \quad (1-\varepsilon)^{-1/2} = 1 + O(\varepsilon). \end{aligned}$$

These are proved easily via the identity $(1+\varepsilon)^{-1} = \sum_{k=1}^\infty \varepsilon^k$ and similar. □

Orthogonal Subspace Basis

Lemma 5.6.1. *Let $a_1, \dots, a_n \in \mathbb{R}^d$ be independent random vectors from $\mathcal{N}(0, \frac{1}{n}\operatorname{Id})$ with $d \le n$ and let $A = \sum_{i=1}^n a_i a_i^\top$. Then for every unit vector $x \in \mathbb{R}^d$, with overwhelming probability $1 - d^{-\omega(1)}$,*

$$\left|\langle x, A^{-1}x\rangle - \|x\|^2\right| \le \tilde{O}\left(\frac{d+\sqrt{n}}{n}\right) \cdot \|x\|^2.$$

Proof. Let $x \in \mathbb{R}^d$. By scale invariance, we may assume $\|x\| = 1$.

By standard matrix concentration bounds, the matrix $B = \mathrm{Id} - A$ has spectral norm $\|B\| \le \tilde{O}(d/n)^{1/2}$ w.ov.p. [Ver12, Corollary 5.50]. Since $A^{-1} = (\mathrm{Id} - B)^{-1} = \sum_{k=0}^{\infty} B^k$, the spectral norm of $A^{-1} - \mathrm{Id} - B$ is at most $\sum_{k=2}^{\infty} \|B\|^k$ (whenever the series converges). Hence, $\|A^{-1} - \mathrm{Id} - B\| \le \tilde{O}(d/n)$ w.ov.p..

It follows that it is enough to show that $|\langle x, Bx\rangle| \le \tilde{O}(1/n)^{1/2}$ w.ov.p.. The random variable $n - n\langle x, Bx\rangle = \sum_{i=1}^{n} \langle \sqrt{n} \cdot a_i, x\rangle^2$ is χ^2-distributed with n degrees of freedom. Thus, by standard concentration bounds, $n|\langle x, Bx\rangle| \le \tilde{O}(\sqrt{n})$ w.ov.p. [LM00].

We conclude that with overwhelming probability $1 - d^{-\omega(1)}$,

$$\left|\langle x, A^{-1}x\rangle - \|x\|^2\right| \le |\langle x, Bx\rangle| + \tilde{O}(d/n) \le \tilde{O}\left(\frac{d+\sqrt{n}}{n}\right).$$

□

Lemma 5.6.2. *Let $a_1, \dots, a_n \in \mathbb{R}^d$ be independent random vectors from $\mathcal{N}(0, \frac{1}{n}\mathrm{Id})$ with $d \le n$ and let $A = \sum_{i=1}^{n} a_i a_i^\top$. Then for every index $i \in [n]$, with overwhelming probability $1 - d^{\omega(1)}$,*

$$\left|\langle a_j, A^{-1}a_j\rangle - \|a_j\|^2\right| \le \tilde{O}\left(\frac{d+\sqrt{n}}{n}\right) \cdot \|a_j\|^2.$$

Proof. Let $A_{-j} = \sum_{i \ne j} a_i a_i^\top$. By Sherman–Morrison,

$$A^{-1} = (A_{-j} + a_j a_j^\top)^{-1} = A_{-j}^{-1} - \frac{1}{1 + a_j^\top A_{-j}^{-1} a_j} A_{-j}^{-1} a_j a_j^\top A_{-j}^{-1}$$

Thus, $\langle a_j, A^{-1}a_j\rangle = \langle a_j, A_{-j}^{-1}a_j\rangle - \langle a_j, A_{-j}^{-1}a_j\rangle^2/(1 + \langle a_j, A_{-j}^{-1}a_j\rangle)$. Since $\|\frac{n}{n-1}A_{-j} - \mathrm{Id}\| = \tilde{O}(d/n)^{1/2}$ w.ov.p., we also have $\|A_{-j}^{-1}\| \le 2$ with overwhelming probability. Therefore, w.ov.p.,

$$\left|\langle a_j, A^{-1}a_j\rangle - \langle a_j, A_{-j}^{-1}a_j\rangle\right| \le \langle a_j, A_{-j}^{-1}a_j\rangle^2 \le 4\|a_j\|^4 \le \tilde{O}(d/n) \cdot \|a_j\|^2.$$

At the same time, by Lemma 5.6.1, w.ov.p.,

$$\left|\langle a_j, \tfrac{n}{n-1}A_{-j}^{-1}a_j\rangle - \|a_j\|^2\right| \le \tilde{O}\left(\frac{d+\sqrt{n}}{n}\right) \cdot \|a_j\|^2.$$

We conclude that, w.ov.p.,

$$\begin{aligned}\left|\langle a_j, A^{-1}a_j\rangle - \|a_j\|^2\right| &\le \left|\langle a_j, A^{-1}a_j\rangle - \langle a_j, A_{-j}^{-1}a_j\rangle\right| + \left|\langle a_j, A_{-j}^{-1}a_j\rangle - \tfrac{n-1}{n}\|a_j\|^2\right| + \tfrac{1}{n}\|a_j\|^2 \\ &\le \tilde{O}\left(\frac{d+\sqrt{n}}{n}\right).\end{aligned}$$

□

Lemma 5.6.3. *Let A be a block matrix where one of the diagonal blocks is the 1×1 identity; that is,*

$$A = \begin{pmatrix} \|v\|^2 & c^\top \\ c & B \end{pmatrix} = \begin{pmatrix} 1 & c^\top \\ c & B \end{pmatrix}.$$

for some matrix B and vector c. Let x be a vector which decomposes as $x = (x(1)\ x')$ where $x(1) = \langle x, e_1 \rangle$ for e_1 the first standard basis vector.

Then

$$\langle x, A^{-1}x \rangle = \langle x', \left(B^{-1} + \frac{B^{-1}cc^\top B^{-1}}{1 - c^\top B^{-1}c}\right) x' \rangle + 2x(1)\langle \left(B^{-1} + \frac{B^{-1}cc^\top B^{-1}}{1 - c^\top B^{-1}c}\right) c, x' \rangle + (1 - c^\top B^{-1}c)^{-1} x(1)^2 .$$

Proof. By the formula for block matrix inverses,

$$A^{-1} = \begin{pmatrix} (1 - c^\top B^{-1}c)^{-1} & c^T(B - cc^\top)^{-1} \\ (B - cc^\top)^{-1}c & (B - cc^\top)^{-1} \end{pmatrix}.$$

The result follows by Sherman-Morrison applied to $(B - cc^\top)^{-1}$ and the definition of x. □

Lemma 5.6.4. *Let $v \in \mathbb{R}^n$ be a unit vector and let $b_1, \ldots, b_n \in \mathbb{R}^{d-1}$ have iid entries from $\mathcal{N}(0, 1/n)$. Let $a_i \in \mathbb{R}^d$ be given by $a_i := (v(i)\ b_i)$. Let $A := \sum_i a_i a_i^T$. Let $c \in \mathbb{R}^{d-1}$ be given by $c := \sum_i v(i) b_i$. Then for every index $i \in [n]$, w.ov.p.,*

$$\left| \langle a_i, A^{-1}a_i \rangle - \|a_i\|^2 \right| \le \widetilde{O}\left(\frac{d + \sqrt{n}}{n}\right) \cdot \|a_i\|^2 .$$

Proof. Let $B := \sum_i b_i b_i^T$. By standard concentration, $\|B^{-1} - \mathrm{Id}\| \le \widetilde{O}(d/n)^{1/2}$ w.ov.p. [Ver12, Corollary 5.50]. At the same time, since v has unit norm, the entries of c are iid samples from $\mathcal{N}(0, 1/n)$, and hence $n\|c\|^2$ is χ^2-distributed with d degrees of freedom. Thus w.ov.p. $\|c\|^2 \le \frac{d}{n} + \widetilde{O}(dn)^{-1/2}$. Together these imply the following useful estimates, all of which hold w.ov.p.:

$$|c^\top B^{-1}c| \le \|c\|^2 \|B^{-1}\|_{op} \le \frac{d}{n} + \widetilde{O}\left(\frac{d}{n}\right)^{3/2}$$

$$\|B^{-1}cc^\top B^{-1}\|_{op} \le \|c\|^2 \|B^{-1}\|_{op}^2 \le \frac{d}{n} + \widetilde{O}\left(\frac{d}{n}\right)^{3/2}$$

$$\left\| \frac{B^{-1}cc^\top B^{-1}}{1 - c^\top B^{-1}c} \right\|_{op} \le \frac{d}{n} + \widetilde{O}\left(\frac{d}{n}\right)^{3/2},$$

where the first two use Cauchy-Schwarz and the last follows from the first two.

We turn now to the expansion of $\langle a_i, A^{-1}a_i \rangle$ offered by Lemma 5.6.3,

$$\langle a_i, A^{-1}a_i \rangle = \langle b_i, \left(B^{-1} + \frac{B^{-1}cc^\top B^{-1}}{1 - c^\top B^{-1}c}\right) b_i \rangle \tag{5.6.1}$$

$$+ 2v(i)\langle\left(B^{-1} + \frac{B^{-1}cc^\top B^{-1}}{1 - c^\top B^{-1}c}\right) c, b_i\rangle \tag{5.6.2}$$

$$+ (1 - c^\top B^{-1}c)^{-1}v(i)^2 \,. \tag{5.6.3}$$

Addressing (5.6.1) first, by the above estimates and Lemma 5.6.2 applied to $\langle b_i, B^{-1}b_i\rangle$,

$$\left|\langle b_i, \left(B^{-1} + \frac{B^{-1}cc^\top B^{-1}}{1 - c^\top B^{-1}c}\right) b_i\rangle - \|b_i\|^2\right| \leq \widetilde{O}\left(\frac{d+\sqrt{n}}{n}\right) \cdot \|b_i\|^2$$

w.ov.p.. For (5.6.2), we pull out the important factor of $\|c\|$ and separate $v(i)$ from b_i: w.ov.p.,

$$\begin{aligned}
\left|2v(i)\langle\left(B^{-1} + \frac{B^{-1}cc^\top B^{-1}}{1 - c^\top B^{-1}c}\right) c, b_i\rangle\right| &= \left|2\|c\|v(i)\langle\left(B^{-1} + \frac{B^{-1}cc^\top B^{-1}}{1 - c^\top B^{-1}c}\right) \frac{c}{\|c\|}, b_i\rangle\right| \\
&\leq \left|\|c\|^2 \left(v(i)^2 + \langle\left(B^{-1} + \frac{B^{-1}cc^\top B^{-1}}{1 - c^\top B^{-1}c}\right) \frac{c}{\|c\|}, b_i\rangle^2\right)\right| \\
&\leq \widetilde{O}\left(\frac{d}{n}\right)(v(i)^2 + \|b_i\|^2) \\
&= \widetilde{O}\left(\frac{d}{n}\right)\|a_i\|^2 \,,
\end{aligned}$$

where the last inequality follows from our estimates above and Cauchy-Schwarz.

Finally, for (5.6.3), since $(1 - c^\top B^{-1}c) \geq 1 - \widetilde{O}(d/n)$ w.ov.p., we have that

$$|(1 - c^\top B^{-1}c)^{-1}v(i)^2 - v(i)^2| \leq \widetilde{O}\left(\frac{d}{n}\right) v(i)^2 \,.$$

Putting it all together,

$$\begin{aligned}
\left|\langle a_i, A^{-1}a_i\rangle - \|a_i\|^2\right| &\leq \left|\langle b_i, \left(B^{-1} + \frac{B^{-1}cc^\top B^{-1}}{1 - c^\top B^{-1}c}\right) b_i\rangle - \|b_i\|^2\right| \\
&+ \left|2v(i)\langle\left(B^{-1} + \frac{B^{-1}cc^\top B^{-1}}{1 - c^\top B^{-1}c}\right) c, b_i\rangle\right| \\
&+ |(1 - c^\top B^{-1}c)^{-1}v(i)^2 - v(i)^2| \\
&\leq \widetilde{O}\left(\frac{d+\sqrt{n}}{n}\right) \cdot \|a_i\|^2 \,.
\end{aligned}$$

□

5.7 Concentration Bounds for Overcomplete Tensor Decomposition

We require some facts about the concentration of certain scalar-and matrix-valued random variables, which generally follow from standard concentration arguments. We present proofs here for completeness.

The first lemma captures standard facts about random Gaussians.

Fact 5.7.1. *Let $a_1, \dots, a_n \in \mathbb{R}^d$ be sampled $a_i \sim \mathcal{N}(0, \frac{1}{d}\,\mathrm{Id})$.*

1. *Inner products $|\langle a_i, a_j\rangle|$ are all $\approx 1/\sqrt{d}$:*

$$\mathbb{P}\left\{\langle a_i, a_j\rangle^2 \le \widetilde{O}\left(\frac{1}{d}\right) \;\middle|\; \forall i, j \in [n],\ i \neq j\right\} \ge 1 - n^{-\omega(1)}.$$

2. *Norms are all about $\|a_i\| \approx 1 \pm \widetilde{O}(1/\sqrt{d})$:*

$$\mathbb{P}\left\{1 - \widetilde{O}(1/\sqrt{d}) \le \|a_i\|_2^2 \le 1 + \widetilde{O}(1/\sqrt{d}) \;\middle|\; \forall i \in [n]\right\} \ge 1 - n^{-\omega(1)}\,.$$

3. *Fix a vector $v \in \mathbb{R}^d$. Suppose $g \in \mathbb{R}^d$ is a vector with entries identically distributed $g_i \sim \mathcal{N}(0, \sigma)$. Then $\langle g, v\rangle^2 \approx \sigma^2 \cdot \|v\|_2^2$:*

$$\mathbb{P}\left\{\left|\langle g, v\rangle^2 - \sigma^2 \cdot \|v\|_2^4\right| \le \widetilde{O}(\sigma^2 \cdot \|v\|_4^2)\right\} \ge 1 - n^{-\omega(1)}\,.$$

Proof of Fact 5.7.1. We start with Item 1. Consider the quantity $\langle a_i, a_j\rangle^2$. We calculate the expectation,

$$\mathbb{E}\left[\langle a_i, a_j\rangle^2\right] = \sum_{k,\ell \in [d]} \mathbb{E}\left[a_i(k)a_i(\ell)a_j(k)a_j(\ell)\right] = \sum_{k \in [d]} \mathbb{E}\left[a_i(k)^2\right] \cdot \mathbb{E}\left[a_j(k)^2\right] = d \cdot \frac{1}{d^2} = \frac{1}{d}.$$

Since this is a degree-4 square polynomial in the entries of a_i and a_j, we may apply Lemma A.2.2 to conclude that

$$\mathbb{P}\left(\langle a_i, a_j\rangle^2 \ge t \cdot \frac{1}{d}\right) \le \exp\left(-O(t^{1/2})\right).$$

Applying this fact with $t = \mathrm{polylog}(n)$ and taking a union bound over pairs $i, j \in [n]$ gives us the desired result.

Next is Item 2. Consider the quantity $\|a_i\|_2^2$. We will apply Lemma A.2.2 in order to obtain a tail bound for the value of the polynomial $(\|a_i\|_2^2 - 1)^2$. We have

$$\mathbb{E}\left[(\|a_i\|_2^2 - 1)^2\right] = O\left(\frac{1}{d}\right),$$

and now applying Lemma A.2.2 with the square root of this expectation, we have

$$\mathbb{P}\left(\left|\|a_i\|_2^2 - 1\right| \ge \widetilde{O}(\tfrac{1}{\sqrt{d}})\right) \le n^{-\log n}\,.$$

This gives both bounds for a single a_i. The result now follows from taking a union bound over all i.

Moving on to Item 3, we view the expression $f(g) := (\langle g, v\rangle^2 - \sigma^2\|v\|^2)^2$ as a polynomial in the Gaussian entries of g. The degree of $f(g)$ is 4, and $\mathbb{E}[|f(g)|] = 3\sigma^4 \cdot \|v\|_4^4$, and so we may apply Lemma A.2.2 to conclude that

$$\mathbb{P}\left(|f(g)| \geq t \cdot 3\sigma^4 \cdot \|v\|_4^4\right) \leq \exp(-c_4 t^{1/2}),$$

and taking $t = \mathrm{polylog}(n)$ the conclusion follows. □

We also use the fact that the covariance matrix of a sum of sufficiently many Gaussian outer products concentrates about its expectation.

Fact 5.7.2. *Let $a_1, \ldots, a_n \in \mathbb{R}^d$ be vectors with iid Gaussian entries such that $\mathbb{E}\left[\|a_i\|_2^2\right] = 1$, and $n = \Omega(d)$. Let $\mathcal{E}$ be the event that the sum $\sum_{i\in[n]} a_i a_i^\top$ is close to $\frac{n}{d} \cdot \mathrm{Id}$, that is*

$$\mathbb{P}\left\{\widetilde{\Omega}(n/d) \cdot \mathrm{Id} \leq \sum_{i\in[n]} a_i a_i^\top \leq \widetilde{O}(n/d) \cdot \mathrm{Id}\right\} \geq 1 - n^{-\omega(1)}.$$

Proof of Fact 5.7.2. We apply a truncated matrix Bernstein inequality. For convenience, $A := \sum_{i\in[n]} a_i a_i^\top$ and let $A_i := a_i a_i^\top$ be a single summand. To begin, we calculate the first and second moments of the summands,

$$\mathbb{E}\left[A_i\right] = \frac{1}{d} \cdot \mathrm{Id}$$
$$\mathbb{E}\left[A_i A_i^\top\right] = O\left(\frac{1}{d}\right) \cdot \mathrm{Id}.$$

So we have $\mathbb{E}\left[A\right] = \frac{n}{d} \cdot \mathrm{Id}$ and $\sigma^2(A) = O\left(\frac{n}{d}\right)$.

We now show that each summand is well-approximated by a truncated variable. To calculate the expected norm $\|A_i\|_{op}$, we observe that A_i is rank-1 and thus $\mathbb{E}\left[\|A_i\|_{op}\right] = \mathbb{E}\left[\|a_i\|_2^2\right] = 1$. Applying Lemma A.3.4, we have

$$\mathbb{P}\left(\|A_i\|_{op} \geq \widetilde{O}(1)\right) \leq n^{-\log n},$$

and also

$$\mathbb{E}\left[\|A_i\|_{op} \cdot \mathbb{I}\{\|A_i\|_{op} \geq \widetilde{O}(1)\}\right] \leq n^{-\log n}.$$

Thus, applying the truncated matrix Bernstein inequality from Proposition A.3.3 with $\sigma^2 = O(\frac{n}{d})$, $\beta = \widetilde{O}(1)$, $p = n^{-\log n}$, $q = n^{-\log n}$, and $t = \widetilde{O}\left(\frac{n^{1/2}}{d^{1/2}}\right)$, we have that with overwhelming probability,

$$\left\|A - \frac{n}{d} \cdot \mathrm{Id}\right\|_{op} \leq \widetilde{O}\left(\frac{n^{1/2}}{d^{1/2}}\right).$$

□

We now show that among the terms of the polynomial $\langle g, Ta_i^{\otimes 2}\rangle$, those that depend on a_j with $j \neq i$ have small magnitude. This polynomial appears in the proof that M_{diag} has a noticeable spectral gap.

Lemma (Restatement of Lemma 5.4.6)**.** *Let $a_1, \dots, a_n$ be independently sampled vectors from $\mathcal{N}(0, \frac{1}{d}\operatorname{Id}_d)$, and let g be sampled from $\mathcal{N}(0, \operatorname{Id}_d)$. Let $T = \sum_i a_i(a_i \otimes a_i)^\top$. Then with overwhelming probability, for every $j \in [n]$,*

$$\left|\langle g, T(a_j \otimes a_j)\rangle - \langle g, a_j\rangle \|a_j\|^4\right| \leq \widetilde{O}\left(\frac{\sqrt{n}}{d}\right).$$

Proof. Fixing a_i and g, the terms in the summation are independent, and we may apply a Bernstein inequality. A straightforward calculation shows that the expectation of the sum is 0 and the variance is $\widetilde{O}(\frac{n}{d^2}) \cdot \|g\|^2\|a_i\|^4$. Additionally, each summand is a polynomial in Gaussian variables, the square of which has expectation $\widetilde{O}(\frac{1}{d^2} \cdot \|g\|^2\|a_i\|^4)$. Thus Lemma A.2.2 allows us to truncate each summand appropriately so as to employ Proposition A.3.3. An appropriate choice of logarithmic factors and the concentration of $\|g\|^2$ and $\|a_i\|^2$ due to Fact 5.7.1 gives the result for each $i \in [n]$. A union bound over each choice of i gives the final result. □

Finally, we prove that a matrix which appears in the expression for M_{same} has bounded norm w.ov.p.

Lemma 5.7.3. *Let $a_1, \dots, a_n$ be independent from $\mathcal{N}(0, \frac{1}{d}\operatorname{Id}_d)$. Let $g \sim \mathcal{N}(0, \operatorname{Id}_d)$. Fix $j \in [n]$. Then w.ov.p.*

$$\left\|\sum_{\substack{i\in[n]\\ i\neq j}} \langle g, a_i\rangle \|a_i\|^2 \langle a_i, a_j\rangle \cdot a_i a_i^\top\right\| \leq \widetilde{O}(n/d^2)^{1/2}.$$

Proof. The proof proceeds by truncated matrix Bernstein, since the summands are independent for fixed g, a_j. For this we need to compute the variance:

$$\sigma^2 = \left\|\sum_{\substack{i\in[n]\\ i\neq j}} \mathbb{E}\langle g, a_i\rangle^2 \|a_i\|^6 \langle a_i, a_j\rangle^2 \cdot a_i a_i^\top\right\| \leq O(1/d) \cdot \left\|\sum_{\substack{i\in[n]\\ i\neq j}} \mathbb{E}\, a_i a_i^\top\right\| \leq O(1/d) \cdot n/d \leq O(n/d^2).$$

The norm of each term in the sum is bounded by a constant-degree polynomial of Gaussians. Straightforward calculations show that in expectation each term is $O(\frac{1}{d}\langle g, a_i\rangle)$ in norm; w.ov.p. this is $O(\sigma)$. So Lemma A.2.2 applies to establish the hypothesis of truncated Bernstein Proposition A.3.3. In turn, Proposition A.3.3 yields that w.ov.p.

$$\left\|\sum_{\substack{i\in[n]\\ i\neq j}} \langle g, a_i\rangle \|a_i\|^2 \langle a_i, a_j\rangle \cdot a_i a_i^\top\right\| \leq \widetilde{O}(\sigma) = \widetilde{O}(n/d^2)^{1/2}.$$

□

Proof of Fact 5.7.4

Here we prove the following fact.

Fact 5.7.4. *Let* $\Sigma = \mathbb{E}_{x\sim\mathcal{N}(0,\mathrm{Id}_d)}(xx^\top)^{\otimes 2}$ *and let* $\tilde{\Sigma} = \mathbb{E}_{x\sim\mathcal{N}(0,\mathrm{Id}_d)}(xx^\top)^{\otimes 2}/\|x\|^4$. *Let* $\Phi = \sum_i e_i^{\otimes 2} \in \mathbb{R}^{d^2}$ *and let* Π_{sym} *be the projector to the symmetric subspace of* $\mathbb{R}^{d^2}$ *(the span of vectors of the form* $x^{\otimes 2}$ *for* $x \in \mathbb{R}^d$*). Then*

$$\Sigma = 2\,\Pi_{\mathrm{sym}} + \Phi\Phi^\top, \qquad \tilde{\Sigma} = \tfrac{2}{d^2+2d}\Pi_{\mathrm{sym}} + \tfrac{1}{d^2+2d}\Phi\Phi^\top,$$
$$\Sigma^+ = \tfrac{1}{2}\Pi_{\mathrm{sym}} - \tfrac{1}{2(d+2)}\Phi\Phi^\top, \qquad \tilde{\Sigma}^+ = \tfrac{d^2+2d}{2}\Pi_{\mathrm{sym}} - \tfrac{d}{2}\Phi\Phi^\top.$$

In particular,

$$R = \sqrt{2}\,(\Sigma^+)^{1/2} = \Pi_{\mathrm{sym}} - \tfrac{1}{d}\left(1 - \sqrt{\tfrac{2}{d+2}}\right)\Phi\Phi^\top \quad \textit{has} \quad \|R\| = 1$$

and for any $v \in \mathbb{R}^d$,

$$\|R(v\otimes v)\|_2^2 = \left(1 - \tfrac{1}{d+2}\right)\cdot\|v\|^4.$$

We will derive Fact 5.7.4 as a corollary of a more general claim about rotationally symmetric distributions.

Lemma 5.7.5. *Let* $\mathcal{D}$ *be a distribution over* $\mathbb{R}^d$ *which is rotationally symmetric; that is, for any rotation* R, $x \sim \mathcal{D}$ *is distributed identically to* Rx. *Let* $\Sigma = \mathbb{E}_{x\sim\mathcal{D}}(xx^\top)^{\otimes 2}$, *let* $\Phi = \sum_i e_i^{\otimes 2} \in \mathbb{R}^{d^2}$ *and let* Π_{sym} *be the projector to the symmetric subspace of* $\mathbb{R}^{d^2}$ *(the span of vectors of the form* $x^{\otimes 2}$ *for* $x \in \mathbb{R}^d$*). Then there is a constant* r *so that*

$$\Sigma = 2r\,\Pi_{\mathrm{sym}} + r\,\Phi\Phi^\top.$$

Furthermore, r *is given by*

$$r = \mathbb{E}\langle x,a\rangle^2\langle x,b\rangle^2 = \tfrac{1}{3}\,\mathbb{E}\langle x,a\rangle^4$$

where a, b *are orthogonal unit vectors.*

Proof. First, Σ is symmetric and operates nontrivially only on the symmetric subspace (in other words $\ker\Pi_{\mathrm{sym}} \subseteq \ker\Sigma$). This follows from Σ being an expectation over symmetric matrices whose kernels always contain the complement of the symmetric subspace.

Let $\hat{a}, \hat{b}, \hat{c}, \hat{d} \in \mathbb{R}^d$ be any four orthogonal unit vectors. Let R be any rotation of $\mathbb{R}^d$ that takes $\hat{a}$ to $-\hat{a}$, but fixes $\hat{b}$, $\hat{c}$, and $\hat{d}$ (this rotation exists for $d \geq 5$, but a different argument holds for $d \leq 4$). By rotational symmetry about R, all of these quantities are 0:

$$\mathbb{E}\langle\hat{a},x\rangle\langle\hat{b},x\rangle\langle\hat{c},x\rangle\langle\hat{d},x\rangle = 0,$$

$$\mathbb{E}\langle\hat{a},x\rangle\langle\hat{b},x\rangle\langle\hat{c},x\rangle^2 = 0, \qquad \mathbb{E}\langle\hat{a},x\rangle\langle\hat{b},x\rangle^3 = 0.$$

Furthermore, let Q be a rotation of $\mathbb{R}^d$ that takes $\hat{a}$ to $(\hat{a}+\hat{b})/\sqrt{2}$. Then by rotational symmetry about Q,

$$\mathbb{E}\langle \hat{a}, x\rangle^4 = \mathbb{E}\langle \hat{a}, Qx\rangle^4 = \mathbb{E}\,\tfrac{1}{4}\langle \hat{a}+\hat{b}, x\rangle^4 = \mathbb{E}\,\tfrac{1}{4}[\langle \hat{a}, x\rangle^4 + \langle \hat{b}, x\rangle^4 + 6\langle \hat{a}, x\rangle^2\langle \hat{b}, x\rangle^2]$$

Thus, since $\mathbb{E}\langle \hat{a}, x\rangle^4 = \mathbb{E}\langle \hat{b}, x\rangle^4$ by rotational symmetry, we have

$$\mathbb{E}\langle \hat{a}, x\rangle^4 = 3\,\mathbb{E}\langle \hat{a}, x\rangle^2\langle \hat{b}, x\rangle^2.$$

So let $r := \mathbb{E}\langle \hat{a}, x\rangle^2\langle \hat{b}, x\rangle^2 = \frac{1}{3}\,\mathbb{E}\langle \hat{a}, x\rangle^4$. By rotational symmetry, r is constant over choice of orthogonal unit vectors $\hat{a}$ and $\hat{b}$.

Since Σ operates only on the symmetric subspace, let $u \in \mathbb{R}^{d^2}$ be any unit vector in the symmetric subspace. Such a u unfolds to a symmetric matrix in $\mathbb{R}^{d\times d}$, so that it has an eigendecomposition $u = \sum_{i=1}^d \lambda_i u_i \otimes u_i$. Evaluating $\langle u, \Sigma u\rangle$,

$$\begin{aligned}
\langle u, \Sigma u\rangle &= \sum_{i,j=1}^d \mathbb{E}\,\lambda_i\lambda_j\langle x, u_i\rangle^2\langle x, u_j\rangle^2 \quad \text{other terms are 0 by above} \\
&= 3r\sum_{i=1}^d \lambda_i^2 + r\sum_{i\neq j}\lambda_i\lambda_j \\
&= 2r\sum_{i=1}^d \lambda_i^2 + r\left(\sum_{i=1}^d \lambda_i\right)^2 \\
&= 2r\,\|u\|^2 + r\left(\sum_{i=1}^d \lambda_i\right)^2 \quad \text{Frobenius norm is sum of squared eigenvalues} \\
&= 2r\,\|u\|^2 + r\left(\sum_i u_{i,i}\right)^2 \quad \text{trace is sum of eigenvalues} \\
&= 2r\,\langle u, \Pi_{\mathrm{sym}} u\rangle + r\,\langle u, \Phi\Phi^\top u\rangle\,,
\end{aligned}$$

so therefore $\Sigma = 2r\,\Pi_{\mathrm{sym}} + r\,\Phi\Phi^\top$. □

Proof of Fact 5.7.4. When $x \sim \mathcal{N}(0, \mathrm{Id}_d)$, the expectation $\mathbb{E}\langle x, a\rangle^2\langle x, b\rangle^2 = 1$ is just a product of independent standard Gaussian second moments. Therefore by Lemma 5.7.5, $\Sigma = 2\,\Pi_{\mathrm{sym}} + \Phi\Phi^\top$.

To find $\tilde{\Sigma}$ where x is uniformly distributed on the unit sphere, we compute

$$1 = \mathbb{E}\,\|x\|^4 = \sum_{i,j}\mathbb{E}\,x_i^2x_j^2 = d\;\mathbb{E}\,x_1^4 + (d^2-d)\;\mathbb{E}\,x_1^2x_2^2$$

and use the fact that $\mathbb{E}\,x_1^4 = 3\;\mathbb{E}\,x_1^2$ (by Lemma 5.7.5) to find that $\mathbb{E}\,x_1^2x_2^2 = \frac{1}{d^2+2d}$, and therefore by Lemma 5.7.5, $\tilde{\Sigma} = \frac{2}{d^2+2d}\Pi_{\mathrm{sym}} + \frac{1}{d^2+2d}\Phi\Phi^\top$.

To verify the pseudoinverses, it is enough to check that $MM^+ = \Pi_{\text{sym}}$ for each matrix M and its claimed pseudoinverse M^+.

To show that

$$\|R(v \otimes v)\|_2^2 = \left(1 - \tfrac{1}{d+2}\right) \cdot \|v\|^4 ,$$

for any $v \in \mathbb{R}^d$, we write $\|R(v\otimes v)\|_2^2 = (v\otimes v)^\top R^2 (v\otimes v)$ and use the substitution $R^2 = 2\Sigma^+$, along with the facts that $\Pi_{\text{sym}}(v \otimes v) = v \otimes v$ and $\langle \Phi, v \otimes v\rangle = \|v\|^2$.

□

Now we can prove some concentration claims we deferred:

Lemma (Restatement of Lemma 5.4.11)**.** *Let $a_1, \dots, a_n \sim \mathcal{N}(0, \frac{1}{d}\,\mathrm{Id}_d)$. Let Σ, R be as in Fact 5.4.8. Let $u_i = a_i \otimes a_i$. With overwhelming probability, every $j \in [n]$ satisfies $\sum_{i\neq j}\langle u_j, R^2 u_i\rangle^2 = \widetilde{O}(n/d^2)$ and $|1 - \|Ru_j\|^2| \le \widetilde{O}(1/\sqrt{d})$.*

Proof of Lemma 5.4.11. We prove the first item:

$$\begin{aligned}
\sum_{i\neq j}\langle u_j, R^2 u_i\rangle^2 &= \sum_{i\neq j}\langle u_j, 2\Sigma^+ u_i\rangle^2 \\
&= \sum_{i\neq j}\langle u_j, (\Pi_{\text{sym}} - \tfrac{1}{d+2}\Phi\Phi^\top)u_i\rangle^2 \quad \text{by Fact 5.7.4} \\
&= \sum_{i\neq j}(\langle a_j, a_i\rangle^2 - \tfrac{1}{d+2}\|u_j\|^2\|u_i\|^2)^2 \\
&= \sum_{i\neq j}\widetilde{O}(1/d)^2 \quad \text{w.ov.p. by Fact 5.7.1} \\
&= \widetilde{O}(n/d^2)\,.
\end{aligned}$$

And one direction of the second item, using Fact 5.7.4 and Fact 5.7.1 (the other direction is similar):

$$\|Ru_j\|^2 = \langle u_j, R^2 u_j\rangle = \langle u_j, (\Pi_{\text{sym}} + \tfrac{1}{d+2}\Phi\Phi^\top)u_j\rangle = (1 - \Theta(1/d))\|a_j\|^4 = 1 - \widetilde{O}(1/\sqrt{d})$$

where the last equality holds w.ov.p.. □

Proof of Lemma 5.4.9

To prove Lemma 5.4.9 we will begin by reducing to the case $S = [n]$ via the following.

Lemma 5.7.6. *Let $v_1, \dots, v_n \in \mathbb{R}^d$. Let A_S have columns $\{v_i\}_{i\in S}$. Let Π_S be the projector to $\mathrm{Span}\{v_i\}_{i\in S}$. Suppose there is $c \ge 0$ so that $\|A_{[n]}^\top A_{[n]} - \mathrm{Id}_n\| \le c$. Then for every $S \subseteq [n]$, $\|A_S A_S^\top - \Pi_S\| \le c$*

Proof. If the hypothesized bound $\|A_{[n]}^\top A_{[n]} - \mathrm{Id}_n\| \le c$ holds then for every $S \subseteq [n]$ we get $\|A_S^\top A_S - \mathrm{Id}_{|S|}\| \le c$ since $A_S^\top A_S$ is a principal submatrix of $A_{[n]}^\top A_{[n]}$. If $\|A_S^\top A_S - \mathrm{Id}_{|S|}\| \le c$, then because $A_S A_S^\top$ has the same nonzero eigenvalues as $A_S^\top A_S$, we must have also $\|A_S A_S^\top - \Pi_S\| \le c$. □

It will be convenient to reduce concentration for matrices involving $a_i \otimes a_i$ to analogous matrices where the vectors $a_i \otimes a_i$ are replaced by isotropic vectors of constant norm. The following lemma shows how to do this.

Lemma 5.7.7. *Let* $a \sim \mathcal{N}(0, \frac{1}{d}\operatorname{Id}_d)$. *Let* $\tilde{\Sigma} := \mathbb{E}_{x\sim\mathcal{N}(0,\operatorname{Id}_d)}(xx^\top)^{\otimes 2}/\|x\|^4$. *Then* $u := (\tilde{\Sigma}^+)^{1/2}a \otimes a/\|a\|^2$ *is an isotropic random vector in the symmetric subspace* $\operatorname{Span}\{y \otimes y \mid y \in \mathbb{R}^d\}$ *with* $\|u\| = \sqrt{\dim \operatorname{Span}\{y \otimes y \mid y \in \mathbb{R}^d\}}$.

Proof. The vector u is isotropic by definition so we prove the norm claim. Let $\tilde{\Phi} = \Phi/\|\Phi\|$. By Fact 5.7.4,

$$\tilde{\Sigma}^+ = \tfrac{d^2+2d}{2}\Pi_{\text{sym}} - \tfrac{d}{2}\Phi\Phi^\top$$

Thus,

$$\|u\|^2 = \langle \tfrac{a\otimes a}{\|a\|^2}, \tilde{\Sigma}^+ \tfrac{a\otimes a}{\|a\|^2}\rangle = \tfrac{d^2+2d}{2} - \tfrac{d}{2} = \tfrac{d^2+d}{2} = \dim \operatorname{Span}\{y \otimes y \mid y \in \mathbb{R}^d\}\,. \qquad \square$$

The last ingredient to finish the spectral bound is a bound on the incoherence of independent samples from $(\tilde{\Sigma}^+)^{1/2}$.

Lemma 5.7.8. *Let* $\tilde{\Sigma} = \mathbb{E}_{a\sim\mathcal{N}(0,\operatorname{Id}_d)}(aa^\top \otimes aa^\top)/\|a\|^4$. *Let* $a_1, \dots, a_n \sim \mathcal{N}(0, \operatorname{Id}_d)$ *be independent, and let* $u_i = (\tilde{\Sigma}^+)^{1/2}(a_i \otimes a_i)/\|a_i\|^2$. *Let* $d' = \dim \operatorname{Span}\{y \otimes y \mid y \in \mathbb{R}^d\} = \frac{1}{2}(d^2+d)$. *Then*

$$\tfrac{1}{d'}\,\mathbb{E}\max_i \sum_{j\neq i}\langle u_i, u_j\rangle^2 \le \widetilde{O}(n)\,.$$

Proof. Expanding $\langle u_i, u_j\rangle^2$ and using $\tilde{\Sigma}^+ = \frac{d^2+2d}{2}\Pi_{\text{sym}} - \frac{d}{2}\Phi\Phi^\top$, we get

$$\langle u_i, u_j\rangle^2 = \left(\tfrac{d^2+2d}{2}\langle \tfrac{a_i\otimes a_i}{\|a_i\|^2}, \tfrac{a_j\otimes a_j}{\|a_j\|^2}\rangle - \tfrac{d}{2}\right)^2 = \left(\tfrac{d^2+2d}{2}\cdot \tfrac{\langle a_i, a_j\rangle^2}{\|a_i\|^2\|a_j\|^2} - \tfrac{d}{2}\right)^2$$

From elementary concentration, $\mathbb{E}\max_{i\neq j}\langle a_i, a_j\rangle^2/\|a_i\|^2\|a_j\|^2 \le \widetilde{O}(1/d)$, so the lemma follows by elementary manipulations. $\square$

We need the following bound on the deviation from expectation of a tall matrix with independent columns.

Theorem 5.7.9 (Theorem 5.62 in [Ver12])**.** *Let* A *be an* $N \times n$ *matrix* ($N \ge n$) *whose columns* A_j *are independent isotropic random vectors in* $\mathbb{R}^N$ *with* $\|A_j\|_2 = \sqrt{N}$ *almost surely. Consider the incoherence parameter*

$$m \stackrel{\text{def}}{=} \frac{1}{N}\,\mathbb{E}\max_{i\in[n]}\sum_{j\neq i}\langle A_i, A_j\rangle^2\,.$$

Then $\mathbb{E}\,\|\frac{1}{N}A^TA - \operatorname{Id}\| \le C_0\sqrt{\frac{m\log n}{N}}$.

We are now prepared to handle the case of $S = [n]$ via spectral concentration for matrices with independent columns, Theorem 5.7.9.

Lemma (Restatement of Lemma 5.4.9)**.** *Let $a_1, \dots, a_n \sim \mathcal{N}(0, \frac{1}{d}\mathrm{Id}_d)$ be independent random vectors with $d \le n$. Let $R := \sqrt{2} \cdot ((\mathbb{E}(aa^\top)^{\otimes 2})^+)^{1/2}$ for $a \sim \mathcal{N}(0, \mathrm{Id}_d)$. For $S \subseteq [n]$, let $P_S = \sum_{i \in S} (a_i a_i^\top)^{\otimes 2}$ and let Π_S be the projector into the subspace spanned by $\{Ra_i^{\otimes 2} \mid i \in S\}$. Then, with probability $1 - o(1)$ over the choice of $a_1, \dots, a_n$,*

$$\forall S \subseteq [n]. \quad \left(1 - \widetilde{O}(n/d^{3/2})\right) \cdot \Pi_S \preceq RP_S R \preceq \left(1 + \widetilde{O}(n/d^{3/2})\right) \cdot \Pi_S \,.$$

Proof of Lemma 5.4.9. By Lemma 5.7.6 it is enough to prove the lemma in the case of $S = [n]$. For this we will use Theorem 5.7.9. Let A be the matrix whose columns are given by $a_i \otimes a_i$, so that $P_{[n]} = P = AA^\top$. Because $RAA^\top R$ and $A^\top RRA$ have the same nonzero eigenvalues, it will be enough to show that $\|A^\top R^2 A - \mathrm{Id}\| \le \widetilde{O}(\sqrt{n}/d) + \widetilde{O}(n/d^{3/2})$ with probability $1 - o(1)$. (Since $n \le d$ we have $\sqrt{n}/d = \widetilde{O}(n/d^{3/2})$ so this gives the theorem.)

The columns of RA are independent, given by $R(a_i \otimes a_i)$. However, they do not quite satisfy the normalization conditions needed for Theorem 5.7.9. Let D be the diagonal matrix whose i-th diagonal entry is $\|a_i\|^2$. Let $\tilde{\Sigma} = \mathbb{E}_{x \sim \mathcal{N}(0, \mathrm{Id})} (xx^\top)^{\otimes 2} / \|x\|^4$. Then by Lemma 5.7.7 the matrix $(\tilde{\Sigma}^+)^{1/2} D^{-1} A$ has independent columns from an isotropic distribution with a fixed norm d'. Together with Lemma 5.7.8 this is enough to apply Theorem 5.7.9 to conclude that $\mathbb{E}\,\|\frac{1}{(d')^2} A^\top D^{-1} \tilde{\Sigma}^+ D^{-1} A - \mathrm{Id}\| \le \widetilde{O}(\sqrt{n}/d)$. By Markov's inequality, $\|\frac{1}{(d')^2} A^\top D^{-1} \tilde{\Sigma}^+ D^{-1} A - \mathrm{Id}\| \le \widetilde{O}(\sqrt{n}/d)$ with probability $1 - o(1)$.

We will show next that $\|A^\top R^2 A - \frac{1}{(d')^2} A^\top D^{-1} \tilde{\Sigma}^+ D^{-1} A\| \le \widetilde{O}(n/d^{3/2})$ with probability $1 - o(1)$; the lemma then follows by triangle inequality. The expression inside the norm expands as

$$A^\top (R^2 - \tfrac{1}{(d')^2} D^{-1} \tilde{\Sigma}^+ D^{-1}) A \,.$$

and so

$$\|A^\top R^2 A - \tfrac{1}{(d')^2} A^\top D^{-1} \tilde{\Sigma}^+ D^{-1} A\| \le \|A\|^2 \|R^2 - \tfrac{1}{(d')^2} D^{-1} \tilde{\Sigma}^+ D^{-1}\|$$

By Fact 5.7.1, with overwhelming probability $\|D - \mathrm{Id}\| \le \widetilde{O}(1/\sqrt{d})$. So $\|(1/d')^2 D^{-1} \tilde{\Sigma}^+ D^{-1} - (1/d')^2 \tilde{\Sigma}^+\| \le \widetilde{O}(1/\sqrt{d})$ w.ov.p.. We recall from Fact 5.7.4, given that $R = \sqrt{2} \cdot (\Sigma^+)^{1/2}$, that

$$R^2 = \Pi_{\mathrm{sym}} - \tfrac{1}{d+2} \Phi\Phi^\top \quad \text{and} \quad \tfrac{1}{(d')^2} \tilde{\Sigma}^+ = \tfrac{d+2}{d+1} \Pi_{\mathrm{sym}} - \tfrac{1}{d+1} \Phi\Phi^\top \,.$$

This implies that $\|R^2 - (1/d')^2 \tilde{\Sigma}^+\| \le O(1/d)$. Finally, by an easy application of Proposition A.3.3, $\|A\|^2 = \|\sum_i (a_i a_i^\top)^{\otimes 2}\| \le \widetilde{O}(n/d)$ w.ov.p.. All together, $\|A^\top R^2 A - \frac{1}{(d')^2} A^\top D^{-1} \tilde{\Sigma}^+ D^{-1} A\| \le \widetilde{O}(n/d^{3/2})$. □

5.8 Concentration Bounds for Tensor Principal Component Analysis

For convenience, we restate Lemma 5.5.5 here.

Lemma 5.8.1 (Restatement of Lemma 5.5.5)**.** *For any v, with high probability over $\mathbf{A}$, the following occur:*

$$\left\|\sum_i \operatorname{Tr}(A_i)\cdot A_i\right\| \le O(n^{3/2}\log^2 n)$$
$$\left\|\sum_i v(i)\cdot A_i\right\| \le O(\sqrt{n}\log n)$$
$$\left\|\sum_i \operatorname{Tr}(A_i)v(i)\cdot vv^T\right\| \le O(\sqrt{n}\log n)\,.$$

Proof of Lemma 5.5.5. We begin with the term $\sum_i \operatorname{Tr}(A_i)\cdot A_i$. It is a sum of iid matrices $\operatorname{Tr}(A_i)\cdot A_i$. A routine computation gives $\mathbb{E}\operatorname{Tr}(A_i)\cdot A_i = \mathrm{Id}$. We will use the truncated matrix Bernstein's inequality (Proposition A.3.3) to bound $\|\sum_i \operatorname{Tr}(A_i)A_i\|$.

For notational convenience, let A be distributed like a generic A_i. By a union bound, we have both of the following:

$$\mathbb{P}\left(\|\operatorname{Tr}(A)\cdot A\| \ge tn\right) \le \mathbb{P}\left(|\operatorname{Tr}(A)| \ge \sqrt{tn}\right) + \mathbb{P}\left(\|A\| \ge \sqrt{tn}\right)$$
$$\mathbb{P}\left(\|\operatorname{Tr}(A)\cdot A - \mathrm{Id}\,\| \ge (t+1)n\right) \le \mathbb{P}\left(|\operatorname{Tr}(A)| \ge \sqrt{tn}\right) + \mathbb{P}\left(\|A\| \ge \sqrt{tn}\right).$$

Since $\operatorname{Tr}(A)$ the sum of iid Gaussians, $\mathbb{P}(|\operatorname{Tr}(A)| \ge \sqrt{tn}) \le e^{-c_1 t}$ for some constant c_1. Similarly, since the maximum eigenvalue of a matrix with iid entries has a subgaussian tail, $\mathbb{P}(\|A\| \ge \sqrt{tn}) \le e^{-c_2 t}$ for some c_2. All together, for some c_3, we get $\mathbb{P}(\|\operatorname{Tr}(A)\cdot A\| \ge tn) \le e^{-c_3 t}$ and $\mathbb{P}(\|\operatorname{Tr}(A)\cdot A - \mathrm{Id}\,\| \ge (t+1)n) \le e^{-c_3 t}$.

For a positive parameter β, let $\mathbb{I}_\beta$ be the indicator variable for the event $\|\operatorname{Tr}(A)\cdot A\| \le \beta$. Then

$$\begin{aligned}
\mathbb{E}\,\|\operatorname{Tr}(A)\cdot A\| - \mathbb{E}\,\|\operatorname{Tr}(A)\cdot A\|\,\mathbb{I}_\beta &= \int_0^\infty \left[\mathbb{P}(\|\operatorname{Tr} A\cdot A\| > s) - \mathbb{P}(\|\operatorname{Tr} A\cdot A\|\,\mathbb{I}_\beta > s)\right] \mathrm{d}s \\
&= \beta\,\mathbb{P}(\|\operatorname{Tr} A\cdot A\| > \beta) + \int_\beta^\infty \mathbb{P}(\|\operatorname{Tr} A\cdot A\| > s)\,\mathrm{d}s \\
&\le \beta e^{-c_3\beta/n} + \int_\beta^\infty \mathbb{P}(\|\operatorname{Tr} A\cdot A\| > s)\,\mathrm{d}s \\
&= \beta e^{-c_3\beta/n} + \int_{\beta/n}^\infty \mathbb{P}(\|\operatorname{Tr} A\cdot A\| \ge tn)\,n\,\mathrm{d}t \\
&\le \beta e^{-c_3\beta/n} + \int_{\beta/n}^\infty ne^{-c_3 t}\,\mathrm{d}t \\
&= \beta e^{-c_3\beta/n} + \tfrac{n}{c_3}e^{-c_3\beta/n}\,.
\end{aligned}$$

Thus, for some $\beta = O(n\log n)$ we may take the parameters p, q of Proposition A.3.3 to be $O(n^{-150})$. The only thing that remains is to bound the parameter σ^2. Since $(\mathbb{E}\operatorname{Tr}(A)\cdot A)^2 =$

Id, it is enough just to bound $\|\mathbb{E}\operatorname{Tr}(A)^2 AA^T\|$. We use again a union bound:

$$\mathbb{P}(\|\operatorname{Tr}(A)^2 AA^T\| > tn^2) \leq \mathbb{P}(|\operatorname{Tr}(A)| > t^{1/4}\sqrt{n}) + \mathbb{P}(\|A\| > t^{1/4}\sqrt{n})\,.$$

By a similar argument as before, using the Gaussian tails of $\operatorname{Tr} A$ and $\|A\|$, we get $\mathbb{P}(\|\operatorname{Tr}(A)^2 AA^T\| > tn^2) \leq e^{-c_4\sqrt{t}}$. Then starting out with the triangle inequality,

$$\begin{aligned}
\sigma^2 &= \|n \cdot \mathbb{E}\operatorname{Tr}(A)^2 AA^T\| \\
&\leq n \cdot \mathbb{E}\,\|\operatorname{Tr}(A)^2 AA^T\| \\
&= n \cdot \int_0^\infty \mathbb{P}(\operatorname{Tr}(A)^2 AA^T > s)\,\mathrm{d}s \\
&= n \cdot \int_0^\infty \mathbb{P}(\operatorname{Tr}(A)^2 AA^T > tn^2)\,n^2\,\mathrm{d}t \\
&\leq n \cdot \int_0^\infty e^{-c_4\sqrt{t}}\,n^2\,\mathrm{d}t \\
&= n \cdot \left[-\frac{2n^2(c_4\sqrt{t}+1)}{c_4^2}e^{-c_4\sqrt{t}}\right]_{t=0}^{t=\infty} \\
&\leq O(n^3)\,.
\end{aligned}$$

This gives that with high probability,

$$\left\|\sum_i \operatorname{Tr}(A_i)\cdot A_i\right\| \leq O(n^{3/2}\log^2 n)\,.$$

The other matrices are easier. First of all, we note that the matrix $\sum_i v(i)\cdot A_i$ has independent standard Gaussian entries, so it is standard that with high probability $\|\sum_i v(i)\cdot A_i\| \leq O(\sqrt{n}\log n)$. Second, we have

$$\sum_i v(i)\operatorname{Tr}(A_i)vv^T = vv^T\sum_i v(i)\operatorname{Tr}(A_i).$$

The random variable $\operatorname{Tr}(A_i)$ is a centered Gaussian with variance n, and since v is a unit vector, $\sum_i v(i)\operatorname{Tr}(A_i)$ is also a centered Gaussian with variance n. So with high probability we get

$$\left\|vv^T\sum_i v(i)\operatorname{Tr}(A_i)\right\| = \left|\sum_i v(i)\operatorname{Tr}(A_i)\right| \leq O(\sqrt{n}\log n)$$

by standard estimates. This completes the proof. □

Appendix A

Additional Technical Underpinnings

A.1 Linear Algebra

Here we provide some linear algebraic lemmas which we will need to make use of in proving our results.

This first lemma is closely related to the SoS Cauchy-Schwarz from [BKS14], and the proof is essentially the same.

Lemma A.1.1 (PSD Cauchy-Schwarz)**.** *Let $M \in \mathbb{R}^{d\times d}$, $M \succeq 0$ and symmetric. Let $p_1, \ldots, p_n, q_1, \ldots, q_n \in \mathbb{R}^d$. Then*

$$\langle M, \sum_{i=1}^n p_i q_i^\top \rangle \leq \langle M, \sum_{i=1}^n p_i p_i^\top \rangle^{1/2} \langle M, \sum_{i=1}^n q_i q_i^\top \rangle^{1/2} .$$

In applications, we will have $\sum_i p_i q_i$ as a single block of a larger block matrix containing also the blocks $\sum_i p_i p_i^\top$ and $\sum_i q_i q_i^\top$.

Proof. We first claim that

$$\langle M, \sum_{i=1}^n p_i q_i^\top \rangle \leq \frac{1}{2}\langle M, \sum_{i=1}^n p_i p_i^\top \rangle + \frac{1}{2}\langle M, \sum_{i=1}^n q_i q_i^\top \rangle .$$

To see this, just note that the right-hand side minus the left is exactly

$$\langle M, \sum_{i=1}^n (p_i - q_i)(p_i - q_i)^\top \rangle = \sum_i (p_i - q_i)^\top M (p_i - q_i) \geq 0 .$$

The lemma follows now be applying this inequality to

$$p_i' = \frac{p_i}{\langle M, \sum_{i=1}^n p_i p_i^\top \rangle^{1/2}} \qquad q_i' = \frac{q_i}{\langle M, \sum_{i=1}^n q_i q_i^\top \rangle^{1/2}} . \qquad \square$$

Lemma A.1.2 (Operator Norm Cauchy-Schwarz for Sums)**.** *Let $A_1, \ldots, A_m, B_1, \ldots, B_m$ be real random matrices. Then*

$$\left\| \sum_i \mathbb{E}\, A_i B_i \right\| \leq \left\| \sum_i \mathbb{E}\, A_i^\top A_i \right\|^{1/2} \left\| \sum_i \mathbb{E}\, B_i^\top B_i \right\|^{1/2}.$$

Proof. We have for any unit x, y,

$$\begin{aligned} x^\top \sum_i \mathbb{E}\, A_i B_i x &= \sum_i \mathbb{E}\langle A_i x, B_i y\rangle \\ &\leq \sum_i \mathbb{E}\, \|A_i x\| \|B_i y\| \\ &\leq \sum_i (\mathbb{E}\, \|A_i x\|^2)^{1/2} (\mathbb{E}\, \|B_i x\|^2)^{1/2} \\ &\leq \sqrt{\sum_i \mathbb{E}\, \|A_i x\|^2} \sqrt{\sum_i \mathbb{E}\, \|B_i y\|^2} \\ &= \sqrt{\mathbb{E}\, x^\top \sum_i A_i^\top A_i x} \sqrt{\mathbb{E}\, y^\top \sum_i B_i^\top B_i y} \\ &\leq \left\| \sum_i \mathbb{E}\, A_i^\top A_i \right\|^{1/2} \left\| \sum_i \mathbb{E}\, B_i^\top B_i \right\|^{1/2}. \end{aligned}$$

where the nontrivial inequalities follow from Cauchy-Schwarz for expectations, vectors and scalars, respectively. □

The following lemma allows to argue about the top eigenvector of matrices with spectral gap.

Lemma A.1.3 (Top eigenvector of gapped matrices)**.** *Let M be a symmetric r-by-r matrix and let u, v be a vectors in $\mathbb{R}^r$ with $\|u\| = 1$. Suppose u is a top singular vector of M so that $|\langle u, Mu\rangle| = \|M\|$ and v satisfies for some $\varepsilon > 0$,*

$$\|M - vv^\top\| \leq \|M\| - \varepsilon \cdot \|v\|^2$$

Then, $\langle u, v\rangle^2 \geq \varepsilon \cdot \|v\|^2$.

Proof. We lower bound the quadratic form of $M - vv^\top$ evaluated at u by

$$|\langle u, (M - vv^\top)u\rangle| \geq |\langle u, Mu\rangle| - \langle u, v\rangle^2 = \|M\| - \langle u, v\rangle^2.$$

At the same time, this quadratic form evaluated at u is upper bounded by $\|M\| - \varepsilon \cdot \|v\|^2$. It follows that $\langle u, v\rangle^2 \geq \varepsilon \cdot \|v\|^2$ as desired.

□

The following lemma states that a vector in $\mathbb{R}^{d^2}$ which is close to a symmetric vector $v^{\otimes 2}$, if flattened to a matrix, has top eigenvector correlated with the symmetric vector.

Lemma (Restatement of Lemma 5.4.19)**.** *Let $M \in \mathbb{R}^{d^2 \times d^2}$ be a symmetric matrix with $\|M\| \leq 1$, and let $v \in \mathbb{R}^d$ and $u \in \mathbb{R}^{d^2}$ be vectors. Furthermore, let U be the reshaping of the vector $Mu \in \mathbb{R}^{d^2}$ to a matrix in $\mathbb{R}^{d \times d}$. Fix $c > 0$, and suppose that $\langle Mu, v \otimes v\rangle^2 \geq c^2 \cdot \|u\|^2 \cdot \|v\|^4$. Then U has some left singular vector a and some right singular vector b such that*

$$|\langle a, v\rangle|, |\langle b, v\rangle| \geq c \cdot \|v\| .$$

Furthermore, for any $0 < \alpha < 1$, there are a', b' among the top $\lfloor \frac{1}{\alpha c^2} \rfloor$ singular vectors of U with

$$|\langle a', v\rangle|, |\langle b', v\rangle| \geq \sqrt{1-\alpha} \cdot c \cdot \|v\| .$$

If $c \geq \sqrt{\frac{1}{2}(1+\eta)}$ for some $\eta > 0$, then a, b are amongst the top $\lfloor \frac{(1+\eta)}{\eta c^2} \rfloor$ singular vectors.

Proof. Let $\hat{v} = v/\|v\|$. Let (σ_i, a_i, b_i) be the ith singular value, left and right (unit) singular vectors of U respectively.

Our assumptions imply that

$$\left|\hat{v}^\top U \hat{v}\right| = |\langle Mu, \hat{v} \otimes \hat{v}\rangle| \geq c \cdot \|u\|.$$

Furthermore, we observe that $\|U\|_F = \|Mu\| \leq \|M\| \cdot \|u\|$, and that therefore $\|U\|_F \leq \|u\|$. We thus have that,

$$c \cdot \|u\| \leq \left|\hat{v}^\top U \hat{v}\right| = \left|\sum_{i \in [d]} \sigma_i \cdot \langle \hat{v}, a_i\rangle \langle \hat{v}, b_i\rangle\right| \leq \|u\| \cdot \sqrt{\sum_{i \in [d]} \langle \hat{v}, a_i\rangle^2 \langle \hat{v}, b_i\rangle^2},$$

where to obtain the last inequality we have used Cauchy-Schwarz and our bound on $\|U\|_F$. We may thus conclude that

$$c^2 \leq \sum_{i \in [d]} \langle \hat{v}, a_i\rangle^2 \langle \hat{v}, b_i\rangle^2 \leq \max_{i \in [d]} \langle a_i, \hat{v}\rangle^2 \cdot \sum_{i \in [d]} \langle b_i, \hat{v}\rangle^2 = \max_{i \in [d]} \langle a_i, \hat{v}\rangle^2 , \tag{A.1.1}$$

where we have used the fact that the left singular values of U are orthonormal. The argument is symmetric in the b_i.

Furthermore, we have that

$$\begin{aligned} c^2 \cdot \|u\|^2 &\leq \left|\hat{v}^\top U \hat{v}\right|^2 \\ &= \left|\sum_{i \in [d]} \sigma_i \cdot \langle \hat{v}, a_i\rangle \langle \hat{v}, b_i\rangle\right|^2 \leq \left(\sum_{i \in [d]} \sigma_i^2 \langle \hat{v}, a_i\rangle^2\right) \cdot \left(\sum_{i \in [d]} \langle \hat{v}, b_i\rangle^2\right) = \sum_{i \in [d]} \sigma_i^2 \langle \hat{v}, a_i\rangle^2, \end{aligned}$$

where we have applied Cauchy-Schwarz and the orthonormality of the b_i. In particular,

$$\sum_{i \in [d]} \sigma_i^2 \langle \hat{v}, a_i\rangle^2 \geq c^2 \|u\|^2 \geq c^2 \|U\|_F^2 .$$

On the other hand, let S be the set of $i \in [d]$ for which $\sigma_i^2 \le \alpha c^2 \|U\|_F^2$. By substitution,

$$\sum_{i \in S} \sigma_i^2 \langle \hat{v}, a_i \rangle^2 \le \alpha c^2 \|U\|_F^2 \sum_{i \in S} \langle \hat{v}, a_i \rangle^2 \le \alpha c^2 \|U\|_F^2 \,,$$

where we have used the fact that the right singular vectors are orthonormal. The last two inequalities imply that $S \neq [d]$. Letting $T = [d] \setminus S$, it follows from subtraction that

$$(1-\alpha) c^2 \|U\|_F^2 \le \sum_{i \in T} \sigma_i^2 \langle \hat{v}, a_i \rangle^2 \le \max_{i \in T} \langle \hat{v}, a_i \rangle^2 \sum_{i \in T} \sigma_i^2 = \max_{i \in T} \langle \hat{v}, a_i \rangle^2 \|U\|_F^2 \,,$$

so that $\max_{i \in T} \langle \hat{v}, a_i \rangle^2 \ge (1-\alpha) c^2$. Finally,

$$|T| \cdot \alpha c^2 \|U\|_F^2 \le |T| \cdot \min_{i \in T} \sigma_i^2 \le \sum_{i \in [d]} \sigma_i^2 = \|U\|_F^2 \,,$$

so that $|T| \le \lfloor \frac{1}{\alpha c^2} \rfloor$. Thus, one of the top $\lfloor \frac{1}{\alpha c^2} \rfloor$ right singular vectors a has correlation $|\langle \hat{v}, a \rangle| \ge \sqrt{(1-\alpha)} c$. The same proof holds for the b.

Furthermore, if $c^2 > \frac{1}{2}(1+\eta)$ for some $\eta > 0$, and $(1-\alpha) c^2 > \frac{1}{2}$, then by (A.1.1) it must be that $\max_{i \in T} \langle \hat{v}, a_i \rangle^2 = \max_{i \in [d]} \langle \hat{v}, a_i \rangle^2$, as $\hat{v}$ cannot have square correlation larger than $\frac{1}{2}$ with more than one left singular vector. Taking $\alpha = \frac{\eta}{1+\eta}$ guarantees this. The conclusion follows. □

A.2 Concentration of Scalar Random Variables

We require a number of tools from the literature on concentration of measure.

For scalar-valued polynomials of Gaussians

We need the some concentration bounds for certain polynomials of Gaussian random variables.

The following lemma gives standard bounds on the tails of a standard Gaussian variable—somewhat more precisely than other bounds in this paper. Though there are ample sources, we repeat the proof here for reference.

Lemma A.2.1. *Let $X \sim \mathcal{N}(0,1)$. Then for $t > 0$,*

$$\mathbb{P}\left(X > t\right) \le \frac{e^{-t^2/2}}{t\sqrt{2\pi}},$$

and

$$\mathbb{P}\left(X > t\right) \ge \frac{e^{-t^2/2}}{\sqrt{2\pi}} \cdot \left(\frac{1}{t} - \frac{1}{t^3}\right).$$

Proof. To show the first statement, we apply an integration trick,

$$\begin{aligned}\mathbb{P}(X > t) &= \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-x^2/2} dx \\ &\leq \frac{1}{\sqrt{2\pi}} \int_t^\infty \frac{x}{t} e^{-x^2/2} dx \\ &= \frac{e^{-t^2/2}}{t\sqrt{2\pi}},\end{aligned}$$

where in the third step we have used the fact that $\frac{x}{t} \leq x$ for $t \geq x$. For the second statement, we integrate by parts and repeat the trick,

$$\begin{aligned}\mathbb{P}(X > t) &= \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_t^\infty \frac{1}{x} \cdot x e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \left[-\frac{1}{x} e^{-x^2/2} . \right]_t^\infty - \frac{1}{\sqrt{2\pi}} \int_t^\infty \frac{1}{x^2} \cdot e^{-x^2/2} dx \\ &\geq \frac{1}{\sqrt{2\pi}} \left[-\frac{1}{x} e^{-x^2/2} . \right]_t^\infty - \frac{1}{\sqrt{2\pi}} \int_t^\infty \frac{x}{t^3} \cdot e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \left(\frac{1}{t} - \frac{1}{t^3} \right) e^{-t^2/2}.\end{aligned}$$

This concludes the proof. □

The following is a small modification of Theorem 6.7 from [Jan97] which follows from Remark 6.8 in the same.

Lemma A.2.2. *For each $\ell \geq 1$ there is a universal constant $c_\ell > 0$ such that for every f a degree-ℓ polynomial of standard Gaussian random variables $X_1, \ldots, X_m$ and $t \geq 2$,*

$$\mathbb{P}(|f(X)| > t\,\mathbb{E}\,|f(X)|) \leq e^{-c_\ell t^{2/\ell}}.$$

The same holds (with a different constant c_ℓ) if $\mathbb{E}\,|f(x)|$ is replaced by $(\mathbb{E}\,f(x)^2)^{1/2}$.

In our concentration results, we will need to calculate the expectations of multivariate Gaussian polynomials, many of which share a common form. Below we give an expression for these expectations.

Fact A.2.3. *Let x be a d-dimensional vector with independent identically distributed Gaussian entries with variance σ^2. Let u be a fixed unit vector. Then setting $X = (\|x\|^2 - c)^p \|x\|^{2m} x x^T$, and setting $U = (\|x\|^2 - c)^p \|x\|^{2m} u u^T$, we have*

$$\mathbb{E}[X] = \left(\sum_{0 \leq k \leq p} \binom{p}{k} (-1)^k c^k (d+2) \cdots (d+2p+2m-2k) \sigma^{2(p+m-k+1)} \right) \cdot \mathrm{Id},$$

and

$$\mathbb{E}[U] = \left(\sum_{0 \le k \le p} \binom{p}{k} (-1)^k c^k d(d+2) \cdots (d+2p+2m-2k-2) \sigma^{2(p+m-k)} \right) \cdot uu^T$$

Proof.

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[(\|x\|^2 - c)^p \|x\|^{2m} x_1^2] \cdot \mathrm{Id} \\ &= \mathrm{Id} \cdot \sum_{0 \le k \le p} \binom{p}{k} (-1)^k c^k \, \mathbb{E} \left[\left(\sum_{\ell \in [d]} x_i^2 \right)^{p+m-k} x_1^2 \right] \end{aligned}$$

Since $\left(\sum_{i \in [d]} x_i^2\right)^{p+m-k}$ is symmetric in $x_1, \ldots, x_d$, we have

$$= \mathrm{Id} \cdot \frac{1}{d} \sum_{0 \le k \le p} \binom{p}{k} (-1)^k c^k \, \mathbb{E} \left[\left(\sum_{i \in [d]} x_i^2 \right)^{p+m-k+1} \right]$$

We have reduced the computation to a question of the moments of a Chi-squared variable with d degrees of freedom. Using these moments,

$$\begin{aligned} &= \mathrm{Id} \cdot \frac{1}{d} \sum_{0 \le k \le p} \binom{p}{k} (-1)^k c^k d(d+2) \cdots (d+2p+2m-2k) \sigma^{2(p+m-k+1)} \\ &= \mathrm{Id} \cdot \left(\sum_{0 \le k \le p} \binom{p}{k} (-1)^k c^k (d+2) \cdots (d+2p+2m-2k) \sigma^{2(p+m-k+1)} \right). \end{aligned}$$

A similar computation yields the result about $\mathbb{E}[U]$. □

A.3 Concentration of Matrix-Valued Random Variables

The following proposition relates the top eigenvalue of a matrix to its expected trace.

Proposition (Restatement of Proposition 3.2.4)**.** *Let $n, \ell \in \mathbb{N}$, let $c \in \mathbb{R}$, and let M be an $n \times n$ random matrix. Then*

$$\mathop{\mathbb{E}}_{M}[\mathrm{Tr}((MM^\top)^\ell)] \le \beta \implies \mathbb{P}\left(\|M\| \ge c \cdot \beta^{1/2\ell}\right) \ge 1 - c^{-2\ell}.$$

Proof. For a positive semidefinite matrix P, $\|P\| \le \mathrm{Tr}(P)$. We apply this along with Markov's inequality:

$$\mathbb{P}[\|M\| \ge t] = \mathbb{P}[\|(MM^\top)^\ell\| \ge t^{2\ell}] \le \mathbb{P}[\mathrm{Tr}((MM^\top)^\ell) \ge t^{2\ell}] \le \frac{1}{t^{2\ell}} \mathbb{E}[\mathrm{Tr}((MM^\top)^\ell)] \le \frac{\beta}{t^{2\ell}},$$

and the conclusion follows from taking $t = c\beta^{1/2\ell}$ □

Bound on the norm of a Rademacher matrix

Here, we prove an upper bound on the norm of a Rademacher matrix. Although tighter bounds are known (see e.g. [AKV02], we are off by a constant factor), we include this simpler, looser proof here in an effort to be self-contained.

The following lemma gives a bound on the size of an epsilon net needed to cover the unit sphere.

Lemma A.3.1 (see Lemma 5.2 in [Ver12])**.** *For every $\varepsilon > 0$, the unit Euclidean sphere S^{n-1} equipped with the Euclidean metric has an ε-net with volume at most $\left(1 + \frac{2}{\varepsilon}\right)^n$.*

We are ready to prove our bound.

Theorem A.3.2. *Let A be an $n \times n$ symmetric matrix with i.i.d. Rademacher entries. Then for all $s \geq 0$,*

$$\mathbb{P}\left(|\|A\| - 12\sqrt{n}| \geq s\right) \leq \exp(-t^2/16).$$

Proof. Let Λ be an ε-net over $\mathcal{S}_{n-1}$, with ε to be chosen later. By Lemma A.3.1, we can choose $|\Lambda| \leq (1 + \frac{2}{\varepsilon})^n$. For any fixed $x \in \Lambda$,

$$x^\top A x = \sum_{i<j} 2x_i x_j A_{ij} + \sum_i x_i^2 A_{ii}\,.$$

Each $x_i x_j A_{ij}$ is an independent random variable. We have absolute bounds on the values of each variable, so we can apply a Hoeffding bound to this sum,

$$\mathbb{P}\left(x^\top A x \geq t\right) \leq \exp\left(\frac{-2t^2}{\sum_{i<j}(4x_i x_j)^2 + \sum_i (2x_i^2)^2}\right) = \exp\left(\frac{-2t^2}{8\|x\|_2^4 - 4\|x\|_4^4}\right) \leq \exp\left(\frac{-t^2}{4}\right)\,.$$

Taking a union bound over Λ, we have

$$\mathbb{P}\left(\max_{x\in\Lambda} x^\top A x \geq t\right) \leq (1 + \frac{2}{\varepsilon})^n \cdot \exp\left(\frac{-t^2}{4}\right)\,.$$

To extend the bound to any point $y \in \mathcal{S}_{n-1}$, let y be the maximizer of $y^\top A y$. We note that there must exist some $x \in \Lambda$ so that $\|x - y\| \leq \varepsilon$ by assumption. We have

$$\left|y^\top A y - x^\top A x\right| = \left|y^\top A(y - x) - x^\top A(x - y)\right| \leq 2\varepsilon\|A\|\,,$$

which by the triangle inequality implies

$$\|A\| = y^\top A y \leq \max_{x\in\Lambda}\{x^\top A x\} + 2\varepsilon\|A\| \qquad \Longrightarrow \qquad \|A\| \leq \frac{1}{1 - 2\varepsilon}\max_{x\in\Lambda}\{x^\top A x\}.$$

Taking $\varepsilon = 1/4$ and $t = 2\sqrt{n\log(1 + \frac{2}{\varepsilon})} + s$ concludes the proof. □

Matrix Chernoff Bounds for Gaussian Matrices

On several occasions we will need to apply a Matrix-Bernstein-like theorem to a sum of matrices with an unfortunate tail. To this end, we prove a "truncated Matrix Bernstein Inequality." Our proof uses an standard matrix Bernstein inequality as a black box. The study of inequalities of this variety—on tails of sums of independent matrix-valued random variables— was initiated by Ahlswede and Winter [AW02]. The excellent survey of Tropp [Tro12] provides many results of this kind.

In applications of the following the operator norms of the summands $X_1, \ldots, X_n$ have well-behaved tails and so the truncation is a routine formality. Two corollaries following the proposition and its proof capture truncation for all the matrices we encounter in the present work.

Proposition A.3.3 (Truncated Matrix Bernstein)**.** *Let $X_1, \ldots, X_n \in \mathbb{R}^{d_1 \times d_2}$ be independent random matrices, and suppose that*

$$\mathbb{P}\left[\|X_i - \mathbb{E}[X_i]\|_{op} \geq \beta\right] \leq p \text{ for all } i \in [n].$$

Furthermore, suppose that for each X_i,

$$\|\mathbb{E}[X_i] - \mathbb{E}[X_i \mathbb{I}[\|X_i\|_{op} < \beta]]\| \leq q.$$

Denote

$$\sigma^2 = \max\left\{\left\|\sum_{i\in[n]} \mathbb{E}\left[X_i X_i^T\right] - \mathbb{E}\left[X_i\right]\mathbb{E}\left[X_i^T\right]\right\|_{op}, \left\|\sum_{i\in[n]} \mathbb{E}\left[X_i^T X_i\right] - \mathbb{E}\left[X_i\right]^T \mathbb{E}\left[X_i\right]\right\|_{op}\right\}.$$

Then for $X = \sum_{i\in[n]} X_i$, we have

$$\mathbb{P}\left[\|X - \mathbb{E}[X]\|_{op} \geq t\right] \leq n \cdot p + (d_1 + d_2) \cdot \exp\left(\frac{-(t - nq)^2}{2(\sigma^2 + \beta(t - nq)/3)}\right).$$

Proof. For simplicity we start by centering the variables X_i. Let $\tilde{X}_i = X_i - \mathbb{E}\, X_i$ and $\tilde{X} = \sum_{i\in[n]} \tilde{X}_i$ The proof proceeds by a straightforward application of the noncommutative Bernstein's Inequality. We define variables $Y_1, \ldots, Y_n$, which are the truncated counterparts of the $\tilde{X}_i$s in the following sense:

$$Y_i = \begin{cases} \tilde{X}_i & \|\tilde{X}_i\|_{op} < \beta, \\ 0 & \text{otherwise.} \end{cases}$$

Define $Y = \sum_{i\in[n]} Y_i$. We claim that

$$\left\|\sum_i \mathbb{E}\, Y_i Y_i^T - \mathbb{E}[Y_i]\,\mathbb{E}[Y_i]^T\right\|_{\text{op}} \leq \left\|\sum_i \mathbb{E}\, \tilde{X}_i \tilde{X}_i^T\right\|_{\text{op}} \leq \sigma^2 \text{ and} \tag{A.3.1}$$

$$\left\| \sum_i \mathbb{E}\, Y_i^T Y_i - \mathbb{E}[Y_i]^T\, \mathbb{E}[Y_i] \right\|_{\text{op}} \leq \left\| \sum_i \mathbb{E}\, \tilde{X}_i^T \tilde{X}_i \right\|_{\text{op}} \leq \sigma^2\,, \tag{A.3.2}$$

which, together with the fact that $\|Y_i\| \leq \beta$ almost surely, will allow us to apply the non-commutative Bernstein's inequality to Y. To see (A.3.1) ((A.3.2) is similar), we expand $\mathbb{E}\, Y_i Y_i^T$ as

$$\mathbb{E}\, Y_i Y_i^T = \mathbb{P}\left[\left\| \tilde{X}_i \right\|_{\text{op}} < \beta \right] \mathbb{E}\left[\tilde{X}_i \tilde{X}_i^T \;\middle|\; \left\| \tilde{X}_i \right\|_{\text{op}} < \beta \right].$$

Additionally expanding $\mathbb{E}\left[\tilde{X}_i \tilde{X}_i^T \right]$ as

$$\begin{aligned} \mathbb{E}\left[\tilde{X}_i \tilde{X}_i^T \right] = \mathbb{P}\left[\left\| \tilde{X}_i \right\|_{\text{op}} < \beta \right] \mathbb{E}\left[\tilde{X}_i \tilde{X}_i^T \;\middle|\; \left\| \tilde{X}_i \right\|_{\text{op}} < \beta \right] \\ + \mathbb{P}\left[\left\| \tilde{X}_i \right\|_{\text{op}} \geq \beta \right] \mathbb{E}\left[\tilde{X}_i \tilde{X}_i^T \;\middle|\; \left\| \tilde{X}_i \right\|_{\text{op}} \geq \beta \right], \end{aligned}$$

we note that $\mathbb{E}[\tilde{X}_i \tilde{X}_i^T \mid \left\| \tilde{X}_i \right\|_{\text{op}} \geq \beta]$ is PSD. Thus, $\mathbb{E}[Y_i Y_i^T] \succeq \mathbb{E}[X_i X_i^T]$. But by definition $\mathbb{E}[Y_i Y_i^T]$ is still PSD (and hence $\left\| \sum_i \mathbb{E}[Y_i Y_i^T] \right\|_{\text{op}}$ is given by the maximum eigenvalue of $\mathbb{E}[Y_i Y_i^T]$), so

$$\left\| \sum_i \mathbb{E}\, Y_i Y_i^T \right\|_{\text{op}} \leq \left\| \sum_i \mathbb{E}\, \tilde{X}_i \tilde{X}_i^T \right\|_{\text{op}}.$$

Also PSD are $\mathbb{E}[Y_i]\, \mathbb{E}[Y_i]^T$ and $\mathbb{E}[(Y_i - \mathbb{E}[Y_i])(Y_i - \mathbb{E}[Y_i])^T] = \mathbb{E}[Y_i Y_i^T] - \mathbb{E}[Y_i]\, \mathbb{E}[Y_i]^T$. By the same reasoning again, then, we get $\left\| \sum_i \mathbb{E}\, Y_i Y_i^T - \mathbb{E}[Y_i]\, \mathbb{E}[Y_i]^T \right\|_{\text{op}} \leq \left\| \sum_i \mathbb{E}[Y_i Y_i^T] \right\|_{\text{op}}$. Putting this all together gives (A.3.1).

Now we are ready to apply the non-commutative Bernstein's inequality to Y. We have

$$\mathbb{P}\left[\|Y - \mathbb{E}[Y]\|_{op} \geq \alpha \right] \leq (d_1 + d_2) \cdot \exp\left(\frac{-\alpha^2/2}{\sigma^2 + \beta \cdot \alpha/3} \right).$$

Now, we have

$$\begin{aligned} \mathbb{P}\left[\|X - \mathbb{E}[X]\|_{op} \geq t \right] &= \mathbb{P}\left[\|X - \mathbb{E}[X]\|_{op} \geq t \mid X = Y \right] \cdot \mathbb{P}\left[X = Y \right] \\ &\quad + \mathbb{P}\left[\|X - \mathbb{E}[X]\|_{op} \geq t \mid X \neq Y \right] \cdot \mathbb{P}\left[X \neq Y \right], \\ &\leq \mathbb{P}\left[\|X - \mathbb{E}[X]\|_{op} \geq t \mid X = Y \right] + n \cdot p \end{aligned}$$

by a union bound over the events $\{X_i \neq Y_i\}$. It remains to bound the conditional probability $\mathbb{P}\left[\|X - \mathbb{E}[X]\|_{op} \geq t \mid X = Y \right]$. By assumption, $\| \mathbb{E}[X] - \mathbb{E}[Y] \|_{op} \leq nq$, and so by the triangle inequality,

$$\|X - \mathbb{E}[X]\|_{op} \leq \|X - \mathbb{E}[Y]\|_{op} + \| \mathbb{E}[Y] - \mathbb{E}[X] \|_{op} \leq \|X - \mathbb{E}[Y]\|_{op} + nq.$$

Thus,

$$\begin{aligned}\mathbb{P}\left[\|X - \mathbb{E}[X]\|_{op} \geq t \mid X = Y\right] &\leq \mathbb{P}\left[\|X - \mathbb{E}[Y]\|_{op} + nq \geq t \mid X = Y\right] \\ &= \mathbb{P}\left[\|Y - \mathbb{E}[Y]\|_{op} \geq t - nq \mid X = Y\right].\end{aligned}$$

Putting everything together and setting $\alpha = t - nq$,

$$\mathbb{P}[\|X - \mathbb{E}[X]\|_{op} \geq t] \leq n \cdot p + (d_1 + d_2) \cdot \exp\left(\frac{-(t-nq)^2/2}{\sigma^2 + \beta(t-nq)/3}\right),$$

as desired. □

The following lemma helps achieve the assumptions of Proposition A.3.3 easily for a useful class of thin-tailed random matrices.

Lemma A.3.4. *Suppose that X is a matrix whose entries are polynomials of constant degree ℓ in unknowns x, which we evaluate at independent Gaussians. Let $f(x) := \|X\|_{op}$ and $g(x) := \|XX^T\|_{op}$, and either f is itself a polynomial in x of degree at most 2ℓ or g is a polynomial in x of degree at most 4ℓ. Then if $\beta = R \cdot \alpha$ for $\alpha \geq \min\{\mathbb{E}\left[|f(x)|\right], \sqrt{\mathbb{E}\left[g(x)\right]}\}$ and $R = \text{polylog}(n)$,*

$$\mathbb{P}\left(\|X\|_{op} \geq \beta\right) \leq n^{-\log n}, \tag{A.3.3}$$

and

$$\mathbb{E}\left[\|X \cdot \mathbb{I}\{\|X\|_{op} \geq \beta\}\|_{op}\right] \leq (\beta + \alpha)n^{-\log n}. \tag{A.3.4}$$

Proof. We begin with (A.3.3). Either $f(x)$ is a polynomial of degree at most 2ℓ, or $g(x)$ is a polynomial of degree at most 4ℓ in Gaussian variables. We can thus use Lemma A.2.2 to obtain the following bound,

$$\mathbb{P}\left(|f(x)| \geq t\alpha\right) \leq \exp\left(-ct^{1/(2\ell)}\right), \tag{A.3.5}$$

where c is a universal constant. Taking $t = R = \text{polylog}(n)$ gives us (A.3.3).

We now address (A.3.4). To this end, let $p(t)$ and $P(t)$ be the probability density function and cumulative density function of $\|X\|_{op}$, respectively. We apply Jensen's inequality and instead bound

$$\|\,\mathbb{E}\left[X\,\mathbb{I}\{\|X\|_{op} \geq \beta\}\right]\,\| \leq \mathbb{E}\left[\|X\|_{op}\,\mathbb{I}\{\|X\|_{op} \geq \beta\}\right] \;= \int_0^\infty t \cdot \mathbb{I}\{t \geq \beta\}p(t)dt$$

since the indicator is 0 for $t \leq \beta$,

$$= \int_\beta^\infty (-t)(-p(t))dt$$

integrating by parts,

$$= -t \cdot (1 - P(t))\Big|_{\beta}^{\infty} + \int_{\beta}^{\infty} (1 - P(t))dt$$

and using the equality of $1 - P(t)$ with $\mathbb{P}(\|X\|_{op} \geq t)$ along with (A.3.3),

$$\leq \beta n^{-\log n} + \int_{\beta}^{\infty} \mathbb{P}(\|X\|_{op} \geq t)dt$$

Applying the change of variables $t = \alpha s$ so as to apply (A.3.5),

$$= \beta n^{-\log n} + \alpha \int_{R}^{\infty} \mathbb{P}(\|X\|_{op} \geq \alpha s)ds$$
$$\leq \beta n^{-\log n} + \alpha \int_{R}^{\infty} \exp(-cs^{1/(2\ell)})ds$$

Now applying a change of variables so $s = (\frac{u \log n}{c})^{2\ell}$,

$$= \beta n^{-\log n} + \alpha \int_{\frac{cR^{1/(2\ell)}}{\log n}}^{\infty} n^{-u} \cdot 2\ell \left(\frac{\log n}{c}\right)^{2\ell} u^{2\ell-1} du$$
$$\leq \beta n^{-\log n} + \alpha \int_{\frac{cR^{1/(2\ell)}}{\log n}}^{\infty} n^{-u/2} du\,,$$

where we have used the assumption that ℓ is constant. We can approximate this by a geometric sum,

$$\leq \beta n^{-\log n} + \alpha \sum_{u=\frac{cR^{1/(2\ell)}}{\log n}}^{\infty} n^{-u/2}$$
$$\leq \beta n^{-\log n} + \alpha \cdot n^{-cR^{1/(2\ell)}/(2\log n)}$$

Evaluating at $R = \text{polylog}\, n$ for a sufficiently large polynomial in the log gives us

$$\mathbb{E}\left[\|X \cdot \mathbb{I}\{\|X\|_{op} \geq \beta\}\|_{op}\right] \leq (\beta + \alpha)n^{-\log n},$$

as desired. □

Bibliography

[AAK+07] Noga Alon, Alexandr Andoni, Tali Kaufman, Kevin Matulef, Ronitt Rubinfeld, and Ning Xie, *Testing k-wise and almost k-wise independence*, STOC, ACM, 2007, pp. 496–505.

[AAM+11] Noga Alon, Sanjeev Arora, Rajsekar Manokaran, Dana Moshkovitz, and Omri Weinstein, *Inapproximability of densest κ-subgraph from average case hardness*, Unpublished manuscript (2011).

[ABBG10] Sanjeev Arora, Boaz Barak, Markus Brunnermeier, and Rong Ge, *Computational complexity and information asymmetry in financial products (extended abstract)*, ICS, Tsinghua University Press, 2010, pp. 49–65.

[ABC11] Per Austrin, Mark Braverman, and Eden Chlamtac, *Inapproximability of np-complete variants of nash equilibrium*, APPROX-RANDOM, Lecture Notes in Computer Science, vol. 6845, Springer, 2011, pp. 13–25.

[ABS15] Sanjeev Arora, Boaz Barak, and David Steurer, *Subexponential algorithms for unique games and related problems*, J. ACM **62** (2015), no. 5, 42:1–42:25.

[ABW10a] Benny Applebaum, Boaz Barak, and Avi Wigderson, *Public-key cryptography from different assumptions*, STOC, ACM, 2010, pp. 171–180.

[ABW10b] _____, *Public-key cryptography from different assumptions*, STOC, ACM, 2010, pp. 171–180.

[Ach09] Dimitris Achlioptas, *Random satisfiability*, Handbook of Satisfiability (Armin Biere, Marijn Heule, Hans van Maaren, and Toby Walsh, eds.), Frontiers in Artificial Intelligence and Applications, vol. 185, IOS Press, 2009, pp. 245–270.

[AFH+15] Anima Anandkumar, Dean P. Foster, Daniel J. Hsu, Sham M. Kakade, and Yi-Kai Liu, *A spectral algorithm for latent dirichlet allocation*, Algorithmica **72** (2015), no. 1, 193–214.

[AGH+14] Animashree Anandkumar, Rong Ge, Daniel J. Hsu, Sham M. Kakade, and Matus Telgarsky, *Tensor decompositions for learning latent variable models*, Journal of Machine Learning Research **15** (2014), no. 1, 2773–2832.

[AGH+15] Anima Anandkumar, Rong Ge, Daniel J. Hsu, Sham M. Kakade, and Matus Telgarsky, *Tensor decompositions for learning latent variable models (A survey for ALT)*, ALT, Lecture Notes in Computer Science, vol. 9355, Springer, 2015, pp. 19–38.

[AGHK14] Animashree Anandkumar, Rong Ge, Daniel J. Hsu, and Sham M. Kakade, *A tensor approach to learning mixed membership community models*, Journal of Machine Learning Research **15** (2014), no. 1, 2239–2312.

[AGJ14] Anima Anandkumar, Rong Ge, and Majid Janzamin, *Analyzing tensor power method dynamics: Applications to learning overcomplete latent variable models*, CoRR **abs/1411.1488** (2014).

[AGJ15] Animashree Anandkumar, Rong Ge, and Majid Janzamin, *Learning overcomplete latent variable models through tensor methods*, COLT, JMLR Workshop and Conference Proceedings, vol. 40, JMLR.org, 2015, pp. 36–112.

[AKS98] Noga Alon, Michael Krivelevich, and Benny Sudakov, *Finding a large hidden clique in a random graph*, Random Struct. Algorithms **13** (1998), no. 3-4, 457–466.

[AKV02] Noga Alon, Michael Krivelevich, and Van H. Vu, *On the concentration of eigenvalues of random symmetric matrices*, Israel Journal of Mathematics **131** (2002).

[AOW15] Sarah R. Allen, Ryan O'Donnell, and David Witmer, *How to refute a random CSP*, FOCS, IEEE Computer Society, 2015, pp. 689–708.

[ARV09] Sanjeev Arora, Satish Rao, and Umesh V. Vazirani, *Expander flows, geometric embeddings and graph partitioning*, J. ACM **56** (2009), no. 2, 5:1–5:37.

[AW02] Rudolf Ahlswede and Andreas J. Winter, *Strong converse for identification via quantum channels*, IEEE Trans. Information Theory **48** (2002), no. 3, 569–579.

[Ban10] Nikhil Bansal, *Constructive algorithms for discrepancy minimization*, FOCS, IEEE Computer Society, 2010, pp. 3–10.

[Bar14] Boaz Barak, *Sum of squares upper bounds, lower bounds, and open questions*, Lecture Notes (2014), http://www.boazbarak.org/sos/files/all–notes.pdf.

[BB02] Eli Ben-Sasson and Yonatan Bilu, *A gap in average proof complexity*, Electronic Colloquium on Computational Complexity (ECCC) (2002), no. 003.

[BBH+12] Boaz Barak, Fernando G. S. L. Brandão, Aram Wettroth Harrow, Jonathan A. Kelner, David Steurer, and Yuan Zhou, *Hypercontractivity, sum-of-squares proofs, and their applications*, STOC, ACM, 2012, pp. 307–326.

[BCK15] Boaz Barak, Siu On Chan, and Pravesh K. Kothari, *Sum of squares lower bounds from pairwise independence*, in Servedio and Rubinfeld [SR15], pp. 97–106.

[BCMV14] Aditya Bhaskara, Moses Charikar, Ankur Moitra, and Aravindan Vijayaraghavan, *Smoothed analysis of tensor decompositions*, STOC, ACM, 2014, pp. 594–603.

[BCV+12] Aditya Bhaskara, Moses Charikar, Aravindan Vijayaraghavan, Venkatesan Guruswami, and Yuan Zhou, *Polynomial integrality gaps for strong SDP relaxations of densest* k*-subgraph*, Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2012, Kyoto, Japan, January 17-19, 2012, 2012, pp. 388–405.

[BE76] Bella Bollobas and Paul Erdös, *Cliques in random graphs*, Mathematical Proceedings of the Cambridge Philosophical Society **80** (1976), 419–427.

[BGL16] Vijay Bhattiprolu, Venkatesan Guruswami, and Euiwoong Lee, *Certifying random polynomials over the unit sphere via sum-of-s uares hierarchy*, preprint, 2016.

[BHK+16] Boaz Barak, Samuel B. Hopkins, Jonathan A. Kelner, Pravesh Kothari, Ankur Moitra, and Aaron Potechin, *A nearly tight sum-of-squares lower bound for the planted clique problem*, FOCS, IEEE Computer Society, 2016, pp. 428–437.

[BKPS98] Paul Beame, Richard M. Karp, Toniann Pitassi, and Michael E. Saks, *On the complexity of unsatisfiability proofs for random* k*-cnf formulas*, STOC, ACM, 1998, pp. 561–571.

[BKS13] Boaz Barak, Guy Kindler, and David Steurer, *On the optimality of semidefinite relaxations for average-case and generalized constraint satisfaction*, ITCS, ACM, 2013, pp. 197–214.

[BKS14] Boaz Barak, Jonathan A. Kelner, and David Steurer, *Rounding sum-of-squares relaxations*, STOC, ACM, 2014, pp. 31–40.

[BKS15] ———, *Dictionary learning and tensor decomposition via the sum-of-squares method*, STOC, ACM, 2015, pp. 143–151.

[BM16] Boaz Barak and Ankur Moitra, *Noisy tensor completion via the sum-of-squares hierarchy*, COLT, JMLR Workshop and Conference Proceedings, vol. 49, JMLR.org, 2016, pp. 417–445.

[BMH12] N.W. Bauer, P.J.H. Maas, and W.P.M.H. Heemels, *Stability analysis of networked control systems: A sum of squares approach*, Automatica **48** (2012), no. 8, 1514 – 1524.

[BMMN11] Mark Braverman, Konstantin Makarychev, Yury Makarychev, and Assaf Naor, *The grothendieck constant is strictly smaller than krivine's bound*, FOCS, IEEE Computer Society, 2011, pp. 453–462.

[BR13] Quentin Berthet and Philippe Rigollet, *Complexity theoretic lower bounds for sparse principal component detection*, Conference on Learning Theory, 2013, pp. 1046–1066.

[BRS11] Boaz Barak, Prasad Raghavendra, and David Steurer, *Rounding semidefinite programming hierarchies via global correlation*, FOCS, IEEE Computer Society, 2011, pp. 472–481.

[BV04] Stephen Boyd and Lieven Vandenberghe, *Convex optimization*, Cambridge University Press, New York, NY, USA, 2004.

[BV09] S. Charles Brubaker and Santosh Vempala, *Random tensors and planted cliques*, APPROX-RANDOM, Lecture Notes in Computer Science, vol. 5687, Springer, 2009, pp. 406–419.

[CCF10] Amin Coja-Oghlan, Colin Cooper, and Alan M. Frieze, *An efficient sparse regularity concept*, SIAM J. Discrete Math. **23** (2010), no. 4, 2000–2034.

[CGL07] Amin Coja-Oghlan, Andreas Goerdt, and André Lanka, *Strong refutation heuristics for random k-sat*, Combinatorics, Probability & Computing **16** (2007), no. 1, 5–28.

[Cha96] J. T. Chang, *Full reconstruction of markov models on evolutionary trees: Identifiability and consistency*, Math Biosci. **137** (1996), 51–73.

[CLP02] Andrea Crisanti, Luca Leuzzi, and Giorgio Parisi, *The 3-sat problem with large number of clauses in the ∞-replica symmetry breaking scheme*, Journal of Physics A: Mathematical and General **35** (2002), 481.

[CS88] Vašek Chvátal and Endre Szemerédi, *Many hard examples for resolution*, J. ACM **35** (1988), no. 4, 759–768.

[Dan16] Amit Daniely, *Complexity theoretic limitations on learning halfspaces*, STOC, 2016, pp. 105–117.

[DGGP14] Yael Dekel, Ori Gurel-Gurevich, and Yuval Peres, *Finding hidden cliques in linear time with high probability*, Combinatorics, Probability and Computing **23** (2014), no. 01, 29–49.

[dGJL04] Alexandre d'Aspremont, Laurent El Ghaoui, Michael I. Jordan, and Gert R. G. Lanckriet, *A direct formulation for sparse PCA using semidefinite programming*, NIPS, 2004, pp. 41–48.

[DH14] Laurent Demanet and Paul Hand, *Scaling law for recovering the sparsest element in a subspace*, Information and Inference **3** (2014), no. 4, 295–309.

[dlPMS95] Victor H de la Peña and Stephen J Montgomery-Smith, *Decoupling inequalities for the tail probabilities of multivariate u-statistics*, The Annals of Probability (1995), 806–816.

[DLS13] Amit Daniely, Nati Linial, and Shai Shalev-Shwartz, *More data speeds up training time in learning halfspaces over sparse vectors*, NIPS, 2013, pp. 145–153.

[DLS14a] ———, *The complexity of learning halfspaces using generalized linear methods*, COLT, JMLR Workshop and Conference Proceedings, vol. 35, JMLR.org, 2014, pp. 244–286.

[DLS14b] ———, *From average case complexity to improper learning complexity*, STOC, ACM, 2014, pp. 441–448.

[dlVK07] Wenceslas Fernandez de la Vega and Claire Kenyon-Mathieu, *Linear programming relaxations of maxcut*, SODA, SIAM, 2007, pp. 53–61.

[DM15a] Yash Deshpande and Andrea Montanari, *Finding hidden cliques of size\ sqrt {N/e} in nearly linear time*, Foundations of Computational Mathematics (2015), 1–60.

[DM15b] Yash Deshpande and Andrea Montanari, *Improved sum-of-squares lower bounds for hidden clique and hidden submatrix problems*, COLT, JMLR Workshop and Conference Proceedings, vol. 40, JMLR.org, 2015, pp. 523–562.

[DSS15] Jian Ding, Allan Sly, and Nike Sun, *Proof of the satisfiability conjecture for large k*, in Servedio and Rubinfeld [SR15], pp. 59–68.

[Fei02] Uriel Feige, *Relations between average case complexity and approximation complexity*, Proceedings of the 17th Annual IEEE Conference on Computational Complexity, Montréal, Québec, Canada, May 21-24, 2002, IEEE Computer Society, 2002, p. 5.

[FG01] Joel Friedman and Andreas Goerdt, *Recognizing more unsatisfiable random 3-sat instances efficiently*, Automata, Languages and Programming, 28th International Colloquium, ICALP 2001, Crete, Greece, July 8-12, 2001, Proceedings (Fernando Orejas, Paul G. Spirakis, and Jan van Leeuwen, eds.), Lecture Notes in Computer Science, vol. 2076, Springer, 2001, pp. 310–321.

[FGK05] Joel Friedman, Andreas Goerdt, and Michael Krivelevich, *Recognizing more unsatisfiable random k-sat instances efficiently*, SIAM J. Comput. **35** (2005), no. 2, 408–430.

[FGR+12] Vitaly Feldman, Elena Grigorescu, Lev Reyzin, Santosh Vempala, and Ying Xiao, *Statistical algorithms and a lower bound for planted clique*, Electronic Colloquium on Computational Complexity (ECCC) **19** (2012), 64.

[FK81] Z. Füredi and J. Komlós, *The eigenvalues of random symmetric matrices*, Combinatorica **1** (1981), no. 3, 233–241.

[FK00] Uriel Feige and Robert Krauthgamer, *Finding and certifying a large hidden clique in a semirandom graph*, Random Struct. Algorithms **16** (2000), no. 2, 195–208.

[FK03] ———, *The probable value of the lovasz-schrijver relaxations for maximum independent set*, SIAM Journal on Computing **32** (2003), 2003.

[FK08] Alan M. Frieze and Ravi Kannan, *A new approach to the planted clique problem*, IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2008, December 9-11, 2008, Bangalore, India, 2008, pp. 187–198.

[FKO06] Uriel Feige, Jeong Han Kim, and Eran Ofek, *Witnesses for non-satisfiability of dense random 3cnf formulas*, FOCS, IEEE Computer Society, 2006, pp. 497–508.

[FR10] Uriel Feige and Dorit Ron, *Finding hidden cliques in linear time*, DMTCS Proceedings (2010), no. 01, 189–204.

[Fu98] Xudong Fu, *On the complexity of proof systems*, Ph.D. thesis, University of Toronto, 1998.

[Gal14] François Le Gall, *Powers of tensors and fast matrix multiplication*, ISSAC, ACM, 2014, pp. 296–303.

[GHK15] Rong Ge, Qingqing Huang, and Sham M. Kakade, *Learning mixtures of Gaussians in high dimensions [extended abstract]*, STOC'15—Proceedings of the 2015 ACM Symposium on Theory of Computing, ACM, New York, 2015, pp. 761–770. MR 3388256

[GK01] Andreas Goerdt and Michael Krivelevich, *Efficient recognition of random unsatisfiable k-sat instances by spectral methods*, STACS 2001, 18th Annual Symposium on Theoretical Aspects of Computer Science, Dresden, Germany, February 15-17, 2001, Proceedings (Afonso Ferreira and Horst Reichel, eds.), Lecture Notes in Computer Science, vol. 2010, Springer, 2001, pp. 294–304.

[GM75] Geoffrey R. Grimmett and Colin J. H. McDiarmid, *On colouring random graphs*, Mathematical Proceedings of the Cambridge Philosophical Society **77** (1975), 313–324.

[GM15] Rong Ge and Tengyu Ma, *Decomposing overcomplete 3rd order tensors using sum-of-squares algorithms*, APPROX-RANDOM, LIPIcs, vol. 40, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2015, pp. 829–849.

[Gri01a] Dima Grigoriev, *Complexity of positivstellensatz proofs for the knapsack*, Computational Complexity **10** (2001), no. 2, 139–154.

[Gri01b] ———, *Linear lower bound on degrees of positivstellensatz calculus proofs for the parity*, Theor. Comput. Sci. **259** (2001), no. 1-2, 613–622.

[GS12] Venkatesan Guruswami and Ali Kemal Sinop, *Faster SDP hierarchy solvers for local rounding algorithms*, FOCS, IEEE Computer Society, 2012, pp. 197–206.

[GVX14] Navin Goyal, Santosh Vempala, and Ying Xiao, *Fourier PCA and robust tensor decomposition*, STOC, ACM, 2014, pp. 584–593.

[GW94] Michel X. Goemans and David P. Williamson, *.879-approximation algorithms for MAX CUT and MAX 2sat*, STOC, ACM, 1994, pp. 422–431.

[Har70] Richard A Harshman, *Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis*, UCLA Working Papers in Phonetics **16** (1970), 1–84.

[Hås90] Johan Håstad, *Tensor rank is np-complete*, J. Algorithms **11** (1990), no. 4, 644–654.

[Hås96] ———, *Clique is hard to approximate within* $n^{1\text{-}epsilon}$, FOCS, IEEE Computer Society, 1996, pp. 627–636.

[HK09] Elad Hazan and Robert Krauthgamer, *How hard is it to approximate the best nash equilibrium?*, SODA, SIAM, 2009, pp. 720–727.

[HKP15] Samuel B. Hopkins, Pravesh K. Kothari, and Aaron Potechin, *Sos and planted clique: Tight analysis of mpw moments and an optimal lower bound at degree four*, Manuscript (2015).

[HKZ12] Daniel J. Hsu, Sham M. Kakade, and Tong Zhang, *A spectral algorithm for learning hidden markov models*, J. Comput. Syst. Sci. **78** (2012), no. 5, 1460–1480.

[HL13] Christopher J. Hillar and Lek-Heng Lim, *Most tensor problems are NP-hard*, J. ACM **60** (2013), no. 6, 45.

[HM13] Aram Wettroth Harrow and Ashley Montanaro, *Testing product states, quantum merlin-arthur games and tensor optimization*, J. ACM **60** (2013), no. 1, 3.

[HSS15] Samuel B. Hopkins, Jonathan Shi, and David Steurer, *Tensor principal component analysis via sum-of-square proofs*, COLT, JMLR Workshop and Conference Proceedings, vol. 40, JMLR.org, 2015, pp. 956–1006.

[HSSS16] Samuel B. Hopkins, Tselil Schramm, Jonathan Shi, and David Steurer, *Fast spectral algorithms from sum-of-squares proofs: tensor decomposition and planted sparse vectors*, STOC, ACM, 2016, pp. 178–191.

[Jan97] Svante Janson, *Gaussian hilbert spaces*, Cambridge Tracts in Mathematics, Cambridge University Press, 1997.

[Jer92] Mark Jerrum, *Large cliques elude the metropolis process.*, Random Struct. Algorithms **3** (1992), no. 4, 347–360.

[Kar76] Richard Karp, *The probabilistic analysis of some combinatorial search algorithms*, Algorithms and Complexity: New Directions and Recent Results (1976), 1–19.

[Kar10] Richard M. Karp, *Reducibility among combinatorial problems*, 50 Years of Integer Programming, Springer, 2010, pp. 219–241.

[KB09] Tamara G. Kolda and Brett W. Bader, *Tensor decompositions and applications*, SIAM Review **51** (2009), no. 3, 455–500.

[Kho02] Subhash Khot, *On the power of unique 2-prover 1-round games*, STOC, ACM, 2002, pp. 767–775.

[KKMO04] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell, *Optimal inapproximability results for max-cut and other 2-variable csps?*, FOCS, IEEE Computer Society, 2004, pp. 146–154.

[KMOW17] Pravesh K. Kothari, Ryuhei Mori, Ryan O'Donnell, and David Witmer, *Sum of squares lower bounds for refuting any CSP*, CoRR **abs/1701.04521** (2017).

[KS09] Subhash Khot and Rishi Saket, *SDP integrality gaps with local ell_1-embeddability*, 50th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2009, October 25-27, 2009, Atlanta, Georgia, USA, 2009, pp. 565–574.

[Kuc95] Ludek Kucera, *Expected complexity of graph partitioning problems*, Discrete Applied Mathematics **57** (1995), no. 2-3, 193–212.

[Las01] Jean B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM Journal on Optimization **11** (2001), no. 3, 796–817.

[LCC07] Lieven De Lathauwer, Joséphine Castaing, and Jean-François Cardoso, *Fourth-order cumulant-based blind identification of underdetermined mixtures*, IEEE Trans. Signal Processing **55** (2007), no. 6-2, 2965–2973.

[LM00] B. Laurent and P. Massart, *Adaptive estimation of a quadratic functional by model selection*, Ann. Statist. **28** (2000), no. 5, 1302–1338.

[LR15] Elaine Levey and Thomas Rothvoss, *A lasserre-based* $(1+\varepsilon)$*-approximation for* $Pm \mid p_j = 1, prec \mid C_{\max}$, CoRR **abs/1509.07808** (2015).

[Mat76] David Matula, *The largest clique size in a random graph*, Tech. report, Southern Methodist University, Dallas, 1976.

[MOO05] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz, *Noise stability of functions with low in.uences invariance and optimality*, FOCS, IEEE Computer Society, 2005, pp. 21–30.

[MPW15] Raghu Meka, Aaron Potechin, and Avi Wigderson, *Sum-of-squares lower bounds for planted clique*, STOC, ACM, 2015, pp. 87–96.

[MR06] Elchanan Mossel and Sébastien Roch, *Learning nonsingular phylogenies and hidden markov models*, Ann. Appl. Probab. **16** (2006), no. 2, 583–614.

[MRZ14] Andrea Montanari, Daniel Reichman, and Ofer Zeitouni, *On the limitation of spectral methods: From the gaussian hidden clique problem to rank one perturbations of gaussian tensors*, Arxiv:1411.6149 (2014).

[MS16] Andrea Montanari and Nike Sun, *Spectral algorithms for tensor completion*, CoRR **abs/1612.07866** (2016).

[MSS16] Tengyu Ma, Jonathan Shi, and David Steurer, *Polynomial-time tensor decompositions with sum-of-squares*, FOCS, IEEE Computer Society, 2016, pp. 438–446.

[MW13] Raghu Meka and Avi Wigderson, *Association schemes, non-commutative polynomial concentration, and sum-of-squares lower bounds for planted clique*, Electronic Colloquium on Computational Complexity (ECCC) **20** (2013), 105.

[MW15] Tengyu Ma and Avi Wigderson, *Sum-of-squares lower bounds for sparse PCA*, NIPS, 2015, pp. 1612–1620.

[NJW01] Andrew Y. Ng, Michael I. Jordan, and Yair Weiss, *On spectral clustering: Analysis and an algorithm*, Advances in Neural Information Processing Systems 14 [Neural Information Processing Systems: Natural and Synthetic, NIPS 2001, December 3-8, 2001, Vancouver, British Columbia, Canada], 2001, pp. 849–856.

[NR09] Phong Q. Nguyen and Oded Regev, *Learning a parallelepiped: Cryptanalysis of GGH and NTRU signatures*, J. Cryptology **22** (2009), no. 2, 139–160.

[O'D16] Ryan O'Donnell, *SOS is not obviously automatizable, even approximately*, Electronic Colloquium on Computational Complexity (ECCC) **23** (2016), 141.

[Par00] Pablo A Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*, Ph.D. thesis, California Institute of Technology, 2000.

[PP02] A. Papachristodoulou and S. Prajna, *On the construction of lyapunov functions using the sum of squares decomposition*, Proceedings of the 41st IEEE Conference on Decision and Control, 2002., vol. 3, Dec 2002, pp. 3482–3487 vol.3.

[PS000] *Combinatorial approaches to finding subtle signals in dna sequences.*, vol. 8, 2000.

[PS82] Christos H. Papadimitriou and Kenneth Steiglitz, *Combinatorial optimization: Algorithms and complexity*, Prentice-Hall, 1982.

[QSW14] Qing Qu, Ju Sun, and John Wright, *Finding a sparse vector in a subspace: Linear sparsity using alternating directions*, NIPS, 2014, pp. 3401–3409.

[Rag08] Prasad Raghavendra, *Optimal algorithms and inapproximability results for every csp?*, STOC, ACM, 2008, pp. 245–254.

[RM14] Emile Richard and Andrea Montanari, *A statistical model for tensor PCA*, NIPS, 2014, pp. 2897–2905.

[RS09a] Prasad Raghavendra and David Steurer, *How to round any CSP*, FOCS, IEEE Computer Society, 2009, pp. 586–594.

[RS09b] ———, *Integrality gaps for strong SDP relaxations of UNIQUE GAMES*, 50th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2009, October 25-27, 2009, Atlanta, Georgia, USA, 2009, pp. 575–585.

[RS15] Prasad Raghavendra and Tselil Schramm, *Tight lower bounds for planted clique in the degree-4 sos program*, Preprint (2015).

[RT12] Prasad Raghavendra and Ning Tan, *Approximating csps with global cardinality constraints using SDP hierarchies*, SODA, SIAM, 2012, pp. 373–387.

[RW17] Prasad Raghavendra and Benjamin Weitz, *On the bit complexity of sum-of-squares proofs*, CoRR **abs/1702.05139** (2017).

[Sch08] Grant Schoenebeck, *Linear level lasserre lower bounds for certain k-CSPs*, FOCS, IEEE Computer Society, 2008, pp. 593–602.

[Sho87] N. Z. Shor, *A class of estimates for the global minimum of polynomial functions*, Kibernetika (Kiev) (1987), no. 6, 9–11, 133. MR 940145

[SM00] Jianbo Shi and Jitendra Malik, *Normalized cuts and image segmentation*, IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (2000), no. 8, 888–905.

[SR15] Rocco A. Servedio and Ronitt Rubinfeld (eds.), *Proceedings of the forty-seventh annual ACM on symposium on theory of computing, STOC 2015, portland, or, usa, june 14-17, 2015*, ACM, 2015.

[SS17] Tselil Schramm and David Steurer, *Fast and robust tensor decomposition with applications to dictionary learning*, Proceedings of The 30th Conference on Learning Theory, COLT 2017, Amsterdam, Netherlands, July 3-6, 2017, 2017.

[SWW12] Daniel A. Spielman, Huan Wang, and John Wright, *Exact recovery of sparsely-used dictionaries*, COLT, JMLR Proceedings, vol. 23, JMLR.org, 2012, pp. 37.1–37.18.

[Tao] Terence Tao, *Topics in random matrix theory*, Graduate studies in mathematics, American Mathematical Soc.

[Tro12] Joel A. Tropp, *User-friendly tail bounds for sums of random matrices*, Foundations of Computational Mathematics **12** (2012), no. 4, 389–434.

[TS14] Ryota Tomioka and Taiji Suzuki, *Spectral norm of random tensors*, arXiv preprint arXiv:1407.1870 (2014).

[Tul09] Madhur Tulsiani, *CSP gaps and reductions in the lasserre hierarchy*, Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009, 2009, pp. 303–312.

[Ver12] Roman Vershynin, *Introduction to the non-asymptotic analysis of random matrices*, Compressed sensing, Cambridge Univ. Press, Cambridge, 2012, pp. 210–268. MR 2963170

[Wei17] Benjamin Weitz, *Polynomial proof systems, effective derivations, and their applications in the sum-of-squares hierarchy*, Ph.D. thesis, UC Berkeley, 2017.