# On the Complexity of Unique Games and Graph Expansion

David Steurer

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

Recommended for Acceptance
By the Department of
Computer Science

Advisor: Sanjeev Arora

November 2010

# Abstract

Understanding the complexity of approximating basic optimization problems is one of the grand challenges of theoretical computer science. In recent years, a sequence of works established that Khot's Unique Games Conjecture, if true, would settle the approximability of many of these problems, making this conjecture a central open question of the field.

The results of this thesis shed new light on the plausibility of the Unique Games Conjecture, which asserts that a certain optimization problem, called Unique Games, is hard to approximate in a specific regime.

On the one hand, we give the first confirmation of this assertion for a restricted model of computation that captures the best known approximation algorithms. The results of this thesis also demonstrate an intimate connection between the Unique Games Conjecture and approximability of graph expansion. In particular, we show that the Unique Games Conjecture is true if the expansion of small sets in graphs is hard to approximate in a certain regime. This result gives the first sufficient condition for the truth of the conjecture based on the inapproximability of a natural combinatorial problem.

On the other hand, we develop efficient approximation algorithms for certain classes of Unique Games instances, demonstrating that several previously proposed variants of the Unique Games Conjecture are false. Finally, we develop a subexponential-time algorithm for Unique Games, showing that this problem is significantly easier to approximate than NP-hard problems like Max 3-Sat, Max 3-Lin, and Label Cover, which are unlikely to have subexponential-time algorithm achieving a non-trivial approximation guarantee. This algorithm also shows that the inapproximability results based on the Unique Games Conjecture do not rule out subexponential-time algorithms, opening the possibility for such algorithms for many basic optimization problems like Max Cut and Vertex Cover.

At the heart of our subexponential-time algorithm for Unique Games lies a novel algorithm for approximating the expansion of graphs across different scales, which might have applications beyond Unique Games, especially in the design of divide-and-conquer algorithms.

# Acknowledgments

I am greatly indebted to my advisor Sanjeev Arora, especially for his constant encouragement, his willingness to spent lots of time to discuss about research with me, and his efforts to improve my writing and presenting. His taste for problems, his commitment to research, and his scientific curiosity and creativity were a great source of inspiration. I couldn't have wished for a better advisor.

I am most thankful also to Boaz Barak, Prasad Raghavendra, and Nisheeth Vishnoi for their guidance, encouragement, and collaboration. They had considerable influence on the content of this thesis. I greatly enjoyed the many discussions with them about new problems and ideas.

Many thanks to Sanjeev Arora, Boaz Barak, Moses Charikar, Bernard Chazelle, and Russell Impagliazzo for serving on my thesis committee.

Over the years, I had the great pleasure to collaborate with many terrific researchers. I owe many thanks to my coauthors Sanjeev Arora, Boaz Barak, Markus Bläser, Constantinos Daskalakis, Benjamin Doerr, Venkatesan Guruswami, Moritz Hardt, Ishay Haviv, Thomas Holenstein, Russell Impagliazzo, Subhash Khot, Alexandra Kolla, Johannes Lengler, Yury Makarychev, William Matthews, Prasad Raghavendra, Anup Rao, Oded Regev, Peter Sanders, Prasad Tetali, Madhur Tulsiani, Nisheeth Vishnoi, Avi Wigderson, and Yuan Zhou. I greatly enjoyed working with them and I have learned a lot from them.

I thank all members of the theory groups in Princeton (at the University and at the Institute for Advanced Study) for creating such a great research environment. Thanks also to the members of the Center for Computational Intractability. The lively center meetings every month and the frequent workshops were great opportunities to learn and discuss about new results.

Many thanks to the Algo group at the Labortoire de Recherche en Informatique (LRI) for hosting me during the winter of 2008.

Many thanks also to Microsoft Research New England for hosting me as an intern in the summer of 2009. It was there that Prasad Raghavendra and I started working on graph expansion.

Before coming to Princeton, I spent three terrific years at Saarland University and the Max-Planck-Institute Informatik. Many thanks to all theorists there, especially my advisors Peter Sanders (now at Karlsruher Institute of Technology) and Kurt Mehlhorn.

Special thanks to the competent and helpful administrative staff at the Computer Science Department of Princeton University, especially Mitra Kelly and Melissa Lawson.

Heartfelt thanks to all my friends. They made the time in Princeton very enjoyable.

Finally, I would like to thank my family — my parents Florence and Manfred and my brother Hendrik — for their love and support.

I gratefully acknowledge the National Science Foundation for their financial support.

# Contents

# 1. Introduction

A central goal of theoretical computer science is to understand the complexity of basic computational tasks, that is, the minimum amount of computational resources (especially, time) needed to carry out such tasks.

We will focus on computational tasks arising from combinatorial optimization problems. Here, the goal is to optimize a given objective function over a finite set of solutions. Problems of this kind are ubiquitous in diverse fields, for example, computer networks, artificial intelligence, operation research, and the natural and social sciences. A simple, illustrating example is Max Cut:

**Max Cut:** Given a graph $G$, find a partition of the vertex set of $G$ into two sets so as to maximize the number of edges of $G$ crossing the partition.

With few exceptions, the complexity of finding optimal solutions to combinatorial optimization problems is understood. Either there is a (low-degree) polynomial-time algorithm that finds an optimal solution for the problem, or there is a (low-degree) polynomial-time reduction from 3-Sat to the problem that shows that finding an optimal solution is NP-hard (as hard as solving 3-Sat). Prominent examples of the first category are Minimum Spanning Tree and Maximum Matching. Problems of the first category typically reduce to some form of convex optimization. Unfortunately, most combinatorial optimization problems fall in the second category, ruling out polynomial-time algorithms — and even subexponential-time algorithms — under standard complexity assumptions. For example, finding an optimal solution for Max Cut is NP-hard [Kar72] (and the reduction from 3-Sat takes linear time). Since 3-Sat is unlikely to have a $\exp(o(n))$-time algorithm [IPZ01], this reduction also shows that finding an optimal solution for Max Cut is likely to require exponential time — $\exp(\Omega(n))$ — in the worst case.

Since finding optimal solution is intractable for many problems, it is natural to allow for approximation, raising the general question:

*What is the complexity of finding approximately optimal solutions?*

Understanding the approximability of optimization problems turns out to be challenging. Even after much progress in the last decades, the approximability of basic problems — in particular, Max Cut — remains open.

## 1.1. Approximation Complexity

There are several ways to measure the quality of an approximation. In many cases, the "approximation ratio" is an appropriate measure. (Later, we will come across problems that require less coarse measures of approximation.) We say that an algorithm has *approximation ratio* $\alpha$ if it always computes a solution whose value is within a factor $\alpha$ of the optimal value. (For us, there is no difference between approximation ratios $\alpha$ and $1/\alpha$. Most of the discussion in this chapter will be about maximization problems. In this case, approximation ratios are typically taken to be less than 1.) We allow algorithms to be randomized. In this case, we compare the expected value of the computed solution to the optimal value. (The expectation is only over the randomness of the algorithm and not over a particular input distribution.) All randomized algorithms considered in this thesis can be derandomized without (much) loss in efficiency or approximation.

For Max Cut, a trivial algorithm achieves approximation ratio $1/2$: A random partition cuts half of the edges in expectations, which is within a factor $1/2$ of the optimal value. For a long time, no better approximation algorithm for Max Cut was known, raising the question:

> *Is there an efficient algorithm for* Max Cut *with non-trivial approximation ratio?*

In a beautiful and influential result, Goemans and Williamson [GW95] answered this question positively, giving an algorithm for Max Cut with non-trivial approximation ratio $\alpha_{\mathrm{GW}} \approx 0.878$. This work introduced semidefinite programming (SDP) relaxations as a general technique for designing approximation algorithms. Today, this technique is the basis of the best known approximation algorithms for many basic optimization problems. (Often, no other techniques are known to achieve the same approximation as SDP relaxations.)

The next question regarding the approximability of Max Cut is whether the approximation ratio $\alpha_{\mathrm{GW}}$ is best possible or if it can be improved. In fact, until the early 1990s, it was not ruled out that for every constant $\alpha < 1$, Max Cut has an efficient algorithm with approximation ratio $\alpha$. (This question is not idle because such approximation schemes are known for other NP-hard optimization problems.)

*Is it* NP-*hard to obtain an approximation ratio* 0.9999 *for* Max Cut*?*

A breakthrough result, known as PCP Theorem[1] [ALM+98, AS98], answered this question positively (possibly for a constant closer to 1 than 0.9999). For a large class of problems, this result showed that the approximation ratios of efficient algorithms are bounded away from 1 (assuming P is different from NP).

These initial hardness of approximation results (for Max Cut say), based on the PCP Theorem, are not entirely satisfying. First, the hardness results leave open the possibility of efficient algorithms with significantly better approximation ratios than currently known. Second, the initial reductions showing these hardness results were quite inefficient (the blowup of the reduction was a high-degree polynomial, especially because of the use of Raz's Parallel Reduction Theorem [Raz98]).

Numerous works addressed the first issue. In an influential work [Hås01], Håstad showed optimal inapproximability results for several basic optimization problem, including Max 3-Lin, a generalization of Max Cut to 3-uniform hypergraph. Like for Max Cut, the trivial randomized algorithm for Max 3-Lin achieves approximation ratio $1/2$. Shockingly, this algorithm is optimal for Max 3-Lin (unlike for Max Cut). Håstad shows that no efficient algorithm for Max 3-Lin can achieve a non-trivial approximation ratio (better than $1/2$) unless P = NP. For many other problems, Håstad obtains improved, but not necessarily tight hardness results. For example, his results rule out efficient algorithms for Max Cut with approximation ratio less than $16/17 \approx 0.941$ (using a gadget of Trevisan et al. [TSSW00]). Closing the gap between known approximation algorithms and known hardness results for basic problems like Max Cut remains one of the outstanding open problems of the field. As discussed in the next section, Khot's Unique Games Conjecture offers a unified way to resolve this question for many problem classes. The hardness results based on the Unique Games Conjecture can be seen as a continuation of Håstad's work.

The second issue (inefficiency of initial hardness reductions) led to the development of efficient PCP constructions, culminating in the work of Moshkovitz and Raz [MR08]. Their result implies that Håstad's reductions (from 3-Sat)

[1] Here, PCP stands for probabilistically checkable proofs.

can be carried out in near-linear time, which shows that achieving a non-trivial approximation ratio for MAX 3-LIN has essentially the same complexity as solving 3-SAT (exactly). Concretely, if one assumes the Exponential Time Hypothesis [IPZ01], namely that the decision problem 3-SAT requires time $\exp(\Omega(n))$, then for every constant[2] $\varepsilon > 0$, achieving an approximation ratio $1/2 + \varepsilon$ for MAX 3-LIN requires time $\exp(n^{1-o(1)})$. One of the main results of this thesis is that reductions based on the Unique Games Conjecture are inherently inefficient, and therefore hardness results based on this conjecture cannot rule out subexponential approximation algorithms.

## 1.2. Unique Games Conjecture

UNIQUE GAMES[3] is the following constraint satisfaction problem, generalizing MAX CUT:

**UNIQUE GAMES:** Given a variable set $V$, an alphabet size $k \in \mathbb{N}$, and a list of constraints of the form $x_i - x_j = c \bmod k$ with $i, j \in V$ and $c \in \{1, \ldots, k\}$, find an assignment to the variables $(x_i)_{i \in V}$ so as to satisfy as many of the constraints as possible.

Khot [Kho02] conjectured that this problem is hard to approximate, in the sense that for $\varepsilon > 0$, it is NP-hard to distinguish between the case that at least $1 - \varepsilon$ of the constraints can be satisfied and the case that not more than $\varepsilon$ of the constraints can be satisfied. (An additional condition of the conjecture is that the alphabet size is allowed to grow only mildly with the input size, say $k = o(\log n)$.)

**Unique Games Conjecture:** For every constant $\varepsilon > 0$, the following task is NP-hard: Given a UNIQUE GAMES instance with $n$ variables and alphabet size $k = o(\log n)$, distinguish between the cases,

**YES:** some assignment satisfies at least $1 - \varepsilon$ of the constraints,

**NO:** no assignment satisfies at least $\varepsilon$ of the constraints.

[2] One can choose $\varepsilon$ even slightly subconstant.

[3] In fact, UNIQUE GAMES typically refers to a slightly more general problem. (See Section 2.3 for the more general definition.) For the purpose of the current discussion, it suffices to consider the special case of UNIQUE GAMES described above.

Starting with Khot's work, a sequence of results showed that the truth of this conjecture would imply improved and often optimal hardness of approximation results for many basic problems, including Max Cut. Perhaps most strikingly, Raghavendra [Rag08] showed that assuming the Unique Games Conjecture a simple SDP relaxation achieves an optimal approximation for every constraint satisfaction problem (this class of problems includes Max Cut, Max 3-Lin, and Unique Games). We refer to the survey [Kho10] for a more complete account of the known consequences of the Unique Games Conjecture.

Independent of the truth of the Unique Games Conjecture, the above results demonstrate that for many problems, Unique Games is a common barrier for improving current algorithms (in the sense that improving the current best algorithm for, say, Max Cut requires giving an improved algorithm for Unique Games). We call approximation problems of this kind UG-hard (as hard as approximating Unique Games).

The great number of strong UG-hardness results makes the Unique Games Conjecture a fascinating open question. Towards resolving this conjecture, it makes sense to ask more generally:

> *What is the complexity of approximating* Unique Games*?*

On the one hand, any lower bound on the complexity of Unique Games typically translates to corresponding lower bounds for all UG-hard problems. (This translation is essentially without quantitative loss, because UG-hardness reductions tend to be very efficient — the output of the reduction is only a constant factor larger than the input.) On the other hand, any upper bound on the complexity of Unique Games opens the possibility of similar upper bounds for other UG-hard problems.

The works in this thesis address several aspects of the above question. First, we consider the following aspect:

> *What properties make* Unique Games *instances easy?*

The goal is to develop efficient algorithms for Unique Games that provide good approximation guarantees for instances with certain properties. These properties should not be too restrictive so that more general constraint satisfaction problems (like Label Cover) remain hard even restricted to instances with these properties. Results of this kind give concrete explanations why the Unique Games Conjecture seems difficult to prove. Any approach for proving the conjecture has to avoid producing instances with these property. Another hope in pursuing this question is that algorithms developed for special classes of instances might be useful for general instances, leading to new upper bounds on the complexity of Unique Games and potentially to a refutation of the Unique Games Conjecture. (This hope turns out to be justified; at the end of this section, we discuss a surprising algorithm for Unique Games that uses algorithms developed for special classes of instances.)

In this thesis, we identify two properties that make Unique Games instances easy. (The properties are not overly restrictive in the sense that the more general Label Cover[4] problem remains hard when restricted to instances with these properties.)

In Chapter 3, we show that instances with expanding constraint graphs are easy. The *constraint graph* of a Unique Games instance is a graph on the variable set with an edge between any two variables that appear in the same constraint. Expansion measures how well a graph is connected. (In the next section, we discuss graph expansion in greater detail.) In particular, randomly generated graphs tend to be expanding. Therefore, this result demonstrates that Unique Games is easy on certain, commonly studied input distributions.

In Chapter 4, we show that instances of Unique Games arising from parallel repetition are easy. Parallel repetition is a commonly used reduction for amplifying approximation hardness (e.g., for Label Cover). This result shows that parallel repetition is unlikely to be useful for proving the Unique Games Conjecture. (Previous work [FKO07] suggested an approach for proving the Unique Games Conjecture based on parallel repetition.)

A major shortcoming of the current knowledge about the Unique Games Conjecture is that only few consequences of a refutation of the conjecture are known. Concretely, the following scenario is not ruled out: there exists an efficient algorithm for Unique Games and, at the same time, it is intractable to achieve better approximations for all other UG-hard problems. (In contrast, an efficient algorithm for 3-Sat implies efficient algorithms for all other NP-hard problems.) Hence, the goal is to identify "hard-looking" problems that could be solved efficiently if an efficient algorithm for Unique Games existed, raising the question:

*What "hard-looking" problems reduce to* Unique Games*?*

[4] Label Cover is a more general constraint satisfaction problem than Unique Games, where the constraints are allowed to be projections instead of permutations. The analog of the Unique Games Conjecture for Label Cover is known to be true (e.g., [MR08]).

In Chapter 6, we show that a certain graph expansion problem reduces to Unique Games and is therefore easier to approximate than Unique Games. (Graph expansion problems are well-studied and have a wide range of applications. See the next section for more discussion.) Our reduction shows that any efficient algorithm for Unique Games could be turned into an algorithm for this graph expansion problem. This algorithm would have a significantly better approximation guarantee than the current best algorithms for this problem. Before this work, no consequence of a refutation of the Unique Games Conjecture was known for a problem other than Unique Games itself.

Showing consequences of a refutation of the Unique Games Conjecture gives some evidence for the truth of the conjecture. (Especially, if the consequences are surprising.) Another way to generate evidence for the truth of the conjecture is to prove lower bounds for Unique Games in restricted models of computation, which leads us to the following question:

> *What is the complexity of approximating* Unique Games *in restricted models of computation?*

In the context of approximation problems, it is natural to study models of computations defined by hierarchies of relaxations (typically linear or semidefinite relaxations). Such hierarchies contain relaxations with gradually increasing complexity — from linear complexity to exponential complexity. Relaxations with higher complexity provide better approximations for the optimal value of optimization problems. (The relaxations with highest, i.e., exponential, complexity typically compute the optimal value exactly.)

Starting with the seminal work of Arora, Bollobás, and Lovász [ABL02], lower bounds on the complexity of many approximation problems were obtained in various hierarchies. We show the first super-polynomial lower bound for Unique Games in a hierarchy that captures the best known algorithms for all constraint satisfaction problems (see Chapter 8).

Our lower bounds translate to corresponding lower bounds for most classes of UG-hard problems (via the known UG-hardness reductions). Even for specific UG-hard problem like Max Cut such lower bounds were not known before.

Finally, we address the question about the complexity of approximating Unique Games directly. In Chapter 5, we give an algorithm for Unique Games with subexponential running time. The algorithm achieves an approximation guarantee that the Unique Games Conjecture asserts to be NP-hard to achieve. Concretely, the algorithm distinguishes between the case that at least $1-\varepsilon$ of the constraints are satisfiable and the case that not more than $\varepsilon$ of the constraints are satisfiable in time $\exp(k\, n^{\varepsilon^{1/3}})$, where $n$ is the number of variables and $k$ is alphabet size. We conclude that UG-hardness results cannot rule out subexponential algorithms (unlike hardness results based on Håstad's reductions). The algorithm also shows that any reduction from 3-Sat to Unique Games proving the Unique Games Conjecture must be inefficient (i.e., blow up the instance size by a high-degree polynomial ), assuming 3-Sat does not have subexponential algorithms. In this way, our algorithm rules out certain classes of reductions for proving the Unique Games Conjecture (in particular, reductions of the kind of Håstad's reductions).

At the heart of our subexponential algorithm for Unique Games lies a novel algorithm for approximating the expansion of graphs across different scales. Cheeger's influential work [Che70], relating eigenvalues and expansion, started a great body of research on approximations for graph expansion. The results of this thesis show an intimate connection between the complexity of Unique Games and the complexity of approximating graph expansion. We discuss more of this connection in next sections (especially §1.4).

## 1.3. Graph Expansion

Graphs are ubiquitous structures in computer science, mathematics, and the natural and social sciences. For example, they are useful for modeling various networks like the internet, genetic networks, and social networks.

A fundamental parameter of a graph is its *expansion*. The expansion[5] of a vertex set $S$, denoted $\Phi(S)$, measures how well the vertices in $S$ are connected to the rest of the graph. Formally, $\Phi(S)$ is defined as the number of edges leaving $S$ normalized by the number of the edges incident to the set. (An edge with both endpoints in the set $S$ counts twice in this normalization.) The *expansion* $\Phi$ of a graph is then defined as the minimum expansion of a vertex set with volume at most $1/2$. (In the context of expansion, the volume of a set is defined as the fraction of edges incident to the set.)

A natural computational question is to compute the expansion of a graph. We can formulate this question as optimization problem:

**Sparsest Cut:** Given a graph, find a vertex set $S$ with volume at most $1/2$ so

[5]In this thesis, expansion always refers to edge expansion. In the context of Markov chains, the quantity we call expansion is often called conductance.

as to minimize its expansion $\Phi(S)$.

A related problem is Balanced Separator, where the goal is to find a set with minimum expansion among all sets with volume between $\beta$ and $1/2$ for some balance parameter $\beta$. We refer to optimization problem of this kind (where we are required to find a set with minimum expansion subject to volume constraints) loosely as *graph expansion problems*.

Unfortunately (but not surprisingly), it is NP-hard to solve graph expansion problems exactly, which motivates the following question:

*How well can we approximate graph expansion problems?*

In an influential work, Cheeger [Che70] showed a relation, known as Cheeger's inequality, between the expansion of a graph and the second largest eigenvalue of its adjacency matrix. (Cheeger's proof is for manifolds instead of graphs. Dodziuk [Dod84] and independently Alon–Milman [AM85] and Alon [Alo86] showed the corresponding inequality for graphs.) Cheeger's inequality leads to an algorithm for Sparsest Cut with the following approximation guarantee: Given a graph with expansion $\varepsilon$, the algorithm finds a set with volume at most $1/2$ and expansion at most $2\sqrt{\varepsilon}$. We say that the algorithm achieves an $(\varepsilon, 2\sqrt{\varepsilon})$-*approximation* for every $\varepsilon > 0$.

Leighton and Rao [LR99] gave an algorithm for Sparsest Cut with approximation ratio $O(\log n)$ based on a linear programming relaxation. Arora, Rao, and Vazirani (ARV) [ARV04] improved this approximation ratio to $O(\sqrt{\log n})$ using a semidefinite programming relaxation. Both of these approximations are incomparable to Cheeger's inequality. For graphs with low expansion $\Phi \ll 1/\log n$, the ARV approximation guarantee is strongest. For expansion $\Phi \gg 1/\log n$, Cheeger's inequality gives the best known approximation.

A major open question in combinatorial optimization is whether the approximations of ARV and Cheeger's inequality are optimal in their respective regimes, or if better approximations for Sparsest Cut are achievable. Even assuming the Unique Games Conjecture, no strong hardness result for Sparsest Cut is known. (The best known result is that one cannot achieve approximation ratios arbitrarily close to 1 unless 3-Sat has subexponential algorithms [AMS07].) In this thesis, we introduce a hypothesis about the approximability of the expansion of small sets that allows us to show quantitatively tight hardness results for Sparsest Cut and other expansion problems (in a certain regime).

In the following discussion, we will focus[6] on the case that the size of the graph is very large compared to its expansion (for example, $\Phi \gg 1/\log\log n$). In this case, approximation guarantees depending on the instance size are meaningless. In particular, Cheeger's inequality provides the only known non-trivial approximation for Sparsest Cut in this regime.

## 1.4. Small-Set Expansion Hypothesis

Cheeger's inequality provides a good approximation for Sparsest Cut in the sense that it can distinguish between the case that the expansion of a graph is very close to 0, say $\Phi \leqslant 0.001$, and the case that the expansion is bounded away from 0, say $\Phi \geqslant 0.1$. For the following natural generalization of Sparsest Cut, no such approximation is known:

**Small-Set Expansion:** Given a graph and a scale $\delta$, find a vertex set $S$ with volume at most $\delta$ so as to minimize its expansion $\Phi(S)$.

Concretely, for every constant $\varepsilon > 0$, there exists $\delta > 0$ such that all known polynomial-time algorithms fail to distinguish between the following cases:

**Small-set expansion close to 0:** there exists a vertex set with volume $\delta$ and expansion at most $\varepsilon$,

**Small-set expansion close to 1:** every vertex set with volume $\delta$ has expansion more than $1-\varepsilon$.

The lack of efficient algorithms suggests that this problem may be computationally hard. To formalize this possibility, we introduce the following hypothesis:

**Small-Set Expansion (SSE) Hypothesis:** For every constant $\varepsilon > 0$, there exists a constant $\delta = \delta(\varepsilon) > 0$ such that the following task is NP-hard: Given a graph, distinguish between the cases,

**YES:** the graph contains a vertex set with volume $\delta$ and expansion at most $\varepsilon$,

[6] This restriction is of course not without loss of generality. We restrict ourselves to this regime because it turns out to be closely related to the Unique Games Conjecture and because we can prove tight bounds on the approximability of graph expansion problems in this regime. Understanding the approximability of Sparsest Cut in the regime $\Phi \ll 1/\log n$ remains wide open.

**NO:** the graph does not contain a vertex set with volume $\delta$ and expansion at most $1-\varepsilon$.

In Chapter 6, we show that this hypothesis implies the Unique Games Conjecture. (Combined with results of Chapter 7, we can in fact show that this hypothesis is equivalent to a variant of the Unique Games Conjecture, namely Hypothesis 6.3.) This result gives the first (and so far only) sufficient condition for the truth of the Unique Games Conjecture based on the inapproximability of a problem different than Unique Games.

Given that the SSE hypothesis is a stronger hardness assumption than the Unique Games Conjecture, we can ask if it has further consequences (beyond the consequences of the Unique Games Conjecture).

*What consequences would a confirmation of the SSE hypothesis have?*

In Chapter 7, we show that the SSE hypothesis implies quantitatively tight inapproximability results for many graph expansion problems, in particular Sparsest Cut and Balanced Separator. Concretely, our results imply that it is SSE-hard to beat Cheeger's inequality and achieve an $(\varepsilon, o(\sqrt{\varepsilon}))$-approximation for Sparsest Cut. (We say a problem is SSE-hard if an efficient algorithm for the problem implies that the SSE hypothesis is false.)

There is a strong parallel between this result and the known consequences of the Unique Games Conjecture. The Unique Games Conjecture asserts a qualitative[7] inapproximability for Unique Games, which is generalization of Max Cut. In turn, the conjecture implies tight quantitative inapproximability results for Max Cut and similar problems. Similarly, the SSE hypothesis asserts a qualitative inapproximability for Small-Set Expansion, which generalizes Sparsest Cut. In turn, the hypothesis implies tight quantitative inapproximability results for Sparsest Cut and other graph expansion problems. The value of results of this kind is that they unify the question of improving known algorithms for a class of problems to a question about the qualitative approximability of a single problem.

The significant consequences of a confirmation of the SSE hypothesis make it an interesting open question to prove or refute the hypothesis. As for Unique Games, it makes sense to ask the general question:

*What is the complexity of approximating* Small-Set Expansion*?*

[7] The Unique Games Conjecture is a qualitative hardness assumption in the sense that it does not specify any precise absolute constants or any concrete functional dependencies.

On the one hand, it turns out that the lower bounds for Unique Games in Chapter 8 extend to Small-Set Expansion, confirming the SSE hypothesis in a restricted model of computation (defined by a hierarchy of relaxations that captures the best known approximation algorithms). As for UG-hard problems, this lower bound translates to all SSE-hard problems. These results give the first quantitatively tight super-polynomial lower bounds for graph expansion problems in a hierarchy that captures the best known algorithms.

The results of Chapter 6 and Chapter 7 also show that the SSE hypothesis is *equivalent* to a variant of the Unique Games Conjecture. This variant asserts that Unique Games remains hard to approximate even restricted to instances with constraint graphs satisfying a relatively mild, qualitative expansion property. (In particular, the techniques of Chapter 3 — unique games with expanding constraints graphs — fail on instances with this mild expansion property.)

On the other hand, we show that Small-Set Expansion (like Unique Games) admits approximation algorithms with subexponential running time (see Chapter 5). This result demonstrates that SSE-hardness results do not rule out subexponential algorithms. A concrete possibility is that Sparsest Cut has a subexponential-time algorithm with constant approximation ratio. The subexponential-time algorithm for Small-Set Expansion inspired the subexponential algorithm for Unique Games discussed previously (in §1.2).

## 1.5. Organization of this Thesis

In the following, we outline the structure of this thesis and describe the contributions and results of the individual chapters.

### Part 1 — Algorithms

**Chapter 3: Unique Games with Expanding Constraint Graphs.** We study the approximability of Unique Games in terms of expansion properties of the underlying constraint graph. If the optimal solution satisfies a $1-\varepsilon$ fraction of the constraints and the constraint graph has spectral gap[8] $\lambda$, the algorithm developed in this chapter finds an assignment satisfying a $1-O(\varepsilon/\lambda)$ fraction of constraints. This results demonstrates that Unique Games instances

[8] Cheeger's inequality asserts that spectral gap of a graph is close to its expansion. Concretely, $\lambda/2 \leqslant \Phi \leqslant \sqrt{2\lambda}$. Sometimes $\lambda$ is called *spectral expansion*.

with spectral gap $\lambda \gg \varepsilon$ cannot be hard in the sense of the Unique Games Conjecture.

Our algorithm is based on a novel analysis of the standard SDP relaxation for Unique Games. In contrast to the usual local analysis, our analysis takes into account global properties of the solution to the SDP relaxation.

This chapter is based on joint work with Arora, Khot, Kolla, Tulsiani, and Vishnoi [AKK+08].

**Chapter 4: Parallel-Repeated Unique Games.** We study the approximability of parallel-repeated instances of Unique Games. Parallel repetition is commonly used for hardness amplification (e.g., for Label Cover, a generalization of Unique Games). We give an improved approximation algorithm for parallel-repeated instances of Unique Games. The guarantee of our algorithm matches the best known approximation for unrepeated instances [CMM06a]. In this sense, our result demonstrates that parallel repetition fails to amplify hardness (for Unique Games).

Our algorithm is based on a novel rounding for the standard SDP relaxation of Unique Games. In contrast to previous works, our rounding is able to exploit a tensor-product structure of the solution.

This chapter is based on joint work with Barak, Hardt, Haviv, Rao and Regev [BHH+08] and on the work [Ste10b].

**Chapter 5: Subexponential Approximation Algorithms.** We develop an algorithm for Unique Games with subexponential running time. The Unique Games Conjecture asserts that the approximation provided by our algorithm is NP-hard to achieve. A consequence of our algorithm is that known UG-hardness results do not rule subexponential algorithms.

The main ingredient of our algorithm is a general decomposition for graphs. We show that any graph can be efficiently partitioned into induced subgraphs, each with at most $n^{\beta}$ eigenvalues above $1-\eta$. Furthermore, the partition respects all but $O(\eta/\beta^3)^{1/2}$ of the original edges.

This chapter is based on joint work with Arora and Barak [ABS10].

## Part 2 — Reductions

**Chapter 6: Graph Expansion and the Unique Games Conjecture.** We propose a hypothesis (SSE hypothesis) about the approximability of the expansion of small sets in graphs. The reduction from Small-Set Expansion to Unique

Games developed in this chapter demonstrates that this hypothesis implies the Unique Games Conjecture. This result gives the first sufficient condition for the truth of the Unique Games Conjecture based on a problem different that Unique Games.

Furthermore, we show a partial converse of this implication: An refutation of the SSE hypothesis implies that a stronger variant of the Unique Games Conjecture is false. This stronger variant asserts that Unique Games remains hard to approximate (in the sense of the Unique Games Conjecture) even restricted to instances with constraint graphs satisfying a relatively mild, qualitative expansion property.

This chapter is based on joint work with Raghavendra [RS10].

**Chapter 7: Reductions between Expansion Problems.** Based on the reduction from Small-Set Expansion to Unique Games in Chapter 6, we develop reductions from Small-Set Expansion to other graph expansion problems, in particular Balanced Separator. Under the SSE hypothesis, these reductions imply quantitatively tight inapproximability results for these graph expansion problems.

Furthermore, we show that the SSE hypothesis implies a stronger variant of the Unique Games Conjecture, asserting that Unique Games restricted to instances with constraint graphs that satisfy a certain expansion property.

Compared to similar reductions from Unique Games, a key novelty of these reductions is that, on top of a composition with local gadgets, carefully placed random edges are added to the construction. The random edges are placed such that, on the one hand, the completeness of the reduction is not changed and, on the other hand, the expansion properties in the soundness case are improved.

This chapter is based on joint work with Raghavendra and Tulsiani [RST10b].

## Part 3 — Lower Bounds

**Chapter 8: Limits of Semidefinite Programming.** We show super-polynomial lower bounds for Unique Games in certain hierarchies of SDP relaxations (which capture the best known approximation algorithms for constraint satisfaction problems). Previous lower bounds considered only relaxations of fixed polynomial complexity or hierarchies that do not capture the best known approximation algorithms. Via known reductions, our lower bounds extend to all known UG-hard problems. (Our construction and analysis also yields super-polynomial lower bounds for Small-Set Expansion. Via the reductions in Chapter 7, these lower bounds extend to other SSE-hard graph expansion problems, like Balanced Separator.)

A simple, but important ingredient of our analysis is a robustness theorem for the hierarchies we consider. This robustness theorem asserts that an "approximate solution" to a relaxation in these hierarchies can be turned to a proper solution without changing the objective value by much.

This chapter is based on joint work with Raghavendra [RS09a].

# 2. Background

This chapter serves as a source of references for the following chapters. (However, its content is not assumed in the following.) Interested readers can skim this chapter for the first reading.

## 2.1. Optimization and Approximation

For simplicity, we restrict the discussion to maximization problems. For minimization problems, we use the same notations, but change signs in the definitions appropriately.

**Optimization Problems.** A *generic optimization instance* $\mathfrak{I}$ consists of a set $\Omega$ of *feasible solutions* and an *objective function* $f : \Omega \to \mathbb{R}$. (The symbol $\mathfrak{I}$ is a capital "i" in Fraktur font. It is intended to be pronounced as "i".) If the meaning is clear from the context, we identify the instance $\mathfrak{I}$ with its objective function and write $\mathfrak{I}(x)$ to denote the objective value of $x$. The *(optimal) value* $\operatorname{opt}(\mathfrak{I})$ is defined as maximum objective value of a feasible solution,

$$\operatorname{opt}(\mathfrak{I}) \stackrel{\text{def}}{=} \max_{x \in \Omega} \mathfrak{I}(x).$$

A solution $x^*$ is *optimal* for instance $\mathfrak{I}$ if it achieves the optimal value $\mathfrak{I}(x^*) = \operatorname{opt}(\mathfrak{I})$.

An *optimization problem* $\Pi$ is formally specified as a set of optimization instances. We say that an optimization problem $\Pi$ is *combinatorial* if every instance $\mathfrak{I} \in \Pi$ has only a finite number of feasible solutions. For any optimization problem $\Pi$, we associate the following computational problem Exact-$\Pi$,

**Exact-$\Pi$:** Given an instance $\mathfrak{I} \in \Pi$, find an optimal solution $x^*$ of $\mathfrak{I}$.

The computational problem Exact-$\Pi$ makes only sense if we also fix an succinct encoding of the instances of $\Pi$. For example, the instances of Max Cut are naturally encoded as graphs.

Typically, we use the same notation for the collection of instances and for the computational problem. For example, Max Cut stands both for the set of Max Cut instances and for the computational problem of finding the maximum cut in a graph.

**Approximation Algorithms.** We say that an algorithm $\mathfrak{A}$ achieves a $(c,s)$-*approximation* for an optimization problem $\Pi$ if given an instance $\mathfrak{I} \in \Pi$ with $\operatorname{opt}(\mathfrak{I}) \geqslant c$, the algorithm finds a feasible solution $x = \mathfrak{A}(\mathfrak{I})$ for the instance $\mathfrak{I}$ with objective value $\mathfrak{I}(x) \geqslant s$.

To understand the approximation guarantees of an algorithm $\mathfrak{A}$ for a problem $\Pi$, we want to find for every value $c \in \mathbb{R}$, the largest value $s$ such that the algorithm $\mathfrak{A}$ achieves a $(c,s)$-approximation for $\Pi$.

For example, the Goemans–Williamson algorithm achieves a $(1-\varepsilon,\ 1-O(\sqrt{\varepsilon}))$-approximation for Max Cut for all $\varepsilon > 0$.

If the objective functions of all instances of a problem $\Pi$ are nonnegative, we define the *approximation ratio* $\alpha$ of an algorithm $\mathfrak{A}$ for $\Pi$ as the minimum ratio between the value of the solution found by $\mathfrak{A}$ and the optimal value,

$$\alpha \stackrel{\text{def}}{=} \min_{\substack{\mathfrak{I} \in \Pi \\ x = \mathfrak{A}(\mathfrak{I})}} \frac{\mathfrak{I}(x)}{\operatorname{opt}(\mathfrak{I})} .$$

**Gap-Promise Problems.** Hardness of approximation results are typically based on gap-preserving reductions between optimization problems. In this context, it is convenient to associate the following kind of promise problems with optimization problems. For an optimization problem $\Pi$ and constants $c,s \in \mathbb{R}$ with $c \geqslant s$, define the following promise problem,

**$(c,s)$-Gap-$\Pi$:** Given an instance $\mathfrak{I} \in \Pi$, distinguish between the cases

- **Yes** $\operatorname{opt}(\mathfrak{I}) \geqslant c$,
- **No** $\operatorname{opt}(\mathfrak{I}) < s$.

(In some situations, it makes sense to take $c$ and $s$ as functions of instance parameters, e.g., the size of the instance. In this thesis, we focus on the case that $c$ and $s$ are constants, depending only on the problem $\Pi$.)

## 2.2. Graphs, Expansion, and Eigenvalues

In this work, we consider undirected, weighted graphs and we allow self-loops. In this case, we can represent a *graph* $G$ with vertex set $V$ as a symmetric distribution over pairs $ij$ with $i, j \in V$. The *edges* of $G$ are the pairs $ij$ in the support of this distribution. (Here, symmetry means that for any two vertices $i, j \in V$, the pair $ij$ has the same probability as the pair $ji$.) We will assume finite vertex sets.

We write $ij \sim G$ to denote a random edge $ij$ sampled from $G$. For a vertex $i \in V$, we write $j \sim G(i)$ to denote a random *neighbor* of $i$ in $G$. (A random neighbor of $i$ is obtained by sampling a random edge of $G$ conditioned on the event that the first endpoint of the edge is $i$ and outputting the second endpoint of that edge.)

**Expansion and Expansion Profile.** For two vertex sets $S, T \subseteq V$, we define $G(S, T)$ as the fraction of edges going from $S$ to $T$,

$$G(S,T) \stackrel{\text{def}}{=} \mathop{\mathbb{P}}_{ij \sim G} \{i \in S,\ j \in T\}\,.$$

For a vertex set $S \subseteq V$, we define its *(edge) boundary* $\partial_G(S)$ as the fraction of edges leaving $S$ (going from $S$ to the complement of $S$),

$$\partial_G(S) \stackrel{\text{def}}{=} G(S, V \setminus S)\,.$$

We define its *volume* $\mu_G(S)$ as the fraction of edges going out a vertex in $S$,

$$\mu_G(S) \stackrel{\text{def}}{=} G(S, V)\,.$$

The *expansion* $\Phi_G(S)$ is the ratio of these quantities

$$\Phi_G(S) \stackrel{\text{def}}{=} \frac{\partial_G(S)}{\mu_G(S)}\,.$$

(According to the above expression, the expansion of the empty set is undefined, since both the boundary and the volume is naught. Since we are interested in minimizing the expansion, it is convenient to define the expansion of the empty set to be infinite.)

The quantities defined above have the following probabilistic interpretations,

$$\partial_G(S) = \mathop{\mathbb{P}}_{ij \sim G} \{i \in S,\ j \notin S\}\,,$$
$$\mu_G(S) = \mathop{\mathbb{P}}_{ij \sim G} \{i \in S\}\,,$$
$$\Phi_G(S) = \mathop{\mathbb{P}}_{ij \sim G} \{j \notin S \mid i \in S\}\,.$$

Another probabilistic interpretation of the expansion of the set $S$ is the fraction of neighbors outside of $S$ for a typical vertex in $S$,

$$\Phi_G(S) = \mathop{\mathbb{E}}_{i \sim G} \left[ \mathop{\mathbb{P}}_{j \sim G(i)} \{j \notin S\} \,\middle|\, i \in S \right].$$

Here, $i \sim G$ denotes a random vertex of $G$, obtained by sampling a random edge of $G$ and taking its first endpoint. If the graph $G$ is clear from the context, we usually write $i \sim V$ or $i \sim \mu$ to denote a random vertex of $G$.

The *expansion* $\Phi_G$ of the graph $G$ is the minimum expansion of a set with volume at most $1/2$,

$$\Phi_G \stackrel{\text{def}}{=} \min_{\substack{S \subseteq V \\ \mu_G(S) \leqslant 1/2}} \Phi_G(S).$$

More generally, for $\delta \in [0,1]$, the *expansion at volume* $\delta$, denoted $\Phi_G(\delta)$, is the minimum expansion of a set with volume at most $\delta$,

$$\Phi_G(\delta) \stackrel{\text{def}}{=} \min_{\substack{S \subseteq V \\ \mu_G(S) \leqslant \delta}} \Phi_G(S).$$

The curve $\delta \mapsto \Phi_G(\delta)$ is called the *expansion profile* of $G$. If the graph $G$ is clear from the context, we typically omit the subscript $G$ for the above notations.

Note that for every $\delta > 0$, the expansion profile satisfies $\Phi(\delta) \cdot \delta' = \Phi(1 - \delta) \cdot (1 - \delta')$ for some $\delta' \leqslant \delta$. In particular, $\Phi(1-\delta) \leqslant \delta/(1-\delta)$. Therefore, we typically take $\delta \in [0, 1/2]$.

**Functions on Graphs, Laplacians, and Cheeger Bounds.** For a graph $G$ with vertex set $V$, we write $L_2(V)$ to denote the space $\{f : V \to \mathbb{R}\}$ equipped with natural inner product for $f, g \in L_2(V)$,

$$\langle f, g \rangle \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{i \sim V} f_i g_i\,.$$

(We often write $f_i$ to denote the value $f(i)$ of the function $f$ on vertex $i$.) This inner product induces the norm $\|f\| := \langle f, f\rangle^{1/2}$. We will also be interested in the norms $\|f\|_\infty = \max_{i\in V} f_i$ and $\|f\|_1 = \mathbb{E}_{i\sim V}|f_i|$.

We identify the graph $G$ with the following linear (Markov) operator on $V$,

$$Gf(i) \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{j\sim G(i)} f(j).$$

The matrix corresponding to this operator is the (weighted) adjacency matrix of $G$ normalized so that every row sums to 1. The operator $G$ is self-adjoint with respect to the inner product on $L_2(V)$ (because $\langle f, Gg\rangle = \mathbb{E}_{ij\sim G} f_i g_j$). Since $G$ is self-adjoint, its eigenvalues are real and its eigenfunctions are an orthogonal basis of $L_2(V)$.

For a vertex set $S \subseteq V$, let $\mathbb{1}_S \in L_2(V)$ be the $\{0,1\}$-indicator function of $S$. (For the all-ones functions $\mathbb{1}_V$, we typically drop the subscript.) Since $G\mathbb{1} = \mathbb{1}$, the all-ones functions is an eigenfunction of $G$ with eigenvalue 1. The *Laplacian* $L_G$ of the graph $G$ is the following linear operator on $L_2(V)$,

$$L_G \stackrel{\text{def}}{=} I - G,$$

where $I$ is the identity operator on $L_2(V)$. The Laplacian $L_G$ corresponds to the following quadratic form

$$\langle f, Lf\rangle = \mathop{\mathbb{E}}_{ij\sim G} \tfrac{1}{2}\left(f_i - f_j\right)^2.$$

(We drop the subscript for the Laplacian if the graph is clear from the context.) The following identities hold for all vertex sets $S, T \subseteq V$,

$$G(S,T) = \langle \mathbb{1}_S, G\mathbb{1}_T\rangle,$$
$$\mu(S) = \|\mathbb{1}_S\|^2, \qquad \partial(S) = \langle \mathbb{1}_S, L\mathbb{1}_S\rangle,$$
$$\Phi(S) = \frac{\langle \mathbb{1}_S, L\mathbb{1}_S\rangle}{\|\mathbb{1}_S\|^2}.$$

**Lemma 2.1** (Cheeger's Inequality, [Che70]). *Let $1 = \lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$ be the eigenvalues of $G$. Suppose the gap between the largest and second largest eigenvalue is $\varepsilon = \lambda_1 - \lambda_2$ (the* spectral gap *of $G$). Then,*

$$\varepsilon/2 \leqslant \Phi_G \leqslant \sqrt{2\varepsilon}.$$

The proof of Cheeger's inequality combined with a truncation argument yields the following lemma. For a self-contained proof, see the next subsection §2.2.1.

**Lemma 2.2** (Local Cheeger Bound). *For every function $f \in L_2(V)$, there exists a level set $S \subseteq V$ of the function $f^2$ with volume $\mu(S) \leqslant \delta$ and expansion*

$$\Phi(S) \leqslant \frac{\sqrt{1 - \langle f, Gf\rangle^2/\|f\|^4}}{1 - \|f\|_1^2/\delta\|f\|^2} \leqslant \frac{\sqrt{2\langle f, Lf\rangle/\|f\|^2}}{1 - \|f\|_1^2/\delta\|f\|^2}.$$

The above Cheeger bound is an important ingredient for the results of Chapter 5 (Subexponential Approximation Algorithms). Local variants of Cheeger's inequality (similar to the above bound) appear in several works, for example, [DI98, GMT06, RST10a].

**Spectral Profile.** The *spectral profile* of the graph $G$ refers the curve $\delta \mapsto \Lambda_G(\delta)$, where

$$\Lambda_G(\delta) \stackrel{\text{def}}{=} \min_{\substack{f \in L_2(V) \\ \|f\|_1^2 \leqslant \delta \|f\|^2}} \frac{\langle f, Lf\rangle}{\|f\|^2}.$$

The spectral profile is a lower bound on the expansion profile, $\Lambda_G(\delta) \leqslant \Phi_G(\delta)$, because $f = \mathbb{1}_S$ satisfies $\|f\|_1^2 = \mu(S)\|f\|^2$ for every vertex set $S \subseteq V$. Lemma 2.2 implies an upper bound of the expansion profile in terms of spectral profile. For all $\delta, \gamma \in [0,1)$, the following relations hold

$$\Lambda(\delta) \leqslant \Phi(\delta) \leqslant \sqrt{2\Lambda\big((1-\gamma)\delta\big)/\gamma}.$$

### 2.2.1. Proof of Local Cheeger Bound

In this subsection, we prove the following local Cheeger bound.

**Lemma** (Restatement of Lemma 2.2). *For every function $f \in L_2(V)$, there exists a level set $S \subseteq V$ of the function $f^2$ with volume $\mu(S) \leqslant \delta$ and expansion*

$$\Phi(S) \leqslant \frac{\sqrt{1 - \langle f, Gf\rangle^2/\|f\|^4}}{1 - \|f\|_1^2/\delta\|f\|^2} \leqslant \frac{\sqrt{2\langle f, Lf\rangle/\|f\|^2}}{1 - \|f\|_1^2/\delta\|f\|^2}.$$

Let $f \in L_2(V)$. Suppose $f^2 \leqslant 1$. Consider the following distribution over vertex subsets $S \subseteq V$:

1. Sample $t \in [0,1]$ uniformly at random.
2. Output the set $S = \{i \in V \mid f_i^2 > t\}$.

Note that every set $S$ in the support of this distribution is a level set of the function $f^2$. In the following lemmas, we establish simple properties of this distribution.

**Claim 2.3.** *The expected volume of $S$ satisfies $\mathbb{E}_S \mu(S) = \|f\|^2$.*

*Proof.* We calculate the expected volume as follows

$$\mathop{\mathbb{E}}_S \mu(S) = \mathop{\mathbb{E}}_{i\sim\mu} \mathop{\mathbb{P}}_{t\in[0,1]} \left\{ f_i^2 > t \right\} = \|f\|^2 . \qquad \square$$

**Claim 2.4.** *The second moment of $\mu(S)$ is at most $\mathbb{E}_S \mu(S)^2 \leqslant \|f\|_1^2$.*

*Proof.* We bound the expectation of $\mu(S)^2$ as follows

$$\mathop{\mathbb{E}}_S \mu(S)^2 = \mathop{\mathbb{E}}_{i,j\sim\mu} \mathop{\mathbb{P}}_t \left\{ \min\{f_i^2, f_j^2\} > t \right\} = \mathop{\mathbb{E}}_{i,j\sim\mu} \min\{f_i^2, f_j^2\} \leqslant \mathop{\mathbb{E}}_{i,j\sim\mu} f_i f_j = \|f\|_1^2 . \qquad \square$$

**Claim 2.5.** *Sets with volume larger than $\delta$ contribute to the expected volume at most $\mathbb{E}_S \mu(S) \mathbb{1}_{\mu(S)>\delta} \leqslant \mathbb{E}_S \mu(S)^2/\delta$.*

*Proof.* Immediate because $\mu(S)\mathbb{1}_{\mu(S)>\delta} \leqslant \mu(S)^2/\delta$ holds pointwise. $\square$

**Claim 2.6.** *The expected boundary of $S$ is bounded by*

$$\mathop{\mathbb{E}}_S G(S, V \setminus S) \leqslant \|f\|^2 \sqrt{1 - \langle f, Gf\rangle^2 / \|f\|^4}.$$

*Proof.* We calculate the expected boundary of $S$ and apply the Cauchy–Schwarz inequality,

$$\begin{aligned}
\mathop{\mathbb{E}}_S G(S, V \setminus S) &= \mathop{\mathbb{E}}_{ij\sim G} \mathop{\mathbb{P}}_S \{i \in S \wedge j \notin S\} = \mathop{\mathbb{E}}_{ij\sim G} \mathop{\mathbb{P}}_t \left\{ f_i^2 > t \geqslant f_j^2 \right\} \\
&= \mathop{\mathbb{E}}_{ij\sim G} \max\left\{ f_i^2 - f_j^2, 0 \right\} = \tfrac{1}{2} \mathop{\mathbb{E}}_{ij\sim G} \left| f_i^2 - f_j^2 \right| = \tfrac{1}{2} \mathop{\mathbb{E}}_{ij\sim G} \left| f_i - f_j \right| \cdot \left| f_i + f_j \right| \\
&\leqslant \left( \mathop{\mathbb{E}}_{ij\sim G} \tfrac{1}{2}(f_i - f_j)^2 \cdot \mathop{\mathbb{E}}_{ij\sim G} \tfrac{1}{2}(f_i + f_j)^2 \right)^{1/2} \qquad \text{(using Cauchy–Schwarz)} \\
&= \langle f, (I - G) f\rangle^{1/2} \langle f, (I + G) f\rangle^{1/2} = \sqrt{\|f\|^4 - \langle f, Gf\rangle^2} . \qquad \square
\end{aligned}$$

We combine the previous claims to complete the proof of Lemma 2.2. Let $S^*$ be the level set of $f^2$ with volume at most $\delta$ and minimum expansion. Then,

$$\begin{aligned}\Phi(S^*) &\leqslant \frac{\mathbb{E}_S\, G(S, V\setminus S)\mathbb{1}_{\mu(S)\leqslant\delta}}{\mathbb{E}_S\, \mu(S)\mathbb{1}_{\mu(S)\leqslant\delta}} \\ &\leqslant \frac{\mathbb{E}_S\, G(S,V\setminus S)}{\mathbb{E}_S\, \mu(S) - \mathbb{E}_S\, \mu(S)^2/\delta} \quad \text{(using Claim 2.5)} \\ &\leqslant \frac{\|f\|^2\sqrt{1-\langle f, Gf\rangle^2/\|f\|^4}}{\|f\|^2 - \|f\|_1^2/\delta} \quad \text{(using Claim 2.3, Claim 2.4, and Claim 2.6).}\end{aligned}$$

Therefore, the set $S^*$ satisfies the conclusion of the local Cheeger bound (Lemma 2.2).

## 2.3. Unique Games and Semidefinite Relaxation

**Unique Games.** A *unique game* $\mathcal{U}$ with vertex set $V$ and alphabet $\Sigma$ is a distribution over *constraints* $(u,v,\pi) \in V\times V\times \mathbb{S}_\Sigma$, where $\mathbb{S}_\Sigma$ denotes the set of permutations of $\Sigma$. An *assignment* $x \in \Sigma^V$ *satisfies* a constraint $(u,v,\pi)$ if $x_v = \pi(x_u)$, that is, the permutation $\pi$ maps labels for $u$ to labels for $v$. The *value* $\mathcal{U}(x)$ of an assignment $x$ for $\mathcal{U}$ is the fraction of the constraints of $\mathcal{U}$ satisfied by the assignment $x$, i.e.,

$$\mathcal{U}(x) \stackrel{\text{def}}{=} \mathop{\mathbb{P}}_{(u,v,\pi)\sim\mathcal{U}} \{\pi(x_u) = x_v\}\,.$$

(Here, $(u,v,\pi)\sim\mathcal{U}$ denotes a constraint sampled from $\mathcal{U}$.) The *(optimal) value* $\mathrm{opt}(\mathcal{U})$ is defined as the maximum value of an assignment, i.e.,

$$\mathrm{opt}(\mathcal{U}) = \max_{x\in\Sigma^V} \mathcal{U}(x)\,.$$

We will assume that the distribution over constraints is symmetric, in the sense that a constraint $(u,v,\pi)$ has the same probability as the constraint $(v,u,\pi^{-1})$.

The value of an assignment $x$ corresponds to the success probability of the following probabilistic verification procedure:

1. Sample a random constraint $(u,v,\pi)\sim\mathcal{U}$.

2. Verify that $x_v = \pi(x_u)$.

**Constraint and Label-Extended Graph.** We can associate two graphs with a unique game. Let $\mathfrak{U}$ be a unique game with vertex set $V$ and alphabet $\Sigma$. The *constraint graph* $G(\mathfrak{U})$ is a graph with vertex set $V$. Its edge distribution is obtained as follows:

1. Sample a random constraint $(u,v,\pi) \sim \mathfrak{U}$.

2. Output the edge $uv$.

The *label-extended graph* $\hat{G}(\mathfrak{U})$ is a graph with vertex set $V \times \Sigma$. Its edge distribution is obtained in the following way:

1. Sample a random constraint $(u,v,\pi) \sim \mathfrak{U}$.

2. Sample a random label $i \in \Sigma$.

3. Output an edge between $(u,i)$ and $(v,\pi(i))$.

Sometimes it will be convenient to denote a vertex $(u,i)$ in the label-extended graph by $u_i$.

An assignment $x \in \Sigma^V$ naturally corresponds to a set $S \subseteq V \times \Sigma$ with cardinality $|S| = |V|$ (and therefore, volume $\mu(S) = 1/|\Sigma|$). The value of the assignment $x$ for the unique game $\mathfrak{U}$ corresponds exactly to the expansion of the corresponding set $S$ in the label-extended graph $\hat{G}(\mathfrak{U})$,

$$\mathfrak{U}(x) = 1 - \Phi(S).$$

This correspondence between expansion and the value of an assignment is the basis of the connection between Unique Games and Small-Set Expansion discussed in this thesis.

**Partial Unique Games.** It is sometimes convenient to consider partial assignments for unique games. Let $\mathfrak{U}$ be a unique game with vertex set $V$ and alphabet $\Sigma$. An assignment $x \in (\Sigma \cup \{\perp\})^V$ is *$\alpha$-partial* if at least an $\alpha$ fraction of the vertices are labeled (with symbols from $\Sigma$), i.e.,

$$\mathop{\mathbb{P}}_{(u,v,\pi)\sim\mathfrak{U}} \{x_u \neq \perp\} \geqslant \alpha\,.$$

A partial assignment $x$ satisfies a constraint $(u,v,\pi)$ if both vertices $u$ and $v$ are labeled and their labels satisfy the constraint $x_v = \pi(x_u)$. For conciseness, we write $x_v = \pi(x_u) \in \Sigma$ to denote the event that the partial assignment $x$ satisfies the constraint $(u,v,\pi)$. The *value* $\mathcal{U}(x)$ of a partial assignment $x$ is the fraction of constraints satisfied by $x$,

$$\mathcal{U}(x) \stackrel{\text{def}}{=} \mathop{\mathbb{P}}_{(u,v,\pi)\sim\mathcal{U}} \{x_v = \pi(x_u) \in \Sigma\} .$$

The *$\alpha$-partial value* $\mathrm{opt}_\alpha(\mathcal{U})$ is the maximum value of a $\alpha$-partial assignment normalized by the fraction of labeled vertices,

$$\mathrm{opt}_\alpha(\mathcal{U}) \stackrel{\text{def}}{=} \max\left\{ \frac{1}{\mathbb{P}_{u\sim V}\{x_u \neq \bot\}} \mathcal{U}(x) \;\middle|\; x \in (\Sigma \cup \{\bot\})^V,\ \mathop{\mathbb{P}}_{(u,v,\pi)\sim\mathcal{U}} \{x_u \neq \bot\} \geqslant \alpha \right\} . \tag{2.1}$$

Note that $\mathrm{opt}(\mathcal{U}) = \mathrm{opt}_1(\mathcal{U})$ and $\mathrm{opt}_\alpha(\mathcal{U}) \leqslant \mathrm{opt}_{\alpha'}(\mathcal{U})$ whenever $\alpha \geqslant \alpha'$.

**SDP Relaxation of Unique Games.** With every unique game $\mathcal{U}$, we associate a *semidefinite program* $\mathrm{SDP}(\mathcal{U})$ of polynomial size. (Semidefinite programs are instances of Semidefinite Programming, an efficiently solvable[1] (convex) optimization problem.)

**Relaxation 2.7** (Semidefinite Relaxation for Unique Games).
Given a unique game $\mathcal{U}$ with vertex $V$ and alphabet $\Sigma$, find a collection of vectors $\{\boldsymbol{u}_i\}_{u\in V, i\in\Sigma} \subseteq \mathbb{R}^d$ for some $d \in \mathbb{N}$ so as to

$$\mathrm{SDP}(\mathcal{U}): \quad \text{maximize} \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathcal{U}} \langle \boldsymbol{u}_i, \boldsymbol{v}_{\pi(i)} \rangle \tag{2.2}$$

$$\text{subject to} \quad \sum_{i\in\Sigma} \|\boldsymbol{u}_i\|^2 = 1 \quad (u \in V), \tag{2.3}$$

$$\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = 0 \quad (u \in V, i \neq j \in \Sigma). \tag{2.4}$$

We let $\mathrm{sdp}(\mathcal{U})$ denote the optimal value of the relaxation $\mathrm{SDP}(\mathcal{U})$.

**Γ-Max 2-Lin.** An interesting special case of Unique Games is the Γ-Max 2-Lin problem. (Indeed, in many situations it suffices to consider the simpler Γ-Max 2-Lin problem.) For a finite group $\Gamma$, a unique game $\mathcal{U}$ with alphabet $\Sigma$ is a Γ-Max 2-Lin instance, if we can identify the alphabet $\Sigma$ with the group $\Gamma$ such

[1] More precisely, for every desired accuracy $\varepsilon > 0$, one can efficiently compute a solution with objective value within $\varepsilon$ of the optimal value (even if $\varepsilon$ is exponentially small in the instance size.) Since we are interested in approximation algorithms for combinatorial problems, we can ignore this (arbitrarily small) error and assume that we can find an optimal solution efficiently.

that for every constraint $(u,v,\pi)$ in $\mathfrak{U}$, there exists a group element $c \in \Gamma$ such that $a = \pi(b)$ if and only if $ab^{-1} = c$. For example, in the case that $\Gamma = \mathbb{Z}_k$ is the cyclic group of order $k$, all constraints of $\mathfrak{U}$ have the form $x_u - x_v = c \bmod k$ for some $c \in \mathbb{Z}_k$.

**Parallel Repetition.** Let $\mathfrak{U}$ be a unique game with vertex set $V$ and alphabet $\Sigma$. For $\ell \in \mathbb{N}$, the $\ell$-fold *(parallel) repetition* of $\mathfrak{U}$, denoted $\mathfrak{U}^{\otimes \ell}$, is a unique game with vertex set $V^\ell$ and alphabet $\Sigma^\ell$. (We sometimes write $\mathfrak{U}^\ell$ instead of $\mathfrak{U}^{\otimes \ell}$.) The unique game $\mathfrak{U}^{\otimes \ell}$ corresponds to the following probabilistic verifier for an assignment $X\colon V^\ell \to \Sigma^\ell$:

1. Sample $\ell$ constraints $(u_1,v_1,\pi_1),\ldots,(u_\ell,v_\ell,\pi_\ell) \sim \mathfrak{U}$.
2. Verify that $X(v_r) = \pi_r(X(u_r))$ for all $r \in \{1,\ldots,\ell\}$.

Note that the constraint distribution of $\mathfrak{U}^{\otimes \ell}$ corresponds to the $\ell$-fold product of the constraint distribution of $\mathfrak{U}$.

It is easy to verify that the optimal value of the repeated unique game $\mathfrak{U}^{\otimes \ell}$ is at least $\operatorname{opt}(\mathfrak{U})^\ell$. This lower bound is not always tight. Indeed, Raz [Raz08] showed that for every small $\varepsilon > 0$, there exists a unique games $\mathfrak{U}$ with $\operatorname{opt}(\mathfrak{U}) \leqslant 1-\varepsilon$, but $\operatorname{opt}(\mathfrak{U}^{\otimes \ell}) \geqslant 1 - O(\sqrt{\ell}\,\varepsilon)$.

If the constraint graph of the unique game $\mathfrak{U}$ is bipartite, then an approach of Feige and Lovász [FL92] shows that the semidefinite value of the repeated game is just a function of the semidefinite value of the unrepeated game.

**Theorem 2.8** ([FL92, MS07, KRT08]). *For every unique game $\mathfrak{U}$ with bipartite constraint graph and every $\ell \in \mathbb{N}$,*

$$\operatorname{sdp}(\mathfrak{U}^{\otimes \ell}) = \operatorname{sdp}(\mathfrak{U})^\ell .$$

We note that in many situations it is possible to reduce general unique games to ones with bipartite constraints.

**Squares of Unique Games.** In Part II (Reductions), we often consider "squares" of unique games (either implicitly or explicitly). For a unique game $\mathfrak{U}$ with vertex set $V$ and alphabet $\Sigma$, its square $\mathfrak{U}^2$ is a unique game with the same vertex set and alphabet. It corresponds to the following probabilistic verifier for an assignment $x \in \Sigma^V$,

1. Sample a random vertex $u \sim V$. (We sample $u$ according to the marginal distribution of the first vertex in a random constraint of $\mathfrak{U}$.)

2. Sample two random constraints $(u, v, \pi), (u, v', \pi') \sim \mathfrak{U} \mid u$ incident to $u$. (Here, $\mathfrak{U} \mid u$ denotes the constraint distribution of $\mathfrak{U}$ conditioned on the first vertex being $u$.)

3. Verify that $\pi^{-1}(x_v) = (\pi')^{-1}(x_{v'})$.

It is easy to check that $\mathfrak{U}(x) \geqslant 1-\eta$ implies $\mathfrak{U}^2(x) \geqslant 1-2\eta$. On the other hand, if $\mathfrak{U}^2(x) \geqslant \zeta$, then one can construct an assignment $y$ such that $\mathfrak{U}(y) \geqslant \zeta/4$.

Let $\mathfrak{U}_{\circlearrowleft}$ be the unique game obtained by sampling with probability $1/2$ a random constraint from $\mathfrak{U}$ and sampling with the remaining probability a trivial constraint $(u, u, \mathrm{id})$. Note that $\mathfrak{U}_{\circlearrowleft}(x) = 1/2(1+\mathfrak{U}(x))$ for every assignment $x \in (\Sigma \cup \{\perp\})^V$.

The square of the unique game $\mathfrak{U}_{\circlearrowleft}$ is very close to the unique game $\mathfrak{U}$. In particular, $\mathfrak{U}^2_{\circlearrowleft}(x) = 1/4 + 1/2\mathfrak{U}(x) + 1/4\mathfrak{U}^2(x)$ for every assignment $x$. Therefore, if $\mathfrak{U}(x) \geqslant 1-\eta$, then also $\mathfrak{U}^2_{\circlearrowleft}(x) \geqslant 1-\eta$. On the other hand, if $\mathfrak{U}^2_{\circlearrowleft}(x) \geqslant 1-\eta'$, then $\mathfrak{U}(x) \geqslant 1-2\eta'$. (Similar statements also hold for partial assignments.)

## 2.4. Unique Games Conjecture and Reductions

**Unique Games Conjecture.** In Chapter 1, we presented a simplified form of the Unique Games Conjecture. This simplified is formally equivalent to the original formulation of conjecture in [Kho02]. However, this equivalence is non-trivial (it follows from results of [KKMO07]).

The following statement of the Unique Games Conjecture is very close to the original formulation in [Kho02]:

**Unique Games Conjecture:** For every constant $\varepsilon > 0$, there exists $k = k(\varepsilon) \in \mathbb{N}$ such that given a unique game $\mathfrak{U}$ with alphabet size $k$, it is NP-hard to distinguish between the cases,

**YES:** $\mathrm{opt}(\mathfrak{U}) \geqslant 1-\varepsilon$,

**NO:** $\mathrm{opt}(\mathfrak{U}) < \varepsilon$.

In fewer words, for every $\varepsilon > 0$, there exists $k \in \mathbb{N}$ such that the promise problem $(1-\varepsilon, \varepsilon)$-Gap-Unique Games$(k)$ is NP-hard. (Here, Unique Games$(k)$ denotes the optimization problem given by unique games with alphabet size $k$. See §2.1 for the definition of gap-promise problems.)

To clarify further, we spell out what NP-hardness means in this context: For every constant $\varepsilon > 0$, there exists $k = k(\varepsilon) \in \mathbb{N}$ and a polynomial time reduction from 3-Sat to Unique Games such that

– every satisfiable 3-Sat instance reduces to a unique game with alphabet size $k$ and optimal value at least $1-\varepsilon$,

– every unsatisfiable 3-Sat instance reduces to a unique game with alphabet size $k$ and optimal value less than $\varepsilon$.

**Consequences of the Unique Games Conjecture.** We say a promise problem $\Pi$ is *UG-hard* if there exists a constant $\varepsilon > 0$ and an efficient reduction from Unique Games to $\Pi$ such that every unique game with value at least $1-\varepsilon$ reduces to a YES instance of $\Pi$ and every unique game with value less than $\varepsilon$ reduces to a NO instance of $\Pi$. The reduction needs to be efficient only for constant alphabet size (in particular, the blow-up of the reduction can be exponential in the alphabet size of the unique game). Assuming the Unique Games Conjecture, every UG-hard problem is also NP-hard.

We briefly discuss the known consequences of the Unique Games Conjecture for constraint satisfaction problems (CSPs). This class of problems contains many basic optimization problems, for example, Max Cut, Unique Games($k$) (with fixed alphabet size $k$), Max 3-Lin, and Max 3-Sat. Raghavendra [Rag08] showed that for every CSP $\Pi$ and every constant $c$, there exists a constant $s = s_\Pi(c)$ such that for every constant $\varepsilon > 0$, there exists an efficient algorithm that achieves a $(c+\varepsilon,\ s-\varepsilon)$-approximation and on the other hand, it is UG-hard to achieve a $(c-\varepsilon,\ s+\varepsilon)$-approximation. Hence, Raghavendra's algorithms achieve essentially optimal approximation guarantees for CSPs assuming the Unique Games Conjecture. In [Ste10a], we show that these approximation guarantees for CSPs can in fact be obtained in quasi-linear time.

**Noise Graphs.** An important ingredient of UG-hardness reductions are "noise graphs". Two kinds of noise graphs are relevant for this thesis. The first kind are *Gaussian noise graphs*, denoted $U_\rho$ for a parameter $\rho \in [0,1]$. The vertex set of these graphs is $\mathbb{R}^d$ for some $d \in \mathbb{N}$. The edge distribution of $U_\rho$ on $\mathbb{R}^d$ is given by two $d$-dimensional vectors sampled from the Gaussian distribution on $\mathbb{R}^{2d}$ with mean 0 and covariance matrix

$$\begin{pmatrix} I & \rho I \\ \rho I & I \end{pmatrix}.$$

Alternatively, we can describe the edge distribution as follows:

1. Sample two standard Gaussian vectors $x, y$ in $\mathbb{R}^d$ independently. (The coordinates of $x$ and $y$ are independent Gaussians with mean 0 and variance 1.)

2. Output an edge between $x$ and $\rho x + \sqrt{1-\rho^2}\, y$.

We identify $U_\rho$ with the following linear (Markov) operator on $L_2(\mathbb{R}^d)$ (where $\mathbb{R}^d$ is equipped with the standard Gaussian measure $\gamma$),

$$U_\rho f(x) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} f\left(\rho x + \sqrt{1-\rho^2}\, y\right) \mathrm{d}\gamma(y).$$

The second kind of noise graph, denote $T_\rho$ for some parameter $\rho \in [0,1]$, is defined on vertex set $\Omega^R$, where $\Omega$ is any (finite) probability space and $R \in \mathbb{N}$. The graph $T_\rho$ on $\Omega^R$ has the following edge distribution

1. Sample $x_1, \ldots, x_R$ independently from $\Omega$.

2. For every $r \in [R]$, sample $y_r$ as follows: With probability $\rho$, set $y_r = x_r$ and with the remaining probability $1-\rho$, sample $y_r$ from $\Omega$.

3. Output the edge $xy$, where $x = (x_1, \ldots, x_R)$ and $y = (y_1, \ldots, y_R)$.

The graph $T_\rho$ corresponds to the following linear (Markov) operator on $L_2(\Omega^R)$,

$$T_\rho f(x) \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{y \sim T_\rho(x)} f(y).$$

Here, $y \sim T_\rho(x)$ denotes that $y$ is a random neighbor of $x$ in the graph $T_\rho$, i.e., for every coordinate $r \in [R]$, we set $y_r = x_r$ with probability $\rho$ and choose $y_r \sim \Omega$ with probability $1-\rho$.

We remark that both $U_\rho$ and $T_\rho$ have second largest eigenvalue $\rho$. Since $U_\rho$ and $T_\rho$ are both product operators (acting independently on coordinates), their spectral decomposition is simple to describe. For $k \in \mathbb{N}$, let $Q_k$ be the linear operator on $L_2(\mathbb{R}^d)$ projecting on the subspace spanned by polynomials of degree $k$ that are orthogonal to all polynomials of degree $k-1$. Similarly, let $P_k$ be the linear operator on $L_2(\Omega^R)$ projecting on subspace spanned by functions that depend only on $k$ coordinates and are orthogonal to all functions depending on at most $k-1$ coordinates. The operators $U_\rho$ and $T_\rho$ have the following spectral decompositions

$$U_\rho = \sum_{k\in\mathbb{N}} \rho^k Q_k\,, \tag{2.5}$$

$$T_\rho = \sum_{k\in[R]} \rho^k P_k\,. \tag{2.6}$$

Another useful property of the operators $U_\rho$ and $T_\rho$ is the following semigroup structure: For all $\rho,\rho' \in [0,1]$, it holds that

$$U_\rho U_{\rho'} = U_{\rho\rho'}\,, \tag{2.7}$$

$$T_\rho T_{\rho'} = T_{\rho\rho'}\,. \tag{2.8}$$

**Influence.** Let $f \in L_2(\Omega^R)$ for some finite probability space $\Omega$. The *influence* of coordinate $i$, denoted $\mathrm{Inf}_i f$, is the squared norm of $f$ projected on the space of functions that depend only on coordinate $i$ and are orthogonal to constant functions. Equivalently, $\mathrm{Inf}_i f$ is the typical variance of $f(x)$ for $x_i \sim \Omega$,

$$\mathrm{Inf}_i f \stackrel{\mathrm{def}}{=} \mathop{\mathbb{E}}_{x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_R\sim\Omega} \left( \mathop{\mathbb{E}}_{x_i\sim\Omega} f(x)^2 - \Big( \mathop{\mathbb{E}}_{x_i\sim\Omega} f(x)\Big)^2 \right).$$

From the above definition, it is straight-forward to check that $\mathrm{Inf}_i f$ is convex in $f$ (using Cauchy–Schwarz).

The *total influence* of a function $\sum_{i\in[R]} \mathrm{Inf}_i f$ has the following expansion in terms of the projections $P_k f$,

$$\sum_{i\in[R]} \mathrm{Inf}_i f = \sum_k k\|P_k f\|^2.$$

A consequence is that "noised" functions can have only few influential coordinates.

**Fact 2.9.** *Let $f \in L_2(\Omega^R)$ for a finite probability space $\Omega$ and $R \in \mathbb{N}$. Then, for every $\rho \in [0,1)$,*

$$\sum_{i\in[R]} \mathrm{Inf}_i T_\rho f \leqslant \tfrac{1}{\log(1/\rho)} \|f\|^2$$

*In particular, if $\|f\|^2 = 1$, the number of coordinates with influence larger than $\tau > 0$ is bounded by $\frac{1}{\tau \log(1/\rho)}$.*

*Proof.* The total influence of $T_\rho f$ is equal to $\sum_{k\in[R]} k\rho^k \|P_k f\|^2 \leqslant \|f\|^2 \max_k k\rho^k \leqslant \|f\|^2/\log(1/\rho)$. □

**Gaussian Noise Stability.** Let $\Gamma_\rho\colon [0,1] \to [0,1]$ be the *stability profile* of the Gaussian noise graph $U_\rho$ on $\mathbb{R}$, that is,

$$\Gamma_\rho(\delta) \stackrel{\text{def}}{=} \max_{S\subseteq\mathbb{R},\ \mu(S)\leqslant\delta} U_\rho(S,S),$$

(Recall that the notation $U_\rho(S,S)$ denotes the probability that a random edge of $U_\rho$ has both endpoints in $S$.) The stability profile of $U_\rho$ is related to its expansion profile in the following way $\Phi_{U_\rho}(\delta) = 1 - \Gamma_\rho(\delta)/\delta$. A result of C. Borell [Bor85] shows that vertex sets of the form $[t,\infty)$ have maximum stability in $U_\rho$,

$$\Gamma_\rho(\delta) = \mathop{\mathbb{P}}_{xy\sim U_\rho}\{x \geqslant t,\ y \geqslant t\} \quad \text{for } t\in\mathbb{R} \text{ such that } \mathop{\mathbb{P}}_{x\sim U_\rho}\{x\geqslant t\} = \delta.$$

The Gaussian noise stability trivially satisfies $\delta^2 \leqslant \Gamma_\rho(\delta) \leqslant \delta$ for all $\delta > 0$. (This relation holds for any regular graph.) Let $\rho = 1-\varepsilon$. Using basic properties of the Gaussian distribution, one can show the following basic bounds on $\Gamma_\rho(\delta)$. For all $\varepsilon,\delta > 0$, the Gaussian noise stability satisfies $\Gamma_\rho(\delta)/\delta \leqslant \delta^{\varepsilon/2}$ [KKMO07]. If $\delta \gg 2^{-\varepsilon}$, then $\Gamma_\rho(\delta)/\delta \approx 1 - O(\sqrt{\varepsilon\log(1/\delta)})$ (e.g. [CMM06b]).

**Invariance Principle.** In this work, the following simple form of the invariance principle suffices: The graph $T_\rho$ acts on bounded functions without influential coordinates in approximately the same way as the graph $U_\rho$. The formal statement is as follows:

**Theorem 2.10** (Invariance Principle, [MOO05]). *For every finite probability space $\Omega$ and constants $\rho\in[0,1)$, $\eta > 0$, there exists constants $\tau,\gamma > 0$ such that for every function $f \in L_2(\Omega^R)$ with $0 \leqslant f \leqslant 1$, either*

$$\langle f, T_\rho f\rangle \leqslant \Gamma_\rho(\mathbb{E} f) + \eta,$$

*or* $\operatorname{Inf}_i T_{1-\gamma} f > \tau$ *for some coordinate* $i\in[R]$.

**Influence Decoding.** To describe this construction, we introduce two notations: (1) For a vertex $u$ of a unique game $\mathfrak{U}$, we write $(u,v,\pi) \sim U \mid u$ to denote a random constraint incident to $u$ in $\mathfrak{U}$. (In other words, we condition the constraint distribution on the event that $u$ is the first vertex of the constraint.) (2) For a vector $x \in \Omega^R$ (where $\Omega$ can be any set) and a permutation $\pi \in \mathbb{S}_R$ of the coordinates, we write $\pi.x$ to denote the vector in $\Omega^R$ obtained by permuting the coordinates of $x$ according to $\pi$. Formally, for all coordinates $i \in [R]$, we set $(\pi.x)_{\pi(i)} = x_i$.

The following lemma gives a sufficient condition for when we can decode a good assignment for a unique game based on influential coordinates of functions associated with the vertices of the unique game.

**Lemma 2.11** (Influence Decoding). *Let $\mathfrak{U}$ be a unique game with vertex set $V$ and alphabet $[R]$. For some probability space $\Omega$, let $\{f_u\}_{u\in V}$ be a collections of normalized functions in $L_2(\Omega^R)$. Consider functions $g_u$ in $L_2(\Omega^R)$ defined by*

$$g_u(x) = \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}|u} f_v(\pi.x).$$

*Then, for all $\gamma, \tau > 0$, there exists $c_{\gamma,\tau} > 0$ (in fact, $c_{\gamma,\tau} = \mathrm{poly}(\gamma,\tau)$) such that*

$$\mathrm{opt}(\mathfrak{U}) \geqslant c_{\gamma,\tau} \cdot \mathop{\mathbb{P}}_{u\sim V}\left\{\max_{i\in[R]} \mathrm{Inf}_i\, T_{1-\gamma} g_u > \tau\right\}.$$

*Proof.* We construct an assignment for the unique game $\mathfrak{U}$ by the following probabilistic procedure:

1. For every vertex $u \in V$, randomly choose one of the following ways to determine the label $x_u$:
   a) Assign $x_u$ randomly from the set
   $$L_f(u) := \{i \in [R] \mid \mathrm{Inf}_i\, T_{1-\gamma} f_u > \tau\}.$$
   (If the set is empty, assign $x_u = \bot$.)
   b) Assign $x_u$ randomly from the set
   $$L_g(u) := \{i \in [R] \mid \mathrm{Inf}_i\, T_{1-\gamma} g_u > \tau\}.$$
   (If the set is empty, assign $x_u = \bot$.)

First, we note that the sets $L_f(u)$ and $L_g(u)$ cannot be large. (By Fact 2.9 at most $1/\tau\gamma$ coordinates of say $T_{1-\gamma} f_u$ can have influence larger that $\tau$) To estimate the expected fraction of constraints satisfied by $x_u$, consider a vertex $u$ such that $\mathrm{Inf}_i\, T_{1-\gamma} g_u > \tau$ for some label $i \in [R]$. Since $i \in L_g(u)$, we assign $x_u = i$ with probability at least $1/2 \cdot 1/|L_g(u)| \geqslant \tau\gamma/2$.

Next, we estimate the probability that $x_v = \pi(i)$ for a constraint $(u,v,\pi) \sim \mathfrak{U} \mid u$. Using the convexity of $\mathrm{Inf}_i$ and the fact that coordinate $i$ for the function $x \mapsto f_u(\pi.x)$ corresponds to coordinate $\pi(i)$ for the function $f$, we derive that

$$\tau < \mathrm{Inf}_i\, T_{1-\gamma} g_u \leqslant \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}|u} \mathrm{Inf}_{\pi(i)}\, T_{1-\gamma} f_v\,,$$

which means that $\pi(i) \in L_f(v)$ with probability at least $\tau$ (over the choice of $(u,v,\pi) \sim \mathfrak{U} \mid u$). Hence, the assignment $x$ satisfies in expectation at least

$$\begin{aligned}\mathop{\mathbb{P}}_{(u,v,\pi)\sim\mathfrak{U}|u} \{x_v = \pi(x_u)\} &\geqslant \mathop{\mathbb{P}}_{(u,v,\pi)\sim UG|u} \{x_u = i \wedge x_v = \pi(i)\} \\ &\geqslant (\tau\gamma/2) \mathop{\mathbb{P}}_{(u,v,\pi)\sim UG|u} \{x_v = \pi(i)\} \\ &\geqslant (\tau^2\gamma^2/4) \mathop{\mathbb{P}}_{(u,v,\pi)\sim UG|u} \left\{\pi(i) \in L_f(v)\right\} \geqslant (\tau^3\gamma^2/4)\end{aligned}$$

of the constraints $(u,v,\pi)$ incident to $u$.

The expected value of the assignment $x$ is therefore at least

$$\begin{aligned}\mathop{\mathbb{P}}_{(u,v,\pi)\sim\mathfrak{U}} \{x_v = \pi(x_u)\} &= \mathop{\mathbb{E}}_{u\sim V} \mathop{\mathbb{P}}_{(u,v,\pi)\sim\mathfrak{U}|u} \{x_v = \pi(x_u)\} \\ &\geqslant (\tau^3\gamma^2/4) \mathop{\mathbb{P}}_{u\sim V} \left\{\exists i.\ \mathrm{Inf}_i\, T_{1-\gamma} g_u > \tau\right\}. \end{aligned}$$

□

# Part I.

# Algorithms

# 3. Unique Games with Expanding Constraint Graphs

Many constraint satisfaction problems have natural candidates for hard input distributions. For example, Feige's [Fei02] Random 3-Sat hypothesis asserts that it is intractable to distinguish near-satisfiable 3-Sat instances and *random* 3-Sat instances of appropriate density (which are far from being satisfiable with high probability). A striking evidence for this hypothesis is that even very strong semidefinite relaxations ($\Omega(n)$-levels of Lasserre's hierarchy [Las01]) fail to distinguish between the two cases [Sch08].

In this chapter, we present an efficient algorithm for Unique Games that provides a good approximation whenever the underlying constraint graph is mildly expanding. Since random instances tend to be expanding, this algorithm shows that Unique Games is easy for many input distributions (for example, Erdős–Renyi graphs with logarithmic average degree or random regular graphs with constant degree).

Given a unique game with optimal value $1-\varepsilon$, our algorithm finds an assignment of value $1-O(\varepsilon/\lambda)$, where $\lambda \in [0,2]$ is the spectral gap of the constraint graph.

In contrast to previous approximation algorithms for Unique Games, the approximation guarantee of this algorithm does not degrade with the size of the alphabet or the number of variables of the unique game.

## 3.1. Main Result

Let $\mathfrak{U}$ be a unique game with vertex set $V = [n]$ and alphabet $\Sigma = [k]$. Let $\lambda \in [0,2]$ be the spectral gap of the constraint graph of $\mathfrak{U}$. In this section, we present an efficient algorithm that can satisfy a constant fraction of the constraints of $\mathfrak{U}$ if $\lambda$ is significantly larger than the fraction of constraints violated in an optimal assignment for $\mathfrak{U}$. The algorithm is based on the standard semidefinite relaxation for Unique Games (see Relaxation 2.7). In the following, we denote this relaxation by $\mathrm{SDP}(\mathfrak{U})$ and its optimal value by

$\mathrm{sdp}(\mathfrak{U})$. (We refer to the optimal value of the relaxation $\mathrm{SDP}(\mathfrak{U})$ sometimes as the *semidefinite value* of $\mathfrak{U}$.) For the convenience of the reader, we restate the relaxation $\mathrm{SDP}(\mathfrak{U})$.

Given a unique game $\mathfrak{U}$ with vertex $V$ and alphabet $\Sigma$, find a collection of vectors $\{\boldsymbol{u}_i\}_{u\in V, i\in\Sigma} \subseteq \mathbb{R}^d$ for some $d \in \mathbb{N}$ so as to

$$\mathrm{SDP}(\mathfrak{U}): \quad \text{maximize} \quad \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}} \langle \boldsymbol{u}_i, \boldsymbol{v}_{\pi(i)} \rangle \tag{3.1}$$

$$\text{subject to} \quad \sum_{i\in\Sigma} \|\boldsymbol{u}_i\|^2 = 1 \quad (u \in V), \tag{3.2}$$

$$\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = 0 \quad (u \in V, i \neq j \in \Sigma). \tag{3.3}$$

**Theorem 3.1.** *Given a unique game $\mathfrak{U}$ with spectral gap $\lambda$ and semidefinite value $\mathrm{sdp}(\mathfrak{U}) \geqslant 1-\varepsilon$, we can efficiently compute an assignment $x$ for $\mathfrak{U}$ of value $\mathfrak{U}(x) \geqslant 1 - O(\varepsilon/\lambda)$.*

The main component of the above theorem is a simple (randomized) rounding algorithm that given a solution for $\mathrm{SDP}(\mathfrak{U})$ of value $1-\varepsilon$, outputs an assignments $x$ for $\mathfrak{U}$ of value $\mathfrak{U}(x) \geqslant 1 - O(\varepsilon/\lambda)$.

A solution for $\mathrm{SDP}(\mathfrak{U})$ consists of a vector $\boldsymbol{u}_i \in \mathbb{R}^d$ for every vertex $u \in V$ and label $i \in \Sigma$. To round this solution, we first pick a random "pivot vertex" $w \in V$ and choose a "seed label" $s \in \Sigma$ for $w$ according to the probability distribution given by the weights $\|\boldsymbol{w}_1\|^2, \ldots, \|\boldsymbol{w}_k\|^2$. Next, we propagate this initial choice in one step to every other vertex $u \in V$: Roughly speaking, we assign label $i \in \Sigma$ to vertex $u$ if the vector $\boldsymbol{u}_i$ is "closest" to $\boldsymbol{w}_s$.

We remark that unlike previous rounding algorithms for semidefinite relaxations, this rounding is not based on random projections of the solution vectors. (Previous projection-based algorithms in fact cannot achieve the approximation in Theorem 3.1.)

In the following, we give a more precise description of the rounding algorithm used for Theorem 3.1. For any two vertices $u, v \in V$, the algorithm defines a distance $\Delta(u,v)$, which measures the similarity of the two vector collections $\{\boldsymbol{u}_i\}_{i\in\Sigma}$ and $\{\boldsymbol{v}_i\}_{i\in\Sigma}$ (similar to the so called *earth mover's distance*). These distances will play an important role in the analysis of the rounding algorithm.

**Algorithm 3.2.** (Propagation Rounding)

***Input:*** Vector solution $\{\boldsymbol{u}_i\}_{u\in V, i\in\Sigma}$ for the relaxation SDP($\mathfrak{U}$).

***Output:*** Distribution over $k\cdot n$ assignments for the unique game $\mathfrak{U}$.

The distribution is specified by the following randomized algorithm:

- For every vertex $u \in V$, let $D_u$ be the distribution over $\Sigma$ which outputs label $i \in \Sigma$ with probability $\|\boldsymbol{u}_i\|^2$.
- For any two vertices $u, v \in V$, let

$$\Delta(u,v) \stackrel{\text{def}}{=} \min_{\sigma\in\mathbb{S}_\Sigma} \sum_{i\in\Sigma} \tfrac{1}{2}\|\boldsymbol{u}_i - \boldsymbol{v}_{\sigma(i)}\|^2$$

  and let $\sigma_{v\leftarrow u}$ be a permutation of $\Sigma$, for which the minimum $\Delta(u,v)$ is attained.
- *Pivot vertex:* Sample a vertex $w \in V$.
- *Seed label:* Sample a label $s \sim D_w$.
- *Propagation:* For every vertex $u \in V$, set $x_u := \sigma_{u\leftarrow w}(s)$ and output the assignment $x \in \Sigma^V$.

We will show that starting from a solution for SDP($\mathfrak{U}$) of value $1-\varepsilon$, Algorithm 3.2 computes an assignment $x$ that violates at most $O(\varepsilon/\lambda)$ of the constraints of $\mathfrak{U}$ in expectation. (It is easy to derandomize Algorithm 3.2, because the number of choices for the pivot vertex $w$ and the seed label $s$ is at most $k\cdot n$. Thus, we could modify the algorithm such that it outputs a list of $k\cdot n$ assignments $x^{(w,s)}$ for $\mathfrak{U}$. One of these assignments $x^{(w,s)}$ has value at least $1-O(\varepsilon/\lambda)$ for $\mathfrak{U}$.)

Let us consider a constraint $(u,v,\pi)$ in $\mathfrak{U}$. The following lemma relates the probability that this constraint is violated to three distance terms. The first term is exactly the (negative) contribution of the constraint $(u,v,\pi)$ to the value of our solution for SDP($\mathfrak{U}$). The remaining terms are averages of the distance $\Delta(u,w)$ and $\Delta(v,w)$ over all choices of vertices $w\in V$. We will prove this lemma in the next section (§3.2).

**Lemma 3.3** (Propagation Rounding). *For every constraint $(u,v,\pi)$ of $\mathfrak{U}$,*

$$\mathop{\mathbb{P}}_{x}\left\{x_v \neq \pi(x_u)\right\} \leqslant 3\sum_{i\in\Sigma}\tfrac{1}{2}\|\boldsymbol{u}_i - \boldsymbol{v}_{\pi(i)}\|^2 + 3\mathop{\mathbb{E}}_{w\in V}\Delta(u,w) + 3\mathop{\mathbb{E}}_{w\in V}\Delta(v,w).$$

To prove Theorem 3.1 using the lemma above, we need to bound the expected value of $\Delta(u,v)$ for two random vertices $u,v \in V$. It is easy to show that for a random constraint $(u,v,\pi)$ of $\mathfrak{U}$, the expectation of $\Delta(u,v)$ is at most $\varepsilon$. We will show that the expansion of the constraint graph implies that the typical value of $\Delta(u,v)$ for two random vertices $u,v \in V$ cannot be much larger than the typical value of $\Delta(u,v)$ for a random constraint $(u,v,\pi)$ in $\mathfrak{U}$. Here, the crux is that the distances $\Delta(u,v)$ can be approximated up to constant factors by squared euclidean distances. (We defer the rather technical proof of this fact to Section 3.3. Since the construction that establishes this fact is somewhat mysterious and relies on tensor products of vectors, we refer to it as "tensoring trick".)

**Lemma 3.4** (Tensoring Trick). *There exist unit vectors $X_1,\ldots,X_n \in \mathbb{R}^n$ such that for any two vertices $u,v \in V$,*

$$c \cdot \Delta(u,v) \leqslant \|X_u - X_v\|^2 \leqslant C \cdot \Delta(u,v).$$

*Here, $c$ and $C$ are absolute constants. (We can choose $c = 1/4$ and $C = 20$).*

The spectral gap $\lambda$ of the constraint graph of $\mathfrak{U}$ allows us to relate the typical value of $\|X_u - X_v\|^2$ for two random vertices $u,w \in V$ to the typical value of $\|X_u - X_v\|^2$ for a constraint $(u,v,\pi)$ in $\mathfrak{U}$,

$$\lambda \leqslant \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}} \|X_u - X_v\|^2 / \mathop{\mathbb{E}}_{u,v\in V} \|X_u - X_v\|^2.$$

Together with Lemma 3.4, we can show that in expanding constraint graphs, local correlation implies global correlation. Here, *local correlation* refers to the situation that $\Delta(u,v)$ is small for a typical constraint $(u,v,\pi)$ of $\mathfrak{U}$. *Global correlation* means that $\Delta(u,v)$ is small for two randomly chosen vertices $u,v \in V$.

**Corollary 3.5** (Local Correlation $\Longrightarrow$ Global Correlation).

$$\mathop{\mathbb{E}}_{u,v\in V} \Delta(u,v) \leqslant (C/c\lambda) \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}} \Delta(u,v).$$

By combining Lemma 3.3 and Corollary 3.5, we show in the next lemma that the assignment computed by Algorithm 3.2 satisfies $1 - O(\varepsilon/\lambda)$ of the constraints of $\mathfrak{U}$ in expectation. Theorem 3.1 is implied by this lemma.

**Lemma 3.6.** *The assignment $x$ computed by Algorithm 3.2 given a solution for* SDP($\mathfrak{U}$) *of value $1-\varepsilon$ satisfies*

$$\mathop{\mathbb{E}}_{x} \mathfrak{U}(x) \geqslant 1 - O(\varepsilon/\lambda).$$

*Proof.* Using Lemma 3.3 (Propagation Rounding) and the fact that we have a solution for SDP($\mathfrak{U}$) of value $1-\varepsilon$, the expected value of the assignment $x$ is at least

$$\begin{aligned}\mathop{\mathbb{E}}_{x} \mathfrak{U}(x) &\geqslant 1 - 3 \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}} \left[ \sum_i \tfrac{1}{2}\|u_i - v_{\pi(i)}\|^2 - 3 \mathop{\mathbb{E}}_{w\in V} \Delta(u,w) - 3 \mathop{\mathbb{E}}_{w\in V} \Delta(v,w) \right] \\ &= 1 - 3\varepsilon - 6 \mathop{\mathbb{E}}_{u,v\in V} \Delta(u,v).\end{aligned}$$

(Here, we also used that the constraint graph of $\mathfrak{U}$ is regular.) From Corollary 3.5 (Local Correlation $\Longrightarrow$ Global Correlation), it follows that

$$\mathop{\mathbb{E}}_{u,v\in V} \Delta(u,v) \leqslant (C/c\lambda) \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}} \Delta(u,v) \leqslant (C/c\lambda) \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}} \sum_i \tfrac{1}{2}\|u_i - v_{\pi(i)}\|^2 = (C/c\lambda)\varepsilon\,.$$

We conclude that

$$\mathop{\mathbb{E}}_{x} \mathfrak{U}(x) \geqslant 1 - 3\varepsilon - 6(C/c\lambda)\varepsilon = 1 - O(\varepsilon/\lambda). \qquad \square$$

## 3.2. Propagation Rounding

In this section, we will prove Lemma 3.3, which we used to analyze Algorithm 3.2 in the previous section (§3.1). Let us quickly recall some details of this algorithm. For every vertex pair $u,v \in V$, the algorithm defines a permutation $\sigma_{v\leftarrow u}$ of $\Sigma$. The algorithm outputs an assignment $x$ defined as $x_u = \sigma_{u\leftarrow w}(s)$, where $w$ is a random vertex and $s$ is a random label drawn from the probability distribution $D_w$ given by the weights $\|\boldsymbol{w}_1\|^2, \ldots, \|\boldsymbol{w}_k\|^2$. For $u,v \in V$ and $\sigma \in \mathbb{S}_\Sigma$, we introduce the following notation

$$\Delta(u,v,\sigma) \stackrel{\text{def}}{=} \sum_{i\in\Sigma} \tfrac{1}{2}\|\boldsymbol{u}_i - \boldsymbol{v}_{\sigma(i)}\|^2\,.$$

With this notation, Lemma 3.3 asserts the following bound on the probability that $x$ does not satisfy a constraint $(u,v,\pi)$ of $\mathfrak{U}$,

$$\mathop{\mathbb{P}}_{x}\{x_v \neq \pi(x_u)\} \leqslant 3\Delta(u,v,\pi) + \mathop{\mathbb{E}}_{w\in V} 3\Delta(w,v,\sigma_{v\leftarrow w}) + \mathop{\mathbb{E}}_{w\in V} 3\Delta(w,u,\sigma_{u\leftarrow w}). \tag{3.4}$$

The following lemma implies (3.4) (and therefore Lemma 3.3) by averaging over $w \in V$ and setting $\sigma = \sigma_{u\leftarrow w}$ and $\sigma' = \sigma_{v\leftarrow w}$.

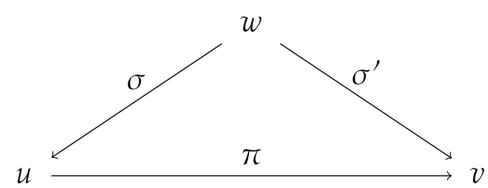


Figure 3.1.: Illustration of Lemma 3.7. The permutation $\pi$ maps labels from vertex $u$ to vertex $v$. The permutations $\sigma$ and $\sigma'$ map labels from vertex $w$ to vertices $u$ and $v$, respectively.

**Lemma 3.7.** *For any vertices $u,v,w \in V$ and permutations $\pi,\sigma,\sigma'$ of $\Sigma$,*

$$\mathop{\mathbb{P}}_{s\sim D_w}\left\{\sigma'(s) \neq \pi\circ\sigma(s)\right\} \leqslant 3\Delta(u,v,\pi) + 3\Delta(w,u,\sigma') + 3\Delta(w,v,\sigma).$$

*Proof.* For simplicity, assume that $\sigma$ and $\sigma'$ are the identity permutation. (We can achieve this situation by reordering the vector collections $\{\boldsymbol{u}_i\}_{i\in\Sigma}$ and $\{\boldsymbol{v}_j\}_{j\in\Sigma}$ and changing $\pi$ accordingly.) Since the vectors $\boldsymbol{w}_1,\ldots,\boldsymbol{w}_k$ are pairwise orthogonal, we have $\|\boldsymbol{w}_s - \boldsymbol{w}_{s'}\|^2 = \|\boldsymbol{w}_s\|^2 + \|\boldsymbol{w}_{s'}\|^2$ for $s \neq s' \in \Sigma$. Therefore,

$$\mathop{\mathbb{P}}_{s\in D_w}\left\{s \neq \pi(s)\right\} = \sum_{s\in\Sigma} \tfrac{1}{2}\|\boldsymbol{w}_s - \boldsymbol{w}_{\pi(s)}\|^2.$$

By the triangle inequality (and Cauchy–Schwarz),

$$\begin{aligned}\|\boldsymbol{w}_s - \boldsymbol{w}_{\pi(s)}\|^2 &\leqslant \left(\|\boldsymbol{w}_s - \boldsymbol{u}_s\| + \|\boldsymbol{u}_s - \boldsymbol{v}_{\pi(s)}\| + \|\boldsymbol{v}_{\pi(s)} - \boldsymbol{w}_{\pi(s)}\|\right)^2 \\ &\leqslant 3\|\boldsymbol{w}_s - \boldsymbol{u}_s\|^2 + 3\|\boldsymbol{u}_s - \boldsymbol{v}_{\pi(s)}\|^2 + 3\|\boldsymbol{v}_{\pi(s)} - \boldsymbol{w}_{\pi(s)}\|^2.\end{aligned}$$

Taking these bounds together,

$$\begin{aligned}\mathop{\mathbb{P}}_{s\in D_w}\left\{s \neq \pi(s)\right\} &\leqslant 3\sum_s \tfrac{1}{2}\|\boldsymbol{w}_s - \boldsymbol{u}_s\|^2 + 3\sum_s \tfrac{1}{2}\|\boldsymbol{u}_s - \boldsymbol{v}_{\pi(s)}\|^2 + 3\sum_s \tfrac{1}{2}\|\boldsymbol{v}_{\pi(s)} - \boldsymbol{w}_{\pi(s)}\|^2 \\ &= 3\Delta(w,u,\mathrm{id}) + 3\Delta(u,v,\pi) + 3\Delta(w,v,\mathrm{id}).\end{aligned}$$

□

## 3.3. Tensoring Trick

In this section, we prove Lemma 3.4, which allowed us in Section 3.1 to relate the approximation guarantee of Algorithm 3.2 on a unique game $\mathfrak{U}$ to the spectral gap of the constraint graph of $\mathfrak{U}$.

**Lemma** (Restatement of Lemma 3.4). *There exist unit vectors $X_1, \ldots, X_n \in \mathbb{R}^n$ such that for any two vertices $u, v \in V$,*

$$c \cdot \Delta(u,v) \leqslant \|X_u - X_v\|^2 \leqslant C \cdot \Delta(u,v).$$

*Here, $c$ and $C$ are absolute constants. (We can choose $c = 1/4$ and $C = 20$).*

For a vertex $u \in V$, we choose $X_u$ as (weighted) sum of high tensor-powers of the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$,

$$X_u := \sum_i \|\boldsymbol{u}_i\| \bar{\boldsymbol{u}}_i^{\otimes t}.$$

Here, $t \in \mathbb{N}$ is a large enough odd integer (say $t = 5$), $\bar{\boldsymbol{u}}_i$ is the unit vector in direction of $\boldsymbol{u}_i$, and $\bar{\boldsymbol{u}}_i^{\otimes t}$ is the $t$-fold tensor product[1] of $\bar{\boldsymbol{u}}_i$ with itself. The vector $X_u$ has unit norm, because the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ are pairwise orthogonal and $\sum_i \|\boldsymbol{u}_i\|^2 = 1$.

We remark that the constructed vectors $X_1, \ldots, X_n$ are elements of $\mathbb{R}^{n^t}$. However, it is straight-forward to find vectors in $\mathbb{R}^n$ with the same pairwise distances and norms (e.g., by passing to the $n$-dimensional subspace spanned by the vectors $X_1, \ldots, X_n$).

Consider two arbitrary vertices $u, v \in V$. Let $A = (A_{ij})$ be the $k$-by-$k$ matrix of inner products between the vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$,

$$A_{ij} := \langle \bar{\boldsymbol{u}}_i, \bar{\boldsymbol{v}}_j \rangle.$$

Let $p, q \in \mathbb{R}^k$ be the (nonnegative) unit vectors

$$p := \big(\|\boldsymbol{u}_1\|, \ldots, \|\boldsymbol{u}_k\|\big), \quad q := \big(\|\boldsymbol{v}_1\|, \ldots, \|\boldsymbol{v}_k\|\big).$$

With these notations, $\|X_u - X_v\|^2$ and $\Delta(u,v)$ satisfy the following identities,

$$\tfrac{1}{2}\|X_u - X_v\|^2 = 1 - \sum_{ij} p_i A_{ij}^t q_j, \tag{3.5}$$

$$\Delta(u,v) = 1 - \max_{\sigma \in \mathbb{S}_\Sigma} \sum_i p_i A_{i\sigma(i)} q_{\sigma(i)}. \tag{3.6}$$

To simplify notation, we will prove Lemma 3.4 in this more general setting of matrices $A$ and unit vectors $p$ and $q$.

[1] The *tensor product* of two vectors $x, y \in \mathbb{R}^d$ is a vector $x \otimes y \in \mathbb{R}^{d \times d}$ with $(x \otimes y)_{i,j} = x_i y_j$.

It is enough for the matrix $A$ to satisfy the following property,

$$\forall j.\ \sum_i A_{ij}^2 \leqslant 1\,, \quad \text{and} \quad \forall i.\ \sum_j A_{ij}^2 \leqslant 1\,. \tag{3.7}$$

(In the setting of Lemma 3.4, this property holds, because by orthogonality $\sum_i \langle \bar{\boldsymbol{u}}_i, \bar{\boldsymbol{v}}_j\rangle^2 \leqslant \|\bar{\boldsymbol{v}}_j\|^2 = 1$ and $\sum_j \langle \bar{\boldsymbol{u}}_i, \bar{\boldsymbol{v}}_j\rangle^2 \leqslant \|\bar{\boldsymbol{u}}_i\|^2 = 1$.)

We describe how to prove Lemma 3.4 using the lemmas in the rest of the section. The first two lemmas are simple technical observations, which are used in the proofs of the main lemmas, Lemma 3.10 and Lemma 3.11. On the one hand, Lemma 3.10 shows that for any permutation $\sigma$,

$$(1 - \textstyle\sum_{ij} p_i A_{ij}^t q_j) \leqslant 4t(1 - \sum_i p_i A_{i\sigma(i)} q_{\sigma(i)})\,,$$

which implies $\|X_u - X_v\|^2 \leqslant 4t\Delta(u,v)$. On the other hand, Lemma 3.11 shows that there exists a permutation $\sigma$ such that

$$(1 - \textstyle\sum_i p_i A_{i\sigma(i)} q_{\sigma(i)}) \leqslant 2\sum_{ij} p_i A_{ij}^t q_j / (1 - 2^{-(t-3)/2})\,,$$

which means $\Delta(u,v) \leqslant 2\|X_u - X_v\|^2/(1 - 2^{-(t-3)/2})$. Together, these two lemmas imply Lemma 3.4.

**Lemma 3.8.** *Let $p, q \in \mathbb{R}^k$ be two unit vectors and $A \in R^{k\times k}$ be a matrix with property (3.7). Suppose that for $\gamma \geqslant 0$ and $s \in \mathbb{N}$ with $s \geqslant 2$,*

$$\sum_i p_i A_{ii}^s q_i = 1 - \gamma\,.$$

*Then,*

$$\sum_{i \neq j} |p_i A_{ij}^s q_j| \leqslant \gamma\,.$$

*Proof.* Since $s \geqslant 2$ and $A$ satisfies (3.7), every row and column sum of the matrix $B = (|A_{ij}|^s)$ is bounded by 1. It follows that the largest singular value of $B$ is at most 1. Hence, $\sum_{ij} |p_i A_{ij}^s q_j| \leqslant \|p\|\|q\| = 1$, which implies the lemma. For the convenience of the reader, we give a direct proof that $\sum_{ij} |p_i A_{ij}^s q_j| \leqslant 1$,

$$\begin{aligned}\sum_{ij} |p_i A_{ij}^s q_j| &\leqslant \sum_{ij} |A_{ij}^s| \cdot \left(\tfrac{1}{2} p_i^2 + \tfrac{1}{2} q_j^2\right) \\ &= \tfrac{1}{2}\sum_i p_i^2 \sum_j |A_{ij}^s| + \tfrac{1}{2}\sum_j q_i^2 \sum_i |A_{ij}^s| \leqslant \tfrac{1}{2}\sum_i p_i^2 + \tfrac{1}{2}\sum_j q_i^2 = 1\,. \qquad \square\end{aligned}$$

**Lemma 3.9.** *Let $p, q \in \mathbb{R}^k$ be two unit vectors and $A \in R^{k\times k}$ be a matrix with property (3.7). Let $I^+$ be the set of indices $i$ such that $p_i A_{ii} q_i \geqslant 0$. Suppose that for $\gamma \geqslant 0$ and an odd number $s \in \mathbb{N}$,*

$$\sum_{i\in I^+} p_i A^s_{ii} q_i = 1-\gamma\,. \tag{3.8}$$

*Then,*

$$\sum_i p_i A^s_{ii} q_i \geqslant 1-2\gamma\,.$$

*Proof.* Note that since $s$ is odd, $p_i A^s_{ii} q_i$ has the same sign as $p_i A_{ii} q_i$ for every index $i$. From (3.8) and the fact that $|A_{ii}| \leqslant 1$, it follows that $\sum_{i\in I^+} |p_i q_i| \geqslant 1-\gamma$. On the other hand, $\sum_i |p_i q_i| \leqslant \|p\| \cdot \|q\| = 1$ by Cauchy–Schwarz. Hence, $\sum_{i\notin I^+} |p_i q_i| \leqslant \gamma$. We conclude that

$$\sum_i p_i A^s_{ii} q_i \geqslant \sum_{i\in I^+} p_i A^s_{ii} q_i - \sum_{i\notin I^+} |p_i q_i| \geqslant 1-2\gamma\,. \qquad \square$$

**Lemma 3.10.** *Let $p, q \in \mathbb{R}^k$ be two unit vectors and $A \in R^{k\times k}$ be a matrix with property (3.7). Suppose that for $\gamma \geqslant 0$,*

$$\sum_i p_i A_{ii} q_i = 1-\gamma\,. \tag{3.9}$$

*Then,*

$$\sum_{ij} p_i A^t_{ij} q_j \geqslant 1-4t\gamma\,.$$

*Proof.* By Lemma 3.8, it is enough to show

$$\sum_i p_i A^t_{ii} q_i \geqslant 1-2t\gamma\,. \tag{3.10}$$

As in Lemma 3.9, let $I^+$ be the set of indices $i$ such that $p_i A_{ii} q_i \geqslant 0$. Note that since $t$ is odd, $p_i A^t_{ii} q_i$ has the same sign as $p_i A_{ii} q_i$. By the convexity of $x \mapsto x^t$ on $\mathbb{R}_+$,

$$\sum_{i\in I^+} p_i A^t_{ii} q_i = \sum_{i\in I^+} |p_i q_i| |A_{ii}|^t \geqslant \Big(\sum_{i\in I^+} |p_i q_i| |A_{ii}|\Big)^t \overset{(3.9)}{\geqslant} (1-\gamma)^t \geqslant 1-t\gamma\,. \tag{3.11}$$

(For the first inequality, we also used that $\sum_i |p_i q_i| \leqslant \|p\| \cdot \|q\| = 1$, by Cauchy–Schwarz.) By Lemma 3.9, the bound on $\sum_{i\in I^+} p_i A^t_{ii} q_i$ in (3.11) implies (3.10) as desired. $\square$

**Lemma 3.11.** *Let $p, q \in \mathbb{R}^k$ be two unit vectors and $A \in R^{k\times k}$ be a matrix with property (3.7). Suppose that for $\gamma \geqslant 0$,*

$$\sum_{ij} p_i A_{ij}^t q_j = 1 - \gamma \,. \tag{3.12}$$

*Then, there exists a permutation $\sigma$ such that*

$$\sum_{ij} p_i A_{i\sigma(i)} q_{\sigma(i)} \geqslant 1 - 2\gamma/(1 - 2^{-(t-3)/2}) \,.$$

*Proof.* Since $A$ satisfies (3.7), every row and column contains at most one entry larger than $\sqrt{1/2}$ in absolute value. Hence, there exists a permutation $\sigma$ such that $j = \sigma(i)$ for any two indices $i, j$ with $a_{ij}^2 > 1/2$. By reordering the columns of $A$, we can assume $\sigma = \mathrm{id}$. Now, $A$ satisfies $a_{ij}^2 \leqslant 1/2$ for all indices $i \neq j$.

As in Lemma 3.9, let $I^+$ be the set of indices $i$ with $p_i A_{ii} q_i \geqslant 0$. Note that $\mathrm{sign}(p_i A_{ii} q_i) = \mathrm{sign}(p_i A_{ii}^3 q_i) = \mathrm{sign}(p_i A_{ii}^t q_i)$ (because $t$ is odd).

Choose $\eta \geqslant 0$ such that $1 - \eta = \sum_{i \in I^+} p_i A_{ii}^3 q_i$. From Lemma 3.8, it follows that $\sum_{i\neq j} |p_i A_{ij}^3 q_j| \leqslant \eta$. Therefore, the contribution of the off-diagonal terms to (3.12) is at most

$$\sum_{i\neq j} p_i A_{ij}^t q_j \leqslant \max_{i \neq j} |A_{ij}|^{t-3} \cdot \sum_{i \neq j} |p_i A_{ij}^3 q_j| \leqslant \underbrace{2^{-(t-3)/2}}_{\alpha_t} \cdot \eta \,. \tag{3.13}$$

Combining (3.12) and (3.13) yields

$$\sum_{i \in I^+} p_i A_{ii}^t q_i \geqslant \sum_{ij} p_i A_{ij}^t q_j - \sum_{i \neq j} p_i A_{ij}^t q_j \geqslant 1 - \gamma - \alpha_t \eta \,.$$

On the other hand,

$$\sum_{i \in I^+} p_i A_{ii}^t q_i \leqslant \sum_{i \in I^+} p_i A_{ii}^3 q_i = 1 - \eta \,.$$

Together these two bounds on $\sum_{i\in I^+} p_i A_{ii}^t q_i$ imply $\eta \leqslant \gamma/(1 - \alpha_t)$. We conclude that

$$\sum_{i \in I^+} p_i A_{ii} q_i \geqslant \sum_{i \in I^+} p_i A_{ii}^t q_i \geqslant 1 - \gamma/(1 - \alpha_t) \,,$$

which, by Lemma 3.9, implies that $\sum_i p_i A_{ii} q_i \geqslant 1 - 2\gamma/(1 - \alpha_t)$ as desired. □

## 3.4. Notes

The material presented in this chapter is based on the paper "Unique Games on Expanding Constraints Graphs is Easy" [AKK+08], joint with Sanjeev Arora, Subhash Khot, Alexandra Kolla, Madhur Tulsiani, and Nisheeth Vishnoi. A preliminary version of the paper appeared at STOC 2008.

### Tensoring Trick

The "tensoring trick" presented in §3.3 originates from a construction in Khot and Vishnoi's influential work [KV05] on integrality gaps in the context of the Unique Games Conjecture. The result presented in Chapter 8 is a continuation of their work. Again the tensoring trick plays an important role in the constructions presented in Chapter 8.

We remark that the conference version of the paper [AKK+08] only establishes a weaker version of Lemma 3.4 (which leads to an additional logarithmic factor in the bound of Theorem 3.1). The proof of the tighter bound is only slightly more involved than the proof of the weaker bound presented in [AKK+08].

### Propagation Rounding

Algorithm 3.2 (Propagation Rounding) is related to one of Trevisan's algorithms for Unique Games [Tre05]. Trevisan's algorithm fixes a spanning tree of the constraint graph of the unique game and propagates labels along the edges of this tree (as in Algorithm 3.2 the propagation is according to a solution of an SDP relaxation). In contrast to Trevisan's algorithms, our algorithm propagates labels in one step to every vertex in the graph. (In other words, the propagation is according to a star graph, a spanning tree of the complete graph with radius 1.)

Some ingredients of the analysis of Algorithm 3.2 are inspired by Trevisan's analysis (especially, parts of the proof of Lemma 3.7).

### Subsequent Work

Makarychev and Makarychev [MM09] established the following improvement of Theorem 3.1. Let $\mathfrak{U}$ be a unique game with optimal value $1-\varepsilon$. Suppose the spectral gap $\lambda$ of the constraint graph of $\mathfrak{U}$ satisfies $\lambda \gg \varepsilon$. Then, the algorithm in [MM09] efficiently finds an assignment for $\mathfrak{U}$ of value at least

$1 - O(\varepsilon/\Phi)$. Since $\Phi \geqslant \lambda$ (and in fact, $\Phi$ can be as large as $\Omega(\sqrt{\lambda})$), their bound is an improvement over the bound in Theorem 3.1. However, it is important to note that their result only applies to instances for which the algorithm in Theorem 3.1 finds an assignment with constant value, say value at least $1/2$. Hence, in the context of the Unique Games Conjecture, the algorithm in [MM09] does not increase the range of "easy instances" for Unique Games.

[AIMS10] and independently [RS10] extended the results in the current chapter to "small-set expanders". The precise notion of small-set expanders in these two works are quantitatively similar, but strictly speaking incomparable. (Surprisingly, the techniques are quite different except for the common root in the results of this chapter.)

Kolla and Tulsiani [KT07] found an alternative proof of (a quantitatively weaker version of) Theorem 3.1. The interesting feature of their proof is that it avoids explicit use of semidefinite programming and instead relies on enumerating certain low-dimensional subspaces (suitably discretized). Kolla [Kol10] noted that this algorithm also works (in slightly super-polynomial time) on interesting instances of Unique Games whose constraint graphs do not have large enough spectral gaps, especially the Khot–Vishnoi instances [KV05]. (The spectral gap of the constraint graphs of these instances is too small for the bound in Theorem 3.1 to be non-trivial.) The results presented in Section 5.2 (Subspace Enumeration) are inspired by the works [KT07] and [Kol10].

# 4. Parallel Repeated Unique Games

Proofs of inapproximability results similar to the Unique Games Conjecture are frequently based on hardness-amplifying self-reductions. A simple example is MAX CLIQUE. To approximate the maximum size of a clique in a graph $G$ within a factor $\alpha < 1$, it is enough to approximate the maximum size of a clique in the tensor-product graph $G^{\otimes t}$ within a factor $\alpha^t$. In this sense, the tensor-product operation amplifies the hardness of MAX CLIQUE.

In the context of UNIQUE GAMES, a natural candidate for a hardness-amplifying self-reduction is *parallel repetition*. (Using parallel repetition, Raz [Raz98] showed that the analog of the Unique Games Conjecture is true for LABEL COVER, a more general problem than UNIQUE GAMES.)

## 4.1. Main Result

Let $\mathfrak{U}$ be a unique game with vertex set $V = [n]$ and alphabet $\Sigma = [k]$. In this section, we lower bound the optimal value $\mathrm{opt}(\mathfrak{U}^\ell)$ of the repeated game in terms of the value of the semidefinite relaxation $\mathrm{SDP}(\mathfrak{U})$. (See Section 2.3 for the definition of parallel repetition.)

**Theorem 4.1.** *There exists an algorithm that given a parameter $\ell \in \mathbb{N}$ and a unique game $\mathfrak{U}$ with semidefinite value $\mathrm{sdp}(\mathfrak{U}) \geqslant 1 - \varepsilon$ and alphabet size $k$, computes in time $\mathrm{poly}(n^\ell)$ an assignment $x$ for the repeated game $\mathfrak{U}^\ell$ of value*

$$\mathfrak{U}^\ell(x) \geqslant 1 - O\left(\sqrt{\ell \varepsilon \log k}\right).$$

Before describing the proof, we record two consequence of this theorem. The first consequence is an approximation algorithm for parallel repeated unique games with the following approximation guarantee:

**Corollary 4.2.** *Suppose $\mathfrak{W}$ is an $\ell$-fold repeated unique game with alphabet size $K$ and $\operatorname{opt}(\mathfrak{W}) \geqslant 1-\eta$. Then, using the algorithm in Theorem 4.1, we can efficiently compute an assignment $x$ for $\mathfrak{W}$ of value*

$$\mathfrak{W}(x) \geqslant 1 - O\Big(\sqrt{\eta \ell^{-1} \log K}\Big). \tag{4.1}$$

For non-repeated games ($\ell = 1$), this algorithm achieves the best known approximation [CMM06a] (up to constants in the $O(\cdot)$-notation). For sufficiently many repetitions ($\ell \gg 1$), the algorithm provides strictly better approximations than [CMM06a]. (Independent of the number of repetitions, the algorithm of [CMM06a] only guarantees to satisfy $1 - O(\sqrt{\eta \log K})$ of the constraints.)

Let us briefly sketch how Theorem 4.1 implies the above approximation algorithm (Corollary 4.2). Suppose that $\mathfrak{W} = \mathfrak{U}^\ell$. Since $\mathfrak{U}$ has alphabet size $k$, the alphabet of $\mathfrak{W}$ has size $K = k^\ell$. By a result of Feige and Lovász [FL92], $\operatorname{sdp}(\mathfrak{W}) = \operatorname{sdp}(\mathfrak{U})^\ell$. Therefore, (in the relevant range of parameters) $\operatorname{sdp}(\mathfrak{U}) \geqslant 1 - O(\eta/\ell)$. It follows that the assignment obtained via Theorem 4.1 satisfies the desired bound (4.1).

The second consequence of Theorem 4.1 is a tight relation between the amortized value of a unique game and its semidefinite value.

**Corollary 4.3.**

$$\operatorname{sdp}(\mathfrak{U})^{O(\log k)} \leqslant \sup_{\ell \in \mathbb{N}} \operatorname{opt}(\mathfrak{U})^{1/\ell} \leqslant \operatorname{sdp}(\mathfrak{U}).$$

The proof of Theorem 4.1 uses an intermediate relaxation, denoted $\mathrm{SDP}_+(\mathfrak{U})$ (defined in Section 4.2). The optimal value of this relaxation, $\operatorname{sdp}_+(\mathfrak{U})$, is sandwiched between $\operatorname{opt}(\mathfrak{U})$ and $\operatorname{sdp}(\mathfrak{U})$. Furthermore, the relaxation satisfies $\operatorname{sdp}_+(\mathfrak{U}^\ell) \geqslant \operatorname{sdp}_+(\mathfrak{U})^\ell$ for every $\ell \in \mathbb{N}$. (Note that both $\operatorname{opt}(\mathfrak{U})$ and $\operatorname{sdp}(\mathfrak{U})$ have the same property.) In Section 4.2, we will show that the approximation guarantee of the intermediate relaxation $\mathrm{SDP}_+(\mathfrak{U})$ is independent of the alphabet size. (In contrast, the approximation guarantee of $\mathrm{SDP}(\mathfrak{U})$ degrades with growing alphabet size.)

**Theorem 4.4.** *Given a solution to $\mathrm{SDP}_+(\mathfrak{U})$ of value $1-\gamma$, we can efficiently compute an assignment $x$ for $\mathfrak{U}$ of value*

$$\mathfrak{U}(x) \geqslant 1 - O(\sqrt{\gamma}).$$

We remark that it is not clear whether optimal solutions for the relaxation $\mathrm{SDP}_+(\mathfrak{U})$ can be computed efficiently (in fact, if the Unique Games Conjecture is true, then this task is NP-hard). To prove Theorem 4.1, we will approximate $\mathrm{SDP}_+(\mathfrak{U})$ using the standard semidefinite relaxation $\mathrm{SDP}(\mathfrak{U})$. (See Section 4.3 for a proof of Theorem 4.5.)

**Theorem 4.5.** *There exists a polynomial-time algorithm that given a unique game $\mathfrak{U}$ with alphabet size $k$ and semidefinite value $\mathrm{sdp}(\mathfrak{U}) = 1-\varepsilon$, computes a solution for $\mathrm{SDP}_+(\mathfrak{U})$ of value at least $1 - O(\varepsilon \log k)$.*

Assuming the previous two theorems (Theorem 4.4 and Theorem 4.5), we can prove Theorem 4.1.

*Proof of Theorem 4.1.* Since $\mathrm{sdp}(\mathfrak{U}) = 1-\varepsilon$, Theorem 4.5 allows us to compute a solution for $\mathrm{SDP}_+(\mathfrak{U})$ of value $1 - O(\varepsilon \log k)$. By taking tensor products, we obtain a solution for $\mathrm{SDP}_+(\mathfrak{U}^\ell)$ of value $(1 - O(\varepsilon \log k))^\ell \geqslant 1 - O(\ell \log k)$ (see Section 4.2). Using Theorem 4.4 (on the unique game $\mathfrak{U}^\ell$), we can round this solution for $\mathrm{SDP}_+(\mathfrak{U}^\ell)$ to an assignment $x$ of value $\mathfrak{U}^\ell(x) \geqslant 1 - O(\ell \log k)^{1/2}$. □

## 4.2. Rounding Nonnegative Vectors

In this section, we introduce the intermediate relaxation $\mathrm{SDP}_+(\mathfrak{U})$ of a unique game $\mathfrak{U}$ and prove Theorem 4.4 (restated below).

**Relaxation 4.6** (Nonnegative Relaxation for Unique Games).
Given a unique game $\mathfrak{U}$ with vertex set $V$ and alphabet $\Sigma$, find a collection of nonnegative functions $\{f_{u,i}\}_{u \in V, i \in \Sigma} \subseteq L_2(\Omega)$ for some finite probability space $(\Omega, \mu)$ so as to

$$\mathrm{SDP}_+(\mathfrak{U}): \quad \text{maximize} \quad \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}} \langle f_{u,i}, f_{v,\pi(i)} \rangle \tag{4.2}$$

$$\text{subject to} \quad \sum_{i \in \Sigma} \|f_{u,i}\|^2 = 1 \quad (u \in V), \tag{4.3}$$

$$\langle f_{u,i}, f_{u,j} \rangle = 0 \quad (u \in V, i \neq j \in \Sigma). \tag{4.4}$$

An important property of the solutions for Relaxation 4.6 is that for every vertex $u \in V$, the functions $f_{u,1}, \ldots, f_{u,k}$ have disjoint support sets. In particular, for every point $\omega \in \Omega$, there exists at most one label such that $f_{u,i}(\omega)$ is positive. This property is the crucial difference from the relaxation $\mathrm{SDP}(\mathfrak{U})$, which allows to prove Theorem 4.4.

As in the previous section §4.1, let $\mathfrak{U}$ be a unique game with vertex set $V = [n]$ and alphabet $\Sigma = [k]$.

**Theorem** (Restatement of Theorem 4.4). *Given a solution to* $\mathrm{SDP}_+(\mathfrak{U})$ *of value* $1-\gamma$*, we can efficiently compute an assignment* $x$ *for* $\mathfrak{U}$ *of value*

$$\mathfrak{U}(x) \geqslant 1 - O(\sqrt{\gamma}).$$

Let $\{f_{u,i}\} \subseteq L_2(\Omega)$ be a solution for $\mathrm{SDP}_+(\mathfrak{U})$ of value $\alpha = 1-\gamma$. Let $M$ be an upper bound on the values of the functions $f_{u,i}^2$. The running time of the algorithm for Theorem 4.4 will depend polynomially on $M$. It is straightforward to ensure that $M$ is polynomial in $n$, the number of vertices of $\mathfrak{U}$.

We remark that for the purpose of proving Theorem 4.4, the reader can replace the function $\varphi(z) = (1-z)/(1+z)$ that appears in the next lemmas by the lower bound $\varphi(z) \geqslant 1-2z$. (See Figure 4.1 for a comparison of the two functions.) The more precise bounds show that our rounding for $\mathrm{SDP}_+(\mathfrak{U})$ gives non-trivial guarantees for all $\alpha > 0$ (whereas the bound stated in Theorem 4.4 is non-trivial only for $\alpha$ close enough to 1).

The first step in the proof of Theorem 4.4 is to construct a distribution over partial assignments $A\colon V \to \Sigma \cup \{\perp\}$ for $\mathfrak{U}$. If a vertex is not labeled by a partial assignment $A$ (i.e., $A$ assigns the null symbol $\perp$ to this vertex), then all constraints containing this vertex count as unsatisfied. The following lemma lower bounds the probability that a constraints is satisfied (normalized by the fraction of labeled vertices) in terms of $L_1$-distances of the functions $f_{u,i}^2$.

**Lemma 4.7.** *There is a polynomial-time randomized algorithm that samples a partial assignment* $A\colon V \to \Sigma \cup \{\perp\}$ *such that for every constraint* $(u,v,\pi)$ *of* $\mathfrak{U}$*,*

$$\mathop{\mathbb{P}}_{A}\Big\{A(u) = \pi\big(A(v)\big) \;\Big|\; A(u) \neq \perp \vee A(v) \neq \perp\Big\} \geqslant \varphi\Big(\tfrac{1}{2}\textstyle\sum_i \|f_{u,i}^2 - f_{v,\pi(i)}^2\|_1\Big),$$

*where* $\varphi(z) = (1-z)/(1+z)$*. Furthermore,* $\mathbb{P}_A\{A(u) \neq \perp\} = 1/M$ *for* $u \in V$.

*Proof.* Sample $\omega \sim \Omega$ and $t \in [0,M]$. Put $A(u) = i$ if $f_{u,i}^2(\omega) > t$ and $A(u) = \perp$ if $f_{u,j}(\omega) \leqslant t$ for all $j \in \Sigma$. (Note that since $f_{u,1},\ldots,f_{u,k}$ have pairwise disjoint supports, there is at most one label $i \in \Sigma$ with $f_{u,i}^2(\omega) > t$.) The partial assignment

$A$ satisfies the constraint $(u,v,\pi)$ of $\mathfrak{U}$ with probability

$$\begin{aligned}\mathop{\mathbb{P}}_{A}\Big\{A(v)=\pi\big(A(u)\big)\in\Sigma\Big\} &= \sum_i \mathop{\mathbb{P}}_{A}\Big\{A(u)=i\wedge A(v)=\pi(i)\Big\}\\ &= \sum_i \mathop{\mathbb{E}}_{\omega\sim\Omega}\tfrac{1}{M}\min\Big\{f_{u,i}^2(\omega),f_{v,\pi(i)}^2(\omega)\Big\}\\ &= \tfrac{1}{M}-\tfrac{1}{2M}\sum_i\|f_{u,i}^2-f_{v,\pi(i)}^2\|_1\,.\end{aligned}$$

On the other hand, we can bound the probability that $A$ assigns a label to one of the vertices $u$ and $v$,

$$\begin{aligned}\mathop{\mathbb{P}}_{A}\Big\{A(u)\neq\perp\vee A(v)\neq\perp\Big\} &= \mathop{\mathbb{E}}_{\omega\sim\Omega}\tfrac{1}{M}\max\Big\{\sum_i f_{u,i}^2(\omega),\sum_i f_{v,\pi(i)}^2(\omega)\Big\}\\ &= \tfrac{1}{M}+\tfrac{1}{2M}\Big\|\sum_i f_{u,i}^2-\sum_i f_{v,\pi(i)}^2\Big\|_1 \leqslant \tfrac{1}{M}+\tfrac{1}{2M}\sum_i\|f_{u,i}^2-f_{v,\pi(i)}^2\|_1\,.\end{aligned}$$

Combining these bounds,

$$\mathop{\mathbb{P}}_{A}\Big\{A(u)=\pi\big(A(v)\big)\;\Big|\;A(u)\neq\perp\vee A(v)\neq\perp\Big\}\geqslant\frac{1-\frac{1}{2}\sum_i\|f_{u,i}^2-f_{v,\pi(i)}^2\|_1}{1+\frac{1}{2}\sum_i\|f_{u,i}^2-f_{v,\pi(i)}^2\|_1}.\qquad\square$$

The next step in the proof of Theorem 4.4 uses correlated sampling to obtain a distribution over (total) assignments $x\in\Sigma^V$ for $\mathfrak{U}$.

**Lemma 4.8.** *We can efficiently sample an assignment $x\in\Sigma$ for $\mathfrak{U}$ such that*

$$\mathop{\mathbb{E}}_{x}\mathfrak{U}(x)\geqslant\varphi\Big(\mathbb{E}_{(u,v,\pi)\sim\mathfrak{U}}\tfrac{1}{2}\textstyle\sum_i\|f_{u,i}^2-f_{v,\pi(i)}^2\|_1\Big),$$

*where $\varphi(z)=(1-z)/(1+z)$ as before.*

*Proof.* Let $\{A^{(r)}\}_{r\in\mathbb{N}}$ be an infinite sequence of partial assignments independently sampled according to Lemma 4.7. (We will argue at the end of the proof that a small number of samples suffices in expectation.) For every vertex $u\in V$, let $r(u)$ be the smallest index $r$ such that $A^{(r)}(u)\neq\perp$. Assign the label $x_u=A^{(r(u))}(u)$ to vertex $u$. (Note that $r(u)$ is finite for every vertex with probability 1.) Consider a constraint $(u,v,\pi)$ of $\mathfrak{U}$. We claim

$$\begin{aligned}\mathop{\mathbb{P}}_{x}\{x_v=\pi(x_u)\} &\geqslant \mathop{\mathbb{P}}_{x,r}\{x_v=\pi(x_u)\wedge r(v)=r(u)\}\\ &= \mathop{\mathbb{P}}_{A}\Big\{A(v)=\pi\big(A(u)\big)\;\Big|\;A(u)\neq\perp\vee A(v)\neq\perp\Big\}\geqslant\varphi\Big(\tfrac{1}{2}\textstyle\sum_i\|f_{u,i}-f_{v,\pi(i)}\|_1\Big).\end{aligned}$$

For the equality, we use that the partial assignments $\{A^{(r)}\}$ are sampled independently. Therefore, we can condition, say, on the event $\min\{r(u), r(v)\} = 1$ without changing the probability. But this event is the same as the event $A^{(1)}(u) \neq \bot \vee A^{(1)}(v) \neq \bot$. The second inequality follows from Lemma 4.7.

Using the convexity of the function $\varphi(x) = (1-x)/(1+x)$, we can bound the expected value of the assignment $x$,

$$\mathop{\mathbb{E}}_x \mathfrak{U}(x) = \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}} \mathop{\mathbb{P}}_x \{x_v = \pi(x_u)\} \geqslant \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}} \varphi\left(\tfrac{1}{2}\textstyle\sum_i \|f_{u,i} - f_{v,\pi(i)}\|_1\right)$$
$$\geqslant \varphi\left(\mathbb{E}_{(u,v,\pi)\sim\mathfrak{U}} \tfrac{1}{2}\textstyle\sum_i \|f_{u,i} - f_{v,\pi(i)}\|_1\right).$$

To argue about the efficiency of the construction, we can estimate the expectation of $\max_{u\in V} r(u)$. (In this way, we bound the expected number of partial assignments we need to sample.) Since $\mathbb{P}_A\{A(u) \neq \bot\} = 1/M$ for every vertex $u \in V$, the expectation of $\max_{u\in V} r(u)$ is at most $M \cdot n$. □

In the final step of the proof of Theorem 4.4, we relate $L_1$-distances of the functions $f_{u,i}^2$ to the objective value $\alpha$ of our solution for $\mathrm{SDP}_+(\mathfrak{U})$.

**Lemma 4.9.**

$$\mathop{\mathbb{E}}_{(u,v,\pi)\sim G} \tfrac{1}{2} \sum_i \|f_{u,i}^2 - f_{v,\pi(i)}^2\|_1 \leqslant (1-\alpha)^{1/2}(1+\alpha)^{1/2}.$$

*Proof*. By Cauchy–Schwarz,

$$\mathop{\mathbb{E}}_{(u,v,\pi)\sim G} \tfrac{1}{2} \sum_i \|f_{u,i}^2 - f_{v,\pi(i)}^2\|_1 = \mathop{\mathbb{E}}_{(u,v,\pi)\sim G} \tfrac{1}{2} \sum_i \left\langle |f_{u,i} - f_{v,\pi(i)}|, |f_{u,i} + f_{v,\pi(i)}|\right\rangle$$
$$\leqslant \left(\mathop{\mathbb{E}}_{(u,v,\pi)\sim G} \tfrac{1}{2} \sum_i \|f_{u,i} - f_{v,\pi(i)i}\|^2\right)^{1/2} \left(\mathop{\mathbb{E}}_{(u,v,\pi)\sim G} \tfrac{1}{2} \sum_i \|f_{u,i} + f_{v,\pi(i)}\|^2\right)^{1/2}.$$

This bound proves the lemma, because $\mathbb{E}_{(u,v,\pi)\sim G} \frac{1}{2}\sum_i \|f_{u,i} - f_{v,\pi(i)i}\|^2 = 1-\alpha$ and $\mathbb{E}_{(u,v,\pi)\sim G} \frac{1}{2}\sum_i \|f_{u,i} + f_{v,\pi(i)i}\|^2 = 1+\alpha$. □

Theorem 4.4 follows by combining the previous lemmas (Lemma 4.8 and Lemma 4.9).

*Proof of Theorem 4.4*. Using Lemma 4.8, we can sample an assignment $x$ with

$$\mathop{\mathbb{E}}_x \mathfrak{U}(x) \geqslant \varphi\left(\mathbb{E}_{(u,v,\pi)\sim\mathfrak{U}} \tfrac{1}{2}\textstyle\sum_i \|f_{u,i}^2 - f_{v,\pi(i)}^2\|_1\right).$$

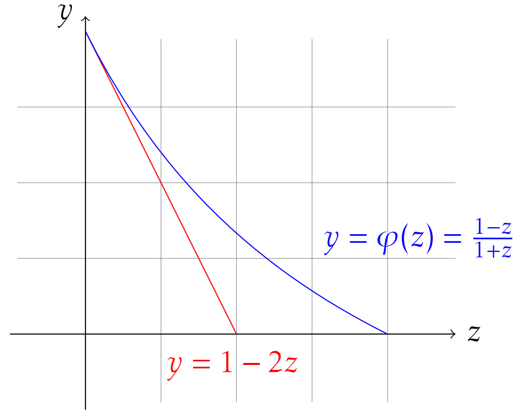


Figure 4.1.: Comparison of $1-2z$ and $(1-z)/(1+z)$

According to Lemma 4.9,

$$\mathbb{E}_{(u,v,\pi)\sim\mathfrak{U}}\,\tfrac{1}{2}\textstyle\sum_i \|f_{u,i}^2 - f_{v,\pi(i)}^2\|_1 \leqslant (1-\alpha)^{1/2}(1+\alpha)^{1/2}$$

Since $\alpha = 1-\gamma$, we have $(1-\alpha)^{1/2}(1+\alpha)^{1/2} \leqslant \sqrt{2\gamma}$. By monotonicity of $\varphi(z) = (1-z)/(1+z)$, we get $\mathbb{E}_x\,\mathfrak{U}(x) \geqslant \varphi(\sqrt{2\gamma}) \geqslant 1-2\sqrt{2\gamma}$. Using standard arguments (independently repeated trials), we can efficiently obtain an assignment $x$ with value $\mathfrak{U}(x) \geqslant 1-O(\sqrt{\gamma})$ with high probability. □

## 4.3. From Arbitrary to Nonnegative Vectors

As in the previous sections, let $\mathfrak{U}$ be a unique game with vertex set $V=[n]$ and alphabet $\Sigma=[k]$.

In this section, we show how to obtain approximate solutions for the non-negative relaxation $\mathrm{SDP}_+(\mathfrak{U})$ from an optimal solution to the semidefinite relaxation $\mathrm{SDP}(\mathfrak{U})$.

**Theorem** (Restatement of Theorem 4.5). *There exists a polynomial-time algorithm that given a unique game $\mathfrak{U}$ with alphabet size $k$ and semidefinite value* $\mathrm{sdp}(\mathfrak{U}) = 1-\varepsilon$, *computes a solution for* $\mathrm{SDP}_+(\mathfrak{U})$ *of value at least* $1-O(\varepsilon\log k)$.

Before describing the general construction, we first illustrate the theorem by proving the following special case.

*Proof of Theorem 4.5 for* $k = 2$. Let $\{\boldsymbol{u}_i\}_{u\in V, i\in\{1,2\}} \subseteq \mathbb{R}^d$ be solution for SDP($\mathfrak{U}$) of value $1-\varepsilon$. Consider an orthonormal basis $e_1,\ldots,e_d$ of $\mathbb{R}^d$. Let $(\Omega,\mu)$ be the uniform distribution over $\{1,\ldots,d\}$. For $u \in V$ and $i \in \{1,2\}$, define a function $f_{u,i} \in L_2(\Omega)$ as follows

$$f_{u,i}(r) = \begin{cases} \sqrt{d}\cdot\Big(\langle e_r,\boldsymbol{u}_i\rangle - \langle e_r,\boldsymbol{u}_{3-i}\rangle\Big) & \text{if } \langle e_r,\boldsymbol{u}_i\rangle > \langle e_r,\boldsymbol{u}_{3-i}\rangle, \\ 0 & \text{otherwise.}\end{cases}$$

(Note that $3-i$ is just the label in $\{1,2\}$ distinct from $i$.) The construction ensures that $f_{u,1}$ and $f_{u,2}$ have disjoint support and are nonnegative for every vertex $u \in V$. Since $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ are orthogonal, the construction also preserves the total mass

$$\|f_{u,1}\|^2 + \|f_{u,2}\|^2 = \sum_{r=1}^{d}\Big(\langle e_r,\boldsymbol{u}_1\rangle - \langle e_r,\boldsymbol{u}_2\rangle\Big)^2 = \|\boldsymbol{u}_1\|^2 + \|\boldsymbol{u}_2\|^2 - \underbrace{2\langle \boldsymbol{u}_1,\boldsymbol{u}_2\rangle}_{0} = 1\,.$$

Next, we show that the construction approximately preserves distances. Consider a constraint $(u,v,\pi)$ of $\mathfrak{U}$. Put $\boldsymbol{u} = \boldsymbol{u}_1 - \boldsymbol{u}_2$ and $\boldsymbol{v} = \boldsymbol{v}_{\pi(1)} - \boldsymbol{v}_{\pi(2)}$. Let $I^+$ be the set of indices $r$ such that $\operatorname{sign}\langle e_r,\boldsymbol{u}\rangle = \operatorname{sign}\langle e_r,\boldsymbol{v}\rangle$, and let $I^-$ be the complement of $I^+$. Then,

$$\sum_{i=\{1,2\}}\|f_{u,i} - f_{v,\pi(i)}\|^2 = \sum_{r\in I^+}\langle e_r,\boldsymbol{u}-\boldsymbol{v}\rangle^2 + \sum_{r\in I^-}\Big(\langle e_r,\boldsymbol{u}\rangle^2 + \langle e_r,\boldsymbol{v}\rangle^2\Big) \leqslant \|\boldsymbol{u}-\boldsymbol{v}\|^2\,.$$

(In the last step, we used that $\langle e_r,\boldsymbol{u}\rangle\langle e_r,\boldsymbol{v}\rangle < 0$ for $r \in I^-$.) By the triangle inequality, $\|\boldsymbol{u}-\boldsymbol{v}\| \leqslant \sum_{i\in\{1,2\}}\|\boldsymbol{u}_i - \boldsymbol{v}_{\pi(i)}\|$. We conclude that $\sum_{i\in\{1,2\}}\|f_{u,i} - f_{v,\pi(i)}\|^2 \leqslant 2\sum_{i\in\{1,2\}}\|\boldsymbol{u}_i - \boldsymbol{v}_{\pi(i)}\|^2$, which implies that the value of the constructed solution $\{f_{u,i}\}$ for $\mathrm{SDP}_+(\mathfrak{U})$ is at least $1-2\varepsilon$. □

For larger alphabets, the proof of Theorem 4.5 proceeds in two steps. In the first step, we construct a nearly feasible solution for $\mathrm{SDP}_+(\mathfrak{U})$. This step loses the factor $O(\log k)$ that appears in the bound of Theorem 4.5. In the second step, we show how to repair a nearly feasible solution of $\mathrm{SDP}_+(\mathfrak{U})$ while approximately preserving the objective value. The construction for this step is quite delicate, but resembles the simple construction for $k = 2$ to some extent.

### 4.3.1. Constructing Nearly Feasible Solutions

We say a collection of nonnegative functions $\{f_{u,i}\}_{u\in V, i\in\Sigma} \subseteq L_2(\Omega)$ is $\gamma$*-infeasible* for $\mathrm{SDP}_+(\mathfrak{U})$ if for every vertex $u \in V$,

$$\sum_i \|f_{u,i}\|^2 = 1\,, \tag{4.5}$$

$$\sum_{i\neq j} \langle f_{u,i}, f_{u,j}\rangle \leqslant \gamma\,. \tag{4.6}$$

In the following, we show that if $\mathrm{sdp}(\mathfrak{U}) = 1-\varepsilon$, then one can efficiently compute a $1/10$-infeasible solution for $\mathrm{SDP}_+(\mathfrak{U})$ of value $1 - O(\varepsilon \log k)$. (The bound stated in the theorem is more general.)

**Theorem 4.10.** *If* $\mathrm{sdp}(\mathfrak{U}) = 1-\varepsilon$*, then for every* $t \geqslant 1$ *we can efficiently compute a* $k\cdot 2^{-t}$*-infeasible solution for* $\mathrm{SDP}_+(\mathfrak{U})$ *of value at least* $1 - O(t)\cdot\varepsilon$*.*

Let $\{\boldsymbol{u}_i\}_{u\in V, i\in\Sigma}$ be a solution for $\mathrm{SDP}(\mathfrak{U})$ of value $1-\varepsilon$. The proof of Theorem 4.10 has two steps. First, we construct nonnegative functions $f_{u,i}$ over the space $\mathbb{R}^d$ (equipped with the $d$-dimensional Lebesgue measure). These functions will have all the desired properties (nearly feasible and high objective value). Then, we discretize the space and obtain functions defined on a finite probability space, while approximately preserving the desired properties. (The second step is quite tedious and not very interesting. We omit its proof.)

The crux of the proof of Theorem 4.10 is the following mapping from $\mathbb{R}^d$ to nonnegative functions in $L_2(\mathbb{R}^d)$,

$$M_\sigma(x) \stackrel{\mathrm{def}}{=} \|x\|\sqrt{T_{\bar{x}}\phi_\sigma}\,. \tag{4.7}$$

Here, $\bar{x} = \|x\|^{-1}x$ is the unit vector in direction $x$, $T_h$ is the translation operator on $L_2(\mathbb{R}^d)$, so that $T_h f(x) = f(x-h)$, and $\phi_\sigma \in L_2(\mathbb{R}^d)$ is the density of the $d$-dimensional Gaussian distribution $N(0,\sigma^2)^d$ with respect to the Lebesgue measure $\lambda^d$ on $\mathbb{R}^d$,

$$\phi_\sigma(x) \stackrel{\mathrm{def}}{=} \frac{1}{(\sigma\sqrt{2\pi})^d} e^{-\|x\|^2/2\sigma^2}\,.$$

Since $\int \phi_\sigma \,\mathrm{d}\lambda^d = 1$ (and the measure $\lambda^d$ is invariant under translation), the mapping $M_\sigma$ is norm-preserving, so that $\|M_\sigma(x)\| = \|x\|$ for every $x \in \mathbb{R}^d$.

The mapping $M_\sigma$ approximately preserves inner products (and thereby distances). The trade-off between the approximation of two scales will be relevant. On the one hand, orthogonal unit vectors are mapped to unit vectors with inner product $2^{-\Omega(\sigma^{-2})}$. On the other hand, the squared distance of vectors is stretched by at most a factor $\sigma^{-2}$.

**Lemma 4.11.** *For any two vectors $x, y \in \mathbb{R}^d$,*

$$\langle M_\sigma(x), M_\sigma(y)\rangle = \|x\|\,\|y\|\,e^{-\|\bar{x}-\bar{y}\|^2/8\sigma^2}. \tag{4.8}$$

*In particular, $\|M_\sigma(x) - M_\sigma(y)\|^2 \leqslant \|x-y\|^2/4\sigma^2$ (for $\sigma \leqslant 1/2$).*

*Proof.* The following identity implies (4.8) (by integration)

$$\sqrt{T_{\bar{x}}\phi_\sigma \cdot T_{\bar{y}}\phi_\sigma} = e^{-\|\bar{x}-\bar{y}\|^2/8\sigma^2}\, T_{\frac{1}{2}(\bar{x}+\bar{y})}\phi_\sigma .$$

To verify the bound $\|M_\sigma(x) - M_\sigma(y)\|^2 \leqslant \|x-y\|^2/4\sigma^2$, we first show a corresponding bound for the unit vectors $\bar{x}$ and $\bar{y}$,

$$\|M_\sigma(\bar{x}) - M_\sigma(\bar{y})\|^2 = 2 - 2e^{-\|\bar{x}-\bar{y}\|^2/8\sigma^2} \leqslant \|\bar{x}-\bar{y}\|^2/4\sigma^2 .$$

Using that the mapping $M_\sigma$ is norm-preserving, we can show the desired bound

$$\begin{aligned}\|M_\sigma(x) - M_\sigma(y)\|^2 &= \big(\|x\| - \|y\|\big)^2 + \|x\|\,\|y\| \cdot \|M_\sigma(\bar{x}) - M_\sigma(\bar{y})\|^2 \\ &\leqslant \big(\|x\| - \|y\|\big)^2/4\sigma^2 + \|x\|\,\|y\| \cdot \|M_\sigma(\bar{x}) - M_\sigma(\bar{y})\|^2/4\sigma^2 = \|x-y\|^2/4\sigma^2 .\end{aligned}$$

(The inequality uses the assumption $\sigma \leqslant 1/2$.) □

To prove Theorem 4.10, we map the vectors $\boldsymbol{u}_i$ of the solution for SDP($\mathfrak{U}$) to nonnegative functions $f_{u,i} := M_\sigma(\boldsymbol{u}_i)$ (with $\sigma^2 \approx 1/t$). Lemma 4.11 shows that the functions $f_{u,i}$ form a $k \cdot 2^{-t}$-infeasible solution for SDP$_+$($\mathfrak{U}$) of value $1 - O(t) \cdot \varepsilon$.

*Proof of Theorem 4.10.* Let $\{\boldsymbol{u}_i\}_{u \in V, i \in \Sigma}$ be a solution for SDP($\mathfrak{U}$) of value $1 - \varepsilon$. Put $f_{u,i} = M_\sigma(\boldsymbol{u}_i)$ (we fix the parameter $\sigma$ later). First, we verify that $\{f_{u,i}\}_{u \in V, i \in \Sigma}$ is $k \cdot e^{-1/4\sigma^2}$-infeasible. Since $M_\sigma$ is norm-preserving, we have

$\sum_{i\in\Sigma}\|f_{u,i}\|^2 = 1$ for every vertex $u \in V$. On the other hand, since the vectors $\{\boldsymbol{u}_i\}_{i\in\Sigma}$ are pairwise orthogonal, Lemma 4.11 shows that

$$\sum_{i\neq j}\langle f_{u,i}, f_{u,j}\rangle = e^{-1/4\sigma^2}\left(\Big(\sum_i\|\boldsymbol{u}_i\|\Big)^2 - 1\right) \leqslant k\cdot e^{-1/4\sigma^2}.$$

(Here, we used Cauchy–Schwarz to bound $\sum_i\|\boldsymbol{u}_i\|$.) The objective value of the functions $f_{u,i}$ is at least $1-\varepsilon/4\sigma^2$, because Lemma 4.11 implies that for every constraint $(u,v,\pi)$ of $\mathfrak{U}$,

$$\sum_i\|f_{u,i} - f_{v,\pi(i)}\|^2 \leqslant \sum_i\|\boldsymbol{u}_i - \boldsymbol{v}_{\pi(i)}\|^2/4\sigma^2.$$

If we choose $\sigma^2 = 1/(4t\ln 2)$, we obtain a $k\cdot 2^{-t}$-infeasible solution for $\mathrm{SDP}_+(\mathfrak{U})$ of value $1-O(t)\varepsilon$, as desired. □

### 4.3.2. Repairing Nearly Feasible Solutions

In the following, we show how to repair nearly feasible solutions for $\mathrm{SDP}_+(\mathfrak{U})$ while approximately preserving the objective value.

**Theorem 4.12.** *Given a $1/10$-infeasible solution for $\mathrm{SDP}_+(\mathfrak{U})$ of value $1-\varepsilon'$, we can efficiently compute a (feasible) solution for $\mathrm{SDP}_+(\mathfrak{U})$ of value $1-O(\varepsilon')$.*

The proof of Theorem 4.12 is quite technical and basically boils down to a careful truncation procedure (which we refer to as "soft truncation", see Lemma 4.13). A significantly simpler truncation procedure shows that given a $\gamma$-infeasible solution of value $1-\varepsilon'$, one can efficiently compute a feasible solution of value $1-O(\varepsilon'+\gamma)$. Due to this additive decrease of the objective value, this simpler truncation is not enough to show Theorem 4.5 ($\mathrm{SDP}(\mathfrak{U})$ vs. $\mathrm{SDP}_+(\mathfrak{U})$). However, it suffices to establish the following slightly weaker result: If $\mathrm{SDP}(\mathfrak{U}) = 1-\varepsilon$, we can efficiently compute a solution for $\mathrm{SDP}_+(\mathfrak{U})$ of value $1-O(\varepsilon\log(k/\varepsilon))$.

Theorem 4.12 follows from the following lemma (which asserts the existence of a mapping that we can use to transform a $1/10$-feasible solution for $\mathrm{SDP}_+(\mathfrak{U})$ to a feasible solution with roughly the same objective value).

**Lemma 4.13** (Soft Truncation). *There exists an efficiently computable mapping $Q\colon L_2(\Omega)^k \to L_2(\Omega)^k$ with the following properties: Let $f_1,\ldots,f_k$ and $g_1,\ldots,g_k$ be*

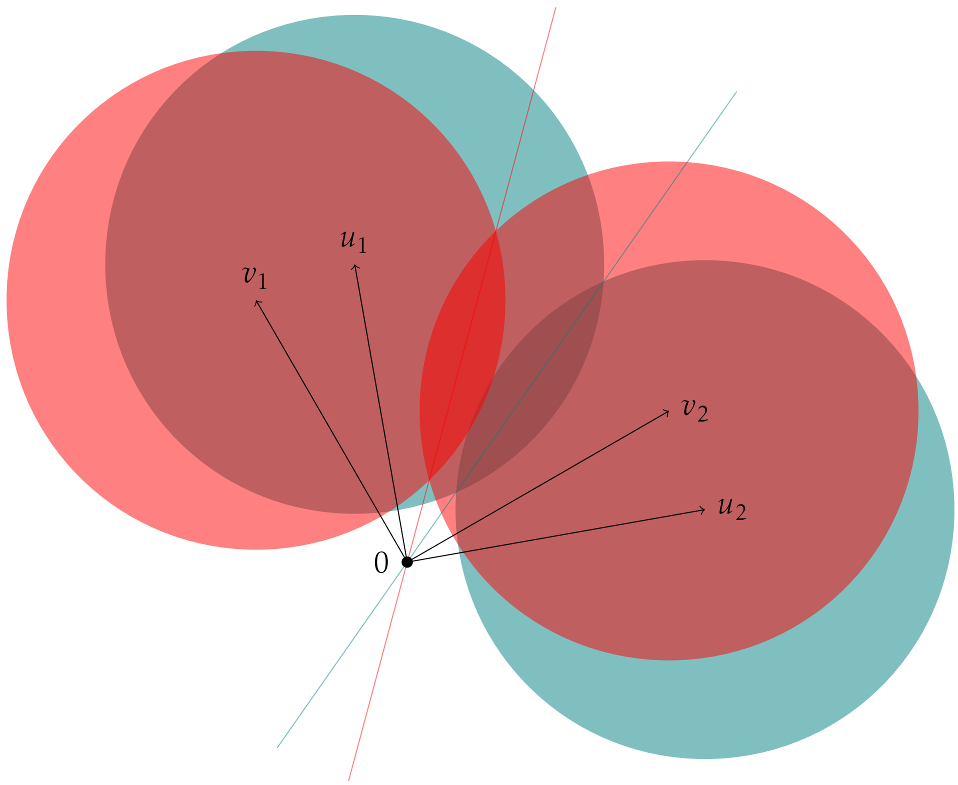


Figure 4.2.: Illustration of the construction for Theorem 4.10. With each vector $u_1, v_1, u_2, v_2$ we associate a Gaussian distribution centered at that vector (depicted as disc in the figure). The infeasibility of the solution for the $\mathrm{SDP}_+$ relaxation corresponds to the overlap of discs with the same color.

*nonnegative functions in* $L_2(\Omega)$ *such that* $\sum_i \|f_i\|^2 = \sum_i \|g_i\|^2 = 1$ *and* $\sum_{i\neq j}\langle f_i, f_j\rangle + \sum_{i\neq j}\langle g_i, g_j\rangle \leqslant \gamma$*. Suppose* $(f_1', \dots, f_k') = Q(f_1, \dots, f_k)$ *and* $(g_1', \dots, g_k') = Q(g_1, \dots, g_k)$*. Then, for every permutation* $\pi$ *of* $[k]$*,*

$$\sum_i \|f_i' - g_{\pi(i)}'\|^2 \leqslant \tfrac{32}{1-4\gamma} \sum_i \|f_i - g_{\pi(i)}\|^2 .$$

*Furthermore,* $\sum_i \|f_i'\|^2 = \sum_i \|g_i'\|^2 = 1$ *and* $\operatorname{supp}(f_i') \cap \operatorname{supp}(f_j') = \operatorname{supp}(g_i') \cap \operatorname{supp}(g_j') = \emptyset$ *for all* $i \neq j$*.*

*Proof.* The mapping $Q$ is the composition of two mappings $Q^{(1)}$ and $Q^{(2)}$. The first mapping $Q^{(1)}$ ensures that the output functions have disjoint support,

$$Q^{(1)}\colon L_2(\Omega)^k \to L_2(\Omega)^k, \quad (f_1, \dots, f_k) \mapsto (f_1', \dots, f_k'), \tag{4.9}$$

$$f_i'(x) = \begin{cases} f_i(x) - \max_{j\neq i} f_j(x) & \text{if } f_i(x) > \max_{j\neq i} f_j(x), \\ 0 & \text{otherwise.} \end{cases} \tag{4.10}$$

The purpose of the second mapping $Q^{(2)}$ is to renormalize the function (so that the squared $L_2$ norms sum to 1),

$$Q^{(2)}\colon L_2(\Omega)^k \to L_2(\Omega)^k, \tag{4.11}$$

$$\Big(f_1, \dots, f_k\Big) \mapsto \Big(\tfrac{1}{\lambda} f_1, \dots, \tfrac{1}{\lambda} f_k\Big), \tag{4.12}$$

$$\text{where } \lambda^2 = \textstyle\sum_i \|f_i\|^2 . \tag{4.13}$$

Let $f_1, \dots, f_k$ and $g_1, \dots, g_k$ be nonnegative functions in $L_2(\Omega)$ satisfying the conditions of the current lemma, i.e.,

- $\sum_{i\neq j}\langle f_i, f_j\rangle + \sum_{i\neq}\langle g_i, g_j\rangle \leqslant \gamma$
- $\sum_i \|f_i\|^2 = \sum_i \|g_i\|^2 = 1$.

Let $(f_1^{(1)}, \dots, f_k^{(1)}) = Q^{(1)}(f_1, \dots, f_k)$ and $(f_1^{(2)}, \dots, f_k^{(2)}) = Q^{(2)}(f_1^{(1)}, \dots, f_k^{(1)})$. Similarly, let $(g_1^{(1)}, \dots, g_k^{(1)}) = Q^{(1)}(g_1, \dots, g_k)$ and $(g_1^{(2)}, \dots, g_k^{(2)}) = Q^{(2)}(g_1^{(1)}, \dots, g_k^{(1)})$. Next, we establish two claims about the effects of the mappings $Q^{(1)}$ and $Q^{(2)}$ (Claim 4.14 and Claim 4.15). The current lemma (Lemma 4.13) follows by combining these two claims.

**Claim 4.14** (Properties of $Q^{(1)}$). *1. For every permutation* $\pi$ *of* $[k]$*,*

$$\sum_i \|f_i^{(1)} - g_{\pi(i)}^{(1)}\|^2 \leqslant 8 \sum_i \|f_i - g_{\pi(i)}\|^2 .$$

2. *For all* $i \neq j$,

$$\operatorname{supp}(f_i^{(1)}) \cap \operatorname{supp}(f_j^{(1)}) = \emptyset .$$

3.

$$1 \geqslant \sum_i \|f_i^{(1)}\|^2 \geqslant 1 - 2\gamma .$$

*Proof.* Item 2 holds, since by construction $\operatorname{supp}(f_i^{(1)}) = \{x \mid f_i(x) > \max_{j\neq i} f_j(x)\}$. To prove Item 3, we observe that $f_i^{(1)}(x)^2 > f_i(x)^2 - 2\sum_{j\neq i} f_i(x) f_j(x)$ and therefore as desired

$$\sum_i \left\|f_i^{(1)}\right\|^2 \geqslant \sum_i \left\|f_i\right\|^2 - 2\sum_{i\neq j} \left\langle f_i, f_j \right\rangle \geqslant 1 - 2\gamma .$$

To prove Item 1, we will show that for every $x \in \Omega$,

$$\sum_i \left(f_i^{(1)}(x) - g_{\pi(i)}^{(1)}(x)\right)^2 \leqslant 8 \sum_i \left(f_i(x) - g_{\pi(i)}(x)\right)^2 . \tag{4.14}$$

Since $Q^{(1)}$ is invariant under permutation of its inputs, we may assume $\pi$ is the identity permutation. At this point, we can verify (4.14) by an exhaustive case distinction. Fix $x \in \Omega$. Let $i_f$ be the index $i$ that maximizes $f_i(x)$. (We may assume the maximizer is unique.) Let $j_f$ be the index such that $f_{j_f}(x) = \max_{j\neq i_f} f_j(x)$. Similarly, define $i_g$ and $j_g$ such that $g_{i_g}(x) = \max_i g_i(x)$ and $g_{j_g}(x) = \max_{j\neq i_g} g_j(x)$. We may assume that $i_f = 1$ and $j_f = 2$. Furthermore, we may assume $i_g, j_g \in \{1,2,3,4\}$. Notice that the sum on the left-hand side of (4.14) has at most two non-zero terms (corresponding to the indices $i \in \{i_f, i_g\} \subseteq \{1,\ldots,4\}$). Hence, to verify (4.14), it is enough to show

$$\max_{i\in\{1,\ldots,4\}} \left|f_i^{(1)}(x) - g_i^{(1)}(x)\right| \leqslant 4 \max_{i\in\{1,\ldots,4\}} \left|f_i(x) - g_i(x)\right| . \tag{4.15}$$

Put $\varepsilon = \max_{i\in\{1,\ldots,4\}} |f_i(x) - g_i(x)|$. Let $q_i(a_1,\ldots,a_4) := \max\{a_i - \max_{j\neq i} a_j, 0\}$. Note that $f_i^{(1)}(x) = q_i\big(f_1(x),\ldots,f_4(x)\big)$ and $g_i^{(1)}(x) = q_i\big(g_1(x),\ldots,g_4(x)\big)$. The functions $q_i$ are 1-Lipschitz in each of their four inputs. It follows as desired that for every $i \in \{1,\ldots,4\}$,

$$\left|f_i^{(1)}(x) - g_i^{(1)}(x)\right| = \left|q_i\big(f_1(x),\ldots,f_4(x)\big) - q_i\big(g_1(x),\ldots,g_4(x)\big)\right| \leqslant 4\varepsilon . \qquad \square$$

**Claim 4.15** (Properties of $Q^{(2)}$). *1. For every permutation $\pi$ of $[k]$,*

$$\sum_i \|f_i^{(2)} - g_{\pi(i)}^{(2)}\|^2 \leqslant \tfrac{4}{1-4\gamma} \sum_i \|f_i^{(1)} - g_{\pi(i)}^{(1)}\|^2 .$$

2. *For all* $i \in \Sigma$,

$$\mathrm{supp}(f_i^{(2)}) = \mathrm{supp}(f_i^{(1)}).$$

3.

$$\sum_i \|f_i^{(2)}\|^2 = 1\,.$$

*Proof*. Again Item 2 and Item 3 follow immediately by definition of the mapping $Q^{(2)}$. To prove Item 1, let $\lambda_f, \lambda_g > 0$ be the multipliers such that $f_i^{(2)} = f_i^{(1)}/\lambda_f$ and $g_i^{(2)} = g_i^{(1)}/\lambda_g$ for all $i \in [k]$. Item 1 of Claim 4.14 shows that $\lambda_f^2$ and $\lambda_g^2$ lie in the interval $[1-2\gamma, 1]$. We estimate the distances between $f_i^{(2)}$ and $g_{\pi(i)}^{(2)}$ as follows,

$$\begin{aligned}
\sum_i \left\|f_i^{(2)} - g_{\pi(i)}^{(2)}\right\|^2 &= \sum_i \left\|\tfrac{1}{\lambda_f}\left(f_i^{(1)} - g_{\pi(i)}^{(1)}\right) + \left(\tfrac{1}{\lambda_f} - \tfrac{1}{\lambda_g}\right) g_i^{(1)}\right\|^2 \\
&\leqslant \tfrac{2}{\lambda_f^2} \sum_i \left\|f_i^{(1)} - g_{\pi(i)}^{(1)}\right\|^2 + 2\left(\tfrac{1}{\lambda_f} - \tfrac{1}{\lambda_g}\right)^2 \sum_i \left\|g_i^{(1)}\right\|^2 \\
&\qquad \text{(using } \|a+b\|^2 \leqslant 2\|a\|^2 + 2\|b\|^2\text{)} \\
&\leqslant \tfrac{2}{1-2\gamma} \sum_i \left\|f_i^{(1)} - g_{\pi(i)}^{(1)}\right\|^2 + 2\left(\tfrac{1}{\lambda_f} - \tfrac{1}{\lambda_g}\right)^2 \quad \text{(using } \textstyle\sum_i \|g_i^{(1)}\|^2 \leqslant 1\text{)}.
\end{aligned}$$

It remains to upper bound the second term on the right-hand side, $(1/\lambda_f - 1/\lambda_g)^2$. Since the function $x \mapsto 1/x$ is $1/a^2$-Lipschitz on an interval of the form $[a, \infty)$, we have

$$\begin{aligned}
\left|\tfrac{1}{\lambda_f} - \tfrac{1}{\lambda_g}\right| &\leqslant \tfrac{1}{1-2\gamma}\left|\lambda_f - \lambda_g\right| \\
&= \tfrac{1}{1-2\gamma}\left|\left(\textstyle\sum_i \|f_i^{(1)}\|^2\right)^{1/2} - \left(\textstyle\sum_i \|g_{\pi(i)}^{(1)}\|^2\right)^{1/2}\right| \\
&\leqslant \tfrac{1}{1-2\gamma}\left(\textstyle\sum_i \left(\|f_i^{(1)}\| - \|g_{\pi(i)}^{(1)}\|\right)^2\right)^{1/2} \quad \text{(using triangle inequality)} \\
&\leqslant \tfrac{1}{1-2\gamma}\left(\textstyle\sum_i \|f_i^{(1)} - g_{\pi(i)}^{(1)}\|^2\right)^{1/2} \quad \text{(using triangle inequality)}.
\end{aligned}$$

Combining the previous two estimates, we get as desired

$$\sum_i \left\|f_i^{(2)} - g_{\pi(i)}^{(2)}\right\|^2 \leqslant \left(\tfrac{2}{1-2\gamma} + \tfrac{2}{(1-2\gamma)^2}\right) \sum_i \left\|f_i^{(1)} - g_{\pi(i)}^{(1)}\right\|^2. \qquad \square$$

□

## 4.4. Notes

The material presented in this chapter is based on the paper "Rounding Parallel Repetitions of Unique Games"[BHH+08], joint with Boaz Barak, Moritz Hardt, Ishay Haviv, Anup Rao, and Oded Regev, and on the paper "Improved Rounding for Parallel-Repeated Unique Games" [Ste10b]. A preliminary version of the first paper appeared at FOCS 2008. The second paper appeared in preliminary form at RANDOM 2010.

### Related Works

The techniques in this chapter are natural generalizations of ideas in Raz's counterexample to strong parallel repetition [Raz08].

Before discussing the connections, we first describe Raz's result and his construction. He showed that for every $\varepsilon > 0$, there exists a unique game $\mathfrak{U}$ (Max Cut on an odd cycle of length $1/\varepsilon$) that satisfies $\operatorname{opt}(\mathfrak{U}) \leqslant 1 - \varepsilon$ and $\operatorname{opt}(\mathfrak{U}^\ell) \geqslant 1 - O(\sqrt{\ell}\,\varepsilon)$. To construct the assignment for the repeated game $\mathfrak{U}^\ell$, he assigned to every vertex of $\mathfrak{U}$ a distribution over assignments for the unique game $\mathfrak{U}$. To every vertex of the repeated game $\mathfrak{U}^\ell$, he assigned the corresponding product distribution. Finally, he obtained an assignment for $\mathfrak{U}^\ell$ using Holenstein's correlated sampling protocol [Hol09].

In retrospect, we can interpret Raz's construction in terms of techniques presented in this chapter. The distributions he assigns to the vertices of the unique game $\mathfrak{U}$ correspond to a solution to the intermediate relaxation $\mathrm{SDP}_+(\mathfrak{U})$. His construction shows that $\mathrm{sdp}_+(\mathfrak{U}) \geqslant 1 - O(\varepsilon^2)$. (In fact, his construction roughly corresponds to the proof of the special case of Theorem 4.5 for $k = 2$ that we presented in the beginning of §4.3.) The distributions he considers for the repeated game $\mathfrak{U}^\ell$ roughly correspond to a (product) solution for the relaxation $\mathrm{SDP}_+(\mathfrak{U}^\ell)$, demonstrating that $\mathrm{sdp}_+(\mathfrak{U}^\ell) \geqslant 1 - O(\ell\varepsilon^2)$. Finally, he rounds these distributions to an assignment for $\mathfrak{U}^\ell$, roughly corresponding to the proof Theorem 4.4 presented in §4.2.

In this light, the main technical contribution of the works [BHH+08] and [Ste10b] is the general construction of solutions to the intermediate relaxation $\mathrm{SDP}_+(\mathfrak{U})$ based on a solution to the relaxation $\mathrm{SDP}(\mathfrak{U})$. (Raz's construction of a solution to $\mathrm{SDP}_+(\mathfrak{U})$ relies on the specific structure of the unique games he considers.)

We note that Kindler et al. [KORW08] extended Raz's construction to the continuous case, leading to a construction of high-dimensional foams with surprisingly small surface area. Alon and Klartag [AK09] unified the results of Raz [Raz08] and Kindler et al. [KORW08] and generalized them further.

## Correlated Sampling

In the proof of Theorem 4.4 (see §4.2), we used a technique called "correlated sampling" to construct an assignment for a unique game from a suitable distribution over partial assignments. This technique was first used by Broder [Bro97] for sketching sets. In the context of rounding relaxations, this technique was introduced by Kleinberg and Tardos [KT02].

Later, Holenstein [Hol09] used this technique in his proof of the Parallel Repetition Theorem. In the context of rounding relaxations of Unique Games, this technique was used by [CMM06b] and [KRT08].

# 5. Subexponential Approximation Algorithms

Impagliazzo, Paturi, and Zane [IPZ01] showed that many NP-hard problems require strongly exponential running time — time $2^{\Omega(n)}$ — assuming 3-Sat requires strongly exponential time. Similarly, efficient PCP construction (the latest one by Moshkovitz and Raz [MR08]) show that, under the same assumption, achieving non-trivial approximation guarantees[1] for problems like Max 3-Lin and Label Cover requires running time $2^{\Omega(n^{1-o(1)})}$.

In this chapter, we show that for Unique Games, non-trivial approximation guarantees can be achieved in subexponential time, demonstrating that Unique Games is significantly easier to approximate than problems like Max 3-Lin and Label Cover. Concretely, we give an algorithm for Unique Games that achieves a $(1-\varepsilon,\ \varepsilon)$-approximation in time exponential in $k \cdot n^{O(\varepsilon^{1/3})}$, where $n$ is the number of vertices and $k$ is the alphabet size.

This algorithm for Unique Games is inspired by a (similar, but somewhat simpler) algorithm for Small-Set Expansion with the following approximation guarantee for every $\beta > 0$: Given a graph $G$ containing a vertex set $S$ with volume $\mu(S) = \delta$ and expansion $\Phi(S) = \varepsilon$, the algorithm finds in time $\exp(n^{\beta}/\delta)$ a vertex set $S'$ with volume close to $\delta$ and expansion $\Phi(S') \leqslant O(\sqrt{\varepsilon/\beta})$.

## 5.1. Main Results

In this section, we describe subexponential approximation algorithms for Small-Set Expansion and Unique Games. In the following, we use $\exp(f)$ for a function $f$ to denote a function that is bounded from above by $2^{C \cdot f}$ for some absolute constant $C \in \mathbb{N}$.

Our algorithm for Small-Set Expansion achieves the following trade-off

[1] Here (as for all of this thesis), we are interested in approximation guarantees independent of the instance size. For Max 3-Lin, a $(1-\varepsilon, 1/2+\varepsilon)$-approximation for some constant $\varepsilon > 0$ is non-trivial. For Label Cover, a $(1, \varepsilon)$-approximation for some constant $\varepsilon$ is non-trivial.

between running time and approximation guarantee. (See §5.1.1 for the proof.)

**Theorem 5.1** (Subexponential Algorithm for Small-Set Expansion).
*There exists an algorithm that given a graph $G$ on $n$ vertices containing a vertex set with volume at most $\delta$ and expansion at most $\varepsilon$, computes in time $\exp(n^\beta/\delta)$ a vertex set $S$ with volume $\mu(S) \leqslant (1+\gamma)\delta$ and expansion $\Phi(S) \leqslant O(\sqrt{\varepsilon/\beta\gamma^9})$. Here, $\beta$ is a parameter of the algorithm that can be chosen as small as $\log\log n/\log n$.*

We remark that for $\beta, \gamma \geqslant \Omega(1)$ (say $\beta = \gamma = 0.001$), the approximation guarantee of our algorithm for Small-Set Expansion matches up to constant factors the approximation guarantee of Cheeger's bound for Sparsest Cut.

Our algorithm for Unique Games achieves the following trade-off between running time and approximation guarantee.

**Theorem 5.2** (Subexponential Algorithm for Unique Games).
*There exists an algorithm that given a unique game $\mathfrak{U}$ with $n$ vertices, alphabet size $k$, and optimal value $\mathrm{opt}(\mathfrak{U}) \geqslant 1-\varepsilon$, computes in time $\exp(kn^\beta)$ an assignment $x$ of value $\mathfrak{U}(x) \geqslant 1 - O(\sqrt{\varepsilon/\beta^3})$. Here, $\beta$ is a parameter of the algorithm that can be chosen as small as $\log\log n/\log n$.*

An immediate consequence of the above theorem is that we can compute in time exponential in $n^{\tilde{O}(\varepsilon^{1/3})}$ an assignment that satisfies a constant fraction (say, 1/2) of the constraints of $\mathfrak{U}$ (where $\mathfrak{U}$ is as in the statement of the theorem). See §5.1.2 for a proof of this special case of Theorem 5.2. (Section 5.5 contains the proof of the general case of Theorem 5.2.)

Both algorithms rely on the notion of threshold rank of a matrix (and the related notion of soft threshold rank). For a symmetric matrix $A$ and $\tau \geqslant 0$, we define the *threshold rank* of $A$ at $\tau$, denoted $\mathrm{rank}_\tau(A)$, as the number of eigenvalues of $A$ larger than $\tau$ in absolute value. The threshold rank at 0 coincides with the usual notion of rank. In general, the threshold rank at $\tau$ satisfies

$$\mathrm{rank}_\tau(A) = \min_{\|B\| \leqslant \tau} \mathrm{rank}(A - B),$$

where $B$ is a symmetric matrix with the same dimension as $A$ and $\|B\|$ denotes the spectral norm of $B$ (the largest eigenvalue in absolute value).

The algorithm for Small-Set Expansion is significantly simpler than the algorithm for Unique Games, but it contains the main ideas.

### 5.1.1. Small Set Expansion

Let $G$ be a graph with vertex set $V = [n]$. For simplicity, assume that $G$ is regular. We identify $G$ with its stochastic adjacency matrix. For $\eta \in [0,1)$, let $U_\eta$ denote the subspace of $L_2(V)$ spanned by the eigenfunctions of $G$ with eigenvalue larger than $1-\eta$. Note that $\dim U_\eta = \mathrm{rank}_{1-\eta/2}(\frac{1}{2} + \frac{1}{2}G) \leqslant \mathrm{rank}_{1-\eta}(G)$.

The next lemma formalizes the following idea: Suppose $G$ contains a set with low expansion. Then, given the projection of the indicator function of this set onto $U_\eta$ (for appropriate $\eta$), we can reconstruct a set with roughly the same volume and expansion. (For the reconstruction, we just try all level sets of the projection.) The lemma also asserts that the reconstruction is robust, i.e., it succeeds even given a function in $U_\eta$ close to the true projection.

**Lemma 5.3.** *Let $\varepsilon, \delta, \eta > 0$ such that $\varepsilon/\eta$ is sufficiently small. Suppose the graph $G$ contains a vertex set with volume at most $\delta$ and expansion at most $\varepsilon$. Then, every $\varepsilon/100$-net of the subspace $U_\eta$ contains a function with a level set $S \subseteq V$ satisfying $\mu(S) \leqslant 1.1\delta$ and $\Phi(S) \leqslant O(\varepsilon)$.*

(A subset $N \subseteq U$ is an *$\varepsilon$-net* of (the unit ball of) a subspace $U$ if for every vector $x \in U$ with norm $\|x\| \leqslant 1$, the net contains a vector $y \in N$ with $\|x-y\| \leqslant \varepsilon$.)

An immediate consequence of the previous lemma is an approximation algorithm for Small-Set Expansion with running time exponential in $\dim U_\eta$. (Here, we use that for a subspace of dimension $d$, one can construct $\varepsilon$-nets of size $(1/\varepsilon)^{O(d)}$.)

The next lemma shows how to find small sets with low expansion when the dimension of $U_\eta$ is large. In this case, there exists a vertex and number of steps $t$ such that one can find a set with small volume and low expansion among the level sets of the probability distribution given by the $t$-step lazy random walk started at this vertex.

**Lemma 5.4.** *Suppose $\dim U_\eta > n^\beta/\delta$ for $n^\beta$ sufficiently large. Then, there exist a vertex $i \in V$ and a number $t \in \{1, \ldots, \eta^{-1}\log n\}$ such that a level set $S$ of $(\frac{1}{2}I + \frac{1}{2}G)^t \mathbb{1}_i$ satisfies $\mu(S) \leqslant \delta$ and $\Phi(S) \leqslant O(\sqrt{\eta/\beta})$.*

Theorem 5.1 (Subexponential Algorithm for Small-Set Expansion) follows by combining the previous lemmas.

*Proof of Theorem 5.1.* Let $\eta = O(\varepsilon/\gamma^9)$. If $\dim U_\eta > n^\beta/\delta$, then, using Lemma 5.4, we can efficiently find a vertex set $S$ with $\mu(S) \leqslant \delta$ and $\Phi(S) \leqslant$

$O(\sqrt{\eta/\beta}) = O(\sqrt{\varepsilon/\beta\gamma^9})$. Otherwise, if $\dim U_\eta < n^\beta/\delta$, we construct an $\varepsilon/100$-net $N$ of $U_\eta$ of size $(1/\varepsilon)^{O(n^\beta/\delta)}$ and enumerate all level sets of functions in $N$. Since the given graph $G$ contains a vertex set with volume at most $\delta$ and expansion at most $\varepsilon$, by Lemma 5.3, one of enumerated level sets $S$ satisfies $\mu(S) \leqslant (1+\gamma)\delta$ and $\Phi(S) \leqslant O(\sqrt{\varepsilon/\gamma^9})$. □

### 5.1.2. Unique Games

Related to the notion of threshold rank, we define the *soft-threshold rank* of a symmetric matrix $A$ at $\tau$ as

$$\operatorname{rank}^*_\tau(A) \stackrel{\text{def}}{=} \inf_{t\in\mathbb{N}} \operatorname{Tr} A^{2t}/\tau^{2t}.$$

The soft-threshold rank upper bounds the threshold rank, $\operatorname{rank}^*_\tau(A) \geqslant \operatorname{rank}_\tau(A)$, because $\operatorname{Tr} A^{2t} \geqslant \tau^{2t} \cdot \operatorname{rank}_\tau(A)$ for all $t \in \mathbb{N}$.

The following theorem gives an algorithm for Unique Games that achieves a good approximation in time exponential in the alphabet size and the soft-threshold rank of the constraint graph (for appropriate threshold value).

**Theorem 5.5.** *Given $\eta > 0$ and a regular unique game $\mathcal{U}$ with $n$ vertices, alphabet size $k$, constraint graph $G$, and optimal value* $\operatorname{opt}(\mathcal{U}) = 1-\varepsilon$, *we can compute in time* $\exp(k \cdot \operatorname{rank}^*_{1-\eta}(G))$ *an assignment $x$ of value $\mathcal{U}(x) \geqslant 1 - O(\varepsilon/\eta)$.*

For a regular graph $G$ and a vertex subset $A$, let $G[A]$ denote the "regularized subgraph" induced by $A$, i.e., we restrict $G$ to the vertex set $A$ and add self-loops to each vertices in $A$ so as to restore its original degree.

The following theorem shows that by changing only a small fraction of its edges, any regular graph can be turned into a vertex disjoint union of regular graphs with few large eigenvalues.

**Theorem 5.6** (Threshold rank decomposition)**.** *There exists a polynomial-time algorithm that given parameters $\eta, \beta > 0$ and a regular graph $G$ with $n$ vertices, computes a partition $\mathcal{P}$ of the vertices of $G$ such that* $\operatorname{rank}^*_{1-\eta}(G[A]) \leqslant n^\beta$ *for all $A \in \mathcal{P}$ and at most an $O(\sqrt{\eta/\beta^3})$ fraction of the edges of $G$ does not respect the partition $\mathcal{P}$.*

Theorem 5.2 (Subexponential algorithm for Unique Games) follows by combining Theorem 5.5 and Theorem 5.6.

*Proof of Theorem 5.2 (special case).* Let $\mathfrak{U}$ be a regular unique game with vertex set $V = [n]$, alphabet size $k$, and optimal value $\mathrm{opt}(\mathfrak{U}) = 1 - \varepsilon$. Let $x$ be an optimal assignment for $\mathfrak{U}$ so that $\mathfrak{U}(x) = 1 - \varepsilon$. Let $\eta > 0$ (we determine this parameter later). Apply Theorem 5.6 to the constraint graph $G$ of $\mathfrak{U}$ and $\mathscr{P} = \{A_1, \ldots, A_r\}$ be the resulting partition of $V$. We replace every constraint that does not respect the partition $\mathscr{P}$ by two self-loops with identity constraints (each of the new constraints has weight half). After replacing constraints that do not respect $\mathscr{P}$ in this way, the unique game $\mathfrak{U}$ decomposes into vertex disjoint unique games $\mathfrak{U}_1, \ldots, \mathfrak{U}_r$ with vertex sets $A_1, \ldots, A_r$, respectively. The constraint graph of $\mathfrak{U}_i$ is $G[A_i]$ (the regularized subgraph of $G$ induced by $A_i$). It holds that $\sum_i \mu(A_i)\mathfrak{U}_i(x) \geqslant \mathfrak{U}(x)$ (since self-loops with identity constraints are always satisfied) and therefore

$$\sum_i \mu(A_i)\,\mathrm{opt}(\mathfrak{U}_i) \geqslant 1 - \varepsilon. \tag{5.1}$$

By Theorem 5.6, we have $\mathrm{rank}^*_{1-\eta}(G[A_i]) \leqslant n^\beta$ for every $i \in [r]$. Hence, using Theorem 5.5, can compute in time $\mathrm{poly}(n)\exp(n^\beta)$ assignments $x^{(1)}, \ldots, x^{(r)}$ for $\mathfrak{U}_1, \ldots, \mathfrak{U}_r$ such that $\mathfrak{U}_i(x^{(i)}) \geqslant 1 - \varepsilon_i/\eta$, where $\varepsilon_i \geqslant 0$ is such that $\mathrm{opt}(\mathfrak{U}_i) = 1 - \varepsilon_i$. From (5.1) it follows that $\sum_i \mu(A_i)\varepsilon_i \leqslant \varepsilon$ and thus, $\sum_i \mu(A_i)\mathfrak{U}_i(x^{(i)}) \geqslant 1 - \varepsilon/\eta$. Let $x'$ be the concatenation of the assignments $x^{(1)}, \ldots, x^{(r)}$. Since at most $O(\sqrt{\eta/\beta^3})$ of the constraints of $\mathfrak{U}$ do not respect the partition $\{A_1, \ldots, A_r\}$, the value of the computed assignment $x'$ is at least $\mathfrak{U}(x') \geqslant 1 - \varepsilon/\eta - O(\sqrt{\eta/\beta^3})$. Hence, in order to satisfy a constant fraction, say 0.9, of the constraints, we can choose $\eta = O(\varepsilon)$ and $\beta = O(\varepsilon^{1/3})$. In this case, the total running time of the algorithm is exponential in $n^{O(\varepsilon^{1/3})}$. □

## 5.2. Subspace Enumeration

In this section, we prove Lemma 5.3 (see §5.2.1) and Theorem 5.5 (see §5.2.2). Before, we prove a general theorem (Theorem 5.7), which we use in the proof of Theorem 5.5 and which also implies a quantitatively weaker version of Lemma 5.3.

For a regular graph $G$ (again identified with its stochastic adjacency matrix) and $\eta \in [0,1)$, define $\mathrm{rank}^+_{1-\eta}(G)$ as the number of eigenvalues of $G$ larger than $1-\eta$ (in contrast to $\mathrm{rank}_{1-\eta}(G)$, we only consider positive eigenvalues). Clearly, $\mathrm{rank}^+_{1-\eta}(G) \leqslant \mathrm{rank}_{1-\eta}(G)$. We can express $\mathrm{rank}^+_{1-\eta}(G)$ in terms of the usual threshold rank, for example

$$\operatorname{rank}^+_{1-\eta}(G) = \operatorname{rank}_{1-\eta/2}\left(\tfrac{1}{2}I + \tfrac{1}{2}G\right).$$

Most lemmas and theorems in this section use $\operatorname{rank}^+_{1-\eta}(G)$ instead of $\operatorname{rank}_{1-\eta}(G)$.

The next theorem shows that given a regular graph and a parameter $\eta$, we can compute a list of vertex sets of length exponential in the threshold rank of the graph at $1-\eta$ such that every vertex set with expansion significantly smaller than $\eta$ is close to one of the sets in the list.

**Theorem 5.7.** *Given a parameter $\eta > 0$ and a regular graph $G$ with vertex set $V = [n]$, we can compute a list of vertex sets $S_1, \ldots, S_r \subseteq V$ with $r \leqslant \exp(\operatorname{rank}^+_{1-\eta}(G) \log n)$ such that for every vertex set $S \subseteq V$, there exists an index $i \in [r]$ with*

$$\frac{\mu(S \Delta S_i)}{\mu(S)} \leqslant 8\,\Phi(S)/\eta\,.$$

*(The computation is efficient in the output size.)*

*Proof.* Let $U_\eta$ be the subspace of $L_2(V)$ spanned by the eigenfunctions of $G$ with eigenvalue at least $1-\eta$. Note $\dim U_\eta = \operatorname{rank}^+_{1-\eta}(G)$. Construct a $1/99n$-net $N$ for (the unit ball of) $U_\eta$ of size $\exp(\dim U_\varepsilon \cdot \log n)$. Compute $S_1, \ldots, S_r$ by enumerating the level sets of the functions in $N$.

To verify the conclusion of the theorem, consider a vertex set $S$ with expansion $\Phi(S) = \varepsilon$. Let $L = I - G$ be the Laplacian of $G$ and let $P_\eta$ be the projector onto $U_\eta$. On the one hand, $\langle \mathbb{1}_S, L\mathbb{1}_S\rangle = \varepsilon \|\mathbb{1}_S\|^2$. On the other hand,

$$\langle \mathbb{1}_S, L\mathbb{1}_S\rangle = \langle P_\eta \mathbb{1}_S, L(P_\eta \mathbb{1}_S)\rangle + \langle \mathbb{1}_S - P_\eta \mathbb{1}_S, L(\mathbb{1}_S - P_\eta \mathbb{1}_S)\rangle \geqslant \eta \|\mathbb{1}_S - P_\eta \mathbb{1}_S\|^2\,.$$

(The inequality uses that $\mathbb{1}_S - P_\eta \mathbb{1}_S$ is orthogonal to $U_\eta$ and thus in the span of the eigenfunctions of $L$ with eigenvalue larger than $\eta$.) It follows that $\|\mathbb{1}_S - P_\eta \mathbb{1}_S\|^2 \leqslant (\varepsilon/\eta)\|\mathbb{1}_S\|^2$. Since $N$ is a $1/99n$-net of $U_\eta$, it contains a function $f \in N$ such that $\|\mathbb{1}_S - f\|^2 < (\varepsilon/\eta)\|\mathbb{1}_S\|^2 + 1/8n$. Consider the index $i \in [r]$ such that $S_i$ is the level set $S_i = \{j \in V \mid f(j) \geqslant 1/2\}$. We bound the volume of the symmetric difference of $S$ and $S_i$,

$$\mu(S \Delta S_i) \leqslant \mathop{\mathbb{P}}_{j \in V}\left\{|\mathbb{1}_S(j) - f(j)| \geqslant 1/2\right\} \leqslant 4\|\mathbb{1}_S - f\|^2 < 4(\varepsilon/\eta)\|\mathbb{1}_S\|^2 + 1/2n\,.$$

It follows that $\mu(S \Delta S_i)$ is at most $8(\varepsilon/\eta)\mu(S)$. (Strictly speaking, we assume here $4(\varepsilon/\eta)\|\mathbb{1}_S\|^2 \geqslant 1/2n$. However, otherwise, we get $\mu(S \Delta S_i) < 1/n$, which means that $S = S_i$ and thus $\mu(S \Delta S_i) = 0$.) $\square$

**Remark 5.8.** For simplicity, the proof of the theorem used a net of width $1/\mathrm{poly}(n)$. For interesting range of the parameters, a net of constant, say $\eta$, width is enough. In this case, the length of the list (and essentially the running time) is bounded by $n \cdot \exp(\mathrm{rank}^+_{1-\eta}(G)\log(1/\eta))$.

### 5.2.1. Small-Set Expansion

Let $G$ be a regular graph with vertex set $V = [n]$. For $\eta \in [0,1]$, let $U_\eta \subseteq L_2(V)$ be the subspace spanned by the eigenfunctions of $G$ with eigenvalue larger than $1-\eta$.

Recall the spectral profile of $G$ (see §2.2) ,

$$\Lambda(\delta) = \min_{\substack{f \in L_2(V)\\ \|f\|_1^2 \leqslant \delta \|f\|^2}} \frac{\langle f, Lf\rangle}{\langle f, f\rangle},$$

where $L = I - G$ is the Laplacian of $G$.

The next lemma shows the following: Suppose $\Lambda(\delta) = \varepsilon$ (which is the case if the given graph $G$ contains a vertex set with volume at most $\delta$ and expansion at most $\varepsilon$). Then, given a fine-enough net of the unit ball of $U_\eta$ for $\eta \gg \varepsilon$, one can efficiently reconstruct a function $h$ that demonstrates that $\Lambda(4/3\,\delta) = O(\varepsilon)$. The constant hidden by the $O(\cdot)$-notation depends on the ratio of $\eta$ and $\varepsilon$ (denoted by $\beta$ in the statement of the lemma). Here, the constant $4/3$ can be replaced by any constant larger than one (leading to larger constants in the $O(\cdot)$-notation).

**Lemma 5.9.** *Let $\beta, \gamma > 0$ be sufficiently small with $\gamma \geqslant \sqrt{\beta}$. Suppose $\Lambda(\delta) = \varepsilon$ and let $\eta = \varepsilon/\beta$. Then for every $\varepsilon/100$-net $N$ of $U_\eta$ , there exists a function $g \in N$ and a threshold value $\tau$ such that the function $h = \max\{g - \tau, 0\} \in L_2(V)$ satisfies*

$$\langle h, Lh\rangle \leqslant (\varepsilon/\beta\gamma)\langle h, h\rangle,$$
$$\mu(\mathrm{supp}(h)) \leqslant (1 + 16\gamma^{1/3})\delta.$$

*Proof.* Let $f$ be a minimizer for the spectral profile $\Lambda(\delta)$, so that $\langle f, Lf\rangle = \varepsilon\|f\|^2$ and $\|f\|_1^2 \leqslant \delta\|f\|^2$. We may assume $f \geqslant 0$ and $\|f\|^2 = 1$. Let $P_\eta$ be the projector onto $U_\eta$. Then, $\|f - P_\eta f\|^2 \leqslant \varepsilon/\eta$. (Same argument as in the proof of Theorem 5.7.) Hence, $N$ contains a function $g$ such that $\|f - g\|^2 \leqslant 2\varepsilon/\eta$. Since $g \in U_\eta$, we have $\langle g, Lg\rangle \leqslant \eta\|g\|^2$. Therefore, the function $h$ also satisfies $\langle h, Lh\rangle \leqslant \eta\|g\|^2$.

For a threshold value $\tau$, we can upper bound the volume of the support of $h$ by roughly $1/\tau^2$. (Later we choose $\tau \approx 1/\sqrt{\delta}$, so that the support has volume roughly $\delta$.) For $\alpha = (\varepsilon/\eta)^{1/3}/2 \leqslant 1/2$,

$$\mu(\operatorname{supp}(h)) = \mathbb{P}\{g_i > \tau\} \leqslant \mathbb{P}\{|f_i - g_i| > \alpha\tau\} + \mathbb{P}\{f_i > (1-\alpha)\tau\}$$
$$\leqslant \tfrac{1}{\alpha^2\tau^2}\|f-g\|^2 + \tfrac{1}{(1-\alpha)^2\tau^2}\|f\|^2 \leqslant \left(1 + 4\alpha + \tfrac{2\varepsilon/\eta}{\alpha^2}\right)\cdot \tfrac{1}{\tau^2} = \left(1 + 10(\varepsilon/\eta)^{1/3}\right)\tfrac{1}{\tau^2}. \quad (5.2)$$

On the other hand, we can lower bound the $L_2$-mass of $h$ in terms of the $L_2$-mass of the function $h' = \max\{f-\tau, 0\}$

$$\|h\|^2 = \mathop{\mathbb{E}}_i \max\{g_i - \tau, 0\}^2 \geqslant \mathop{\mathbb{E}}_i \max\{f_i - \tau, 0\}^2 - 2\mathop{\mathbb{E}}_i \max\{f_i - \tau, 0\}|f_i - g_i| - \mathop{\mathbb{E}}_i (f_i - g_i)^2$$
$$\geqslant \|h'\|^2 - 2\|h'\|\|f-g\| - \|f-g\|^2 \geqslant \|h'\|^2 - 5\sqrt{\varepsilon/\eta}. \quad (5.3)$$

(Here, the first inequality holds for every vertex $i \in V$. The second inequality uses Cauchy–Schwarz. For the third inequality, we use $\|h'\| \leqslant \|f\| = 1$ and $\|f-g\| \leqslant \sqrt{2\varepsilon/\eta}$.) If we choose $\tau = (1-2\gamma^{1/3})^2/\sqrt{\delta}$, then Lemma 5.10 asserts that $\|h'\|^2 \geqslant 8\gamma$. Combined with (5.3), we get $\|h\|^2 \geqslant 3\gamma$. (Recall that $\gamma \geqslant \sqrt{\varepsilon/\eta}$.)

We can now verify that $h$ has the desired properties. On the one hand,

$$\frac{\langle h, Lh\rangle}{\langle h, h\rangle} \leqslant \frac{\eta\|g\|^2}{3\gamma} \leqslant \eta/\gamma.$$

On the other hand, using (5.2),

$$\mu(\operatorname{supp}(h)) \leqslant (1 + 10(\varepsilon/\eta)^{1/3})\tfrac{1}{\tau^2} \leqslant (1 + 16\gamma^{1/3})\delta.$$

(The second inequality uses that $1/\tau^2 \leqslant (1+6\gamma^{1/3})\delta$ for sufficiently small $\gamma$.) □

The following technical lemma is used in the proof of the previous lemma.

**Lemma 5.10.** *Let $(\Omega, \mu)$ be a finite probability space. For a nonnegative function $f \in L_2(\Omega)$ and a threshold value $\tau \leqslant (1-\varepsilon)^2\|f\|^2/\|f\|_1$, consider $g = \max\{f - \tau, 0\}$. Then,*

$$\|g\|^2 \geqslant \varepsilon^3\|f\|^2$$

*Proof.* For $u = \varepsilon^2$, let $A \subseteq \Omega$ be the set of points $\omega \in \Omega$ such that $g(\omega)^2 \geqslant u f(\omega)^2$ and let $B = \Omega \setminus A$ be its complement. For every point $\omega \in B$, the function $f$ satisfies $(f(\omega) - \tau)^2 \leqslant u f(\omega)^2$ and thus $(1-\varepsilon)f(\omega) \leqslant \tau$. Then,

$$\|g\|^2 \geqslant u\langle \mathbb{1}_A, f^2\rangle = u\left(\|f\|^2 - \langle \mathbb{1}_B f, f\rangle\right) \geqslant u\left(\|f\|^2 - \|f\|_1 \cdot \tau/(1-\varepsilon)\right)$$

(The last inequality uses that $\mathbb{1}_B f \leqslant \tau/(1-\varepsilon)$ pointwise.) The claimed lower bound on $\|g\|^2$ follows from the assumption about $\tau$. □

The following lemma follows by combining Lemma 5.9 and Lemma 2.2 (local Cheeger bound).

**Lemma 5.11.** *Let $\beta > 0$ and $\gamma \geqslant 16\beta^{1/6}$ be small enough. Suppose $\Phi(\delta) = \varepsilon$. Then for $\eta = \varepsilon/\beta$, every $\varepsilon/100$-net of the unit ball of $U_\eta$ contains a function with a level set $S$ that satisfies*

$$\mu(S) \leqslant (1+\gamma)\,\delta\,,$$
$$\Phi(S) \leqslant O\Big(\sqrt{\varepsilon/\beta\gamma^3}\Big).$$

In the above lemma, a convenient parameter choice is $\beta = (\gamma/16)^6$. In this case, the vertex set $S$ has expansion $\Phi(S) \leqslant O(\sqrt{\varepsilon/\gamma^9})$.

### 5.2.2. Unique Games

Let $\mathfrak{U}$ be a unique game with vertex set $V = [n]$ and alphabet $\Sigma = [k]$. Let $G = G(\mathfrak{U})$ be the constraint graph of $\mathfrak{U}$ (see §2.3 for a definition).

The *label-extended graph* $\hat{G} = \hat{G}(\mathfrak{U})$ is defined as follows (also see §2.3): The vertex set $\hat{V}$ of $\hat{G}$ consists of all pairs of vertices $u \in V$ and labels $i \in \Sigma$. For ease of notation, we denote such a pair by $u_i$. The edge distribution of $\hat{G}$ is generated as follows: Sample a constraint $(u, v, \pi) \sim \mathfrak{U}$. Sample a label $i \in \Sigma$. Output an edge between $u_i$ and $v_{\pi(i)}$.

For $\eta \in [0,1)$, let $\hat{U}_\eta$ be the subspace of $L_2(\hat{V})$ spanned by the eigenfunctions of $\hat{G}$ with eigenvalue larger than $1-\eta$. The dimension of $\hat{U}_\eta$ is equal to $\operatorname{rank}^+_{1-\eta}(\hat{G})$.

The algorithms described in this subsection run in time exponential in the (positive) threshold rank of the label-extended graph. The following lemma allows us to upper bound (soft-)threshold rank of the label extended graph $\hat{G}$ by $k$ times the soft-threshold rank of the constraint $G$.

**Lemma 5.12.** *For every $t \in \mathbb{N}$,*

$$\operatorname{Tr} \hat{G}^t \leqslant k \operatorname{Tr} G^t\,.$$

*Proof.* Let $u_i \in \hat{V}$ be a label-extended vertex. The diagonal entry of $\hat{G}^t$ corresponding to $u_i$ is equal to the probability that the $t$-step random walk in $\hat{G}$ started at $u_i$ returns to $u_i$. The projection of this random walk corresponds to the $t$-step random walk in the constraint graph $G$ started at $u$. Hence, the probability that the random walk in $\hat{G}$ returns to $u_i$ is at most the probability that the random walk in $G$ return to $u$. It follows that

$$(\hat{G}^t)_{u_i,u_i} \leqslant \sum_{j\in\Sigma}(\hat{G}^t)_{u_i,u_j} = (G^t)_{u,u}\,,$$

which implies $\operatorname{Tr}\hat{G}^t \leqslant k \operatorname{Tr} G^t$ as desired. □

**Lemma 5.13.** *Let $\beta > 0$ be sufficiently small. Suppose $\mathrm{opt}(\mathfrak{U}) = 1-\varepsilon$ and let $\eta = \varepsilon/\beta$. Then every $\varepsilon/100$-net $N$ of the unit ball of $\hat{U}_\eta$, contains a function $g \in N$ such that the function $h \in L_2(\hat{V})$ defined by $h(u_i) = \max\{g(u_i) - \max_{j\neq i} g(u_j), 0\}$ has squared norm at least 1/2 and satisfies*

$$\langle h, \hat{L}h\rangle \leqslant 16(\varepsilon/\beta)\langle h,h\rangle\,,$$

*where $\hat{L} = I - \hat{G}$ is the Laplacian of the label-extended graph $\hat{G}$.*

*Proof.* The proof is similar to the proof of Lemma 5.9. The construction of $h$ is related to the smooth nonnegative orthogonalization in §4.3.2.

Let $x$ be an assignment such that $\mathfrak{U}(x) = 1-\varepsilon$. Consider $f \in L_2(\hat{V})$ such that $f(u_i) = k$ if $x_u = i$ and $f(u_i) = 0$ otherwise (i.e., $f$ is $k$ times the indicator function of the assignment $x$). The function $f$ satisfies $\langle f, Lf\rangle = \varepsilon\langle f,f\rangle$ and $\|f\|^2 = 1$. Following the argument in Lemma 5.9, the net $N$ contains a function $g$ such that $\|f-g\|^2 \leqslant 2\varepsilon/\eta$. Since $g$ is in the unit ball of the subspace $\hat{U}_\eta$, it satisfies $\langle g, \hat{L}g\rangle \leqslant \eta = \varepsilon/\beta$. The analysis of the smooth nonnegative orthogonalization in §4.3.2 shows that

$$\langle h, \hat{L}h\rangle \leqslant 8\langle g, \hat{L}g\rangle\,.$$

It remains to lower bound the squared norm of $h$. It is enough to show that $h$ is close to the function $f$ (which implies that the squared norm of $h$ cannot be much smaller than the squared norm of $f$). Consider the (non-linear) operator $Q$ on $L_2(\hat{V})$ that maps $g$ to $h$ (extended to all of $L_2(\hat{V})$ in the natural way). Notice that $Qf(u_i) = \max\{f(u_i) - \max_{i\neq j} f(u_j), 0\} = f(u_i)$ (the operator $Q$ maps $f$ to itself), since $f$ has only one non-zero among the points $u_1,\ldots,u_k$. Hence, we can estimate the distance $h$ and $f$ in terms of the Lipschitz constant $L_Q$ of the operator $Q$,

$$\|h-f\|^2 = \|Qg - Qf\|^2 \leqslant L_Q^2\|g-f\|^2 \leqslant L_Q^2 \cdot 2\beta\,.$$

It is straight-forward to verify that the Lipschitz constant $L_Q$ is an absolute constant (see the analysis in §4.3.2 for details). It follows that $\|h\|^2 \geqslant \|f\|^2 - O(\|f-h\|) \geqslant 1 - O(\beta)^{1/2}$. Since $\beta$ is assumed to be sufficiently small, we get $\|h\|^2 \geqslant 1/2$. □

The next lemma follows by combining the previous lemma and Lemma 2.2 (local Cheeger bound). If we apply the local Cheeger bound to the function $h$ in the previous lemma, the resulting vertex set in the label-extended graph corresponds to a partial assignment. (The transformation used to construct $h$ in fact ensures that the support of $h$ is a partial assignment.)

**Lemma 5.14.** *Let $\beta > 0$ be sufficiently small. Suppose $\mathrm{opt}(\mathfrak{U}) = 1 - \varepsilon$ and let $\eta = \varepsilon/\beta$. Then, in time $\mathrm{poly}(n)\,(1/\varepsilon)^{O(\dim U_\eta)}$, we can compute a partial assignment $x \in (\Sigma \cup \{\bot\})^V$ with $\mathbb{P}_u\{x_u \neq \bot\} = \alpha > 0$ that satisfies*

$$\mathfrak{U}(x) \geqslant (1 - O(\sqrt{\varepsilon/\beta})) \cdot \alpha\,.$$

*Here, one can also ensure $\alpha \geqslant 1/4$.*

## 5.3. Threshold Rank vs. Small-Set Expansion

In this section we prove Lemma 5.4 (restated below). The general goal is to establish a relation between the threshold rank and the expansion profile of graphs.

Let $G$ be a regular graph with vertex set $V = [n]$. Let $L = I - G$ be the Laplacian of $G$. (Recall that we identify $G$ with its stochastic adjacency matrix.) For $\eta \in [0,1)$, let $U_\eta$ denote the subspace of $L_2(V)$ spanned by the eigenfunctions of $G$ with eigenvalue larger than $1-\eta$.

**Lemma** (Restatement of Lemma 5.4). *Suppose $\dim U_\eta > n^\beta/\delta$ for $n^\beta$ sufficiently large. Then, there exist a vertex $i \in V$ and a number $t \in \{1, \ldots, \eta^{-1}\log n\}$ such that a level set $S$ of $(\frac{1}{2}I + \frac{1}{2}G)^t \mathbb{1}_i$ satisfies $\mu(S) \leqslant \delta$ and $\Phi(S) \leqslant O(\sqrt{\eta/\beta})$.*

An important ingredient of the proof of Lemma 5.4 (and other results in this section) is the following local variant of Cheeger's inequality. (The lemma follows directly from Lemma 2.2 in Chapter 2.)

**Lemma 5.15** (Local Cheeger Bound). *Suppose $f \in L_2(V)$ satisfies $\|f\|_1^2 \leqslant \delta\|f\|^2$ and $\langle f, Lf\rangle \leqslant \varepsilon\|f\|^2$. Then, for every $\gamma > 0$, there exists a level set $S$ of $f$ such that $\mu(S) \leqslant \delta/(1-\gamma)$ and $\Phi(S) \leqslant \sqrt{2\varepsilon}/\gamma$.*

Since $\dim U_\eta = \mathrm{rank}_{1-\eta/2}(\frac{1}{2}I + \frac{1}{2}G) \leqslant \mathrm{rank}^*_{1-\eta/2}(\frac{1}{2}I + \frac{1}{2}G)$, the following lemma combined with Lemma 5.15 (Local Cheeger Bound) implies Lemma 5.4. Note that $\frac{1}{2}I + \frac{1}{2}G = I - L/2$.

**Lemma 5.16.** *Suppose* $\mathrm{rank}^*_{1-\eta/2}(I - L/2) > n^\beta/\delta$ *for* $\eta > 0$ *small enough and* $n^\beta$ *large enough. Then, there exist a vertex* $i \in V$ *and a number* $t \in \{0, \ldots, (\beta/\eta)\log n\}$ *such that the function* $f = (I - L/2)^t \mathbb{1}_i$ *satisfies* $\|f\|_1^2 \leqslant \delta\|f\|^2$ *and* $\langle f, Lf\rangle/\langle f, f\rangle \leqslant 5\eta/\beta$.

*Proof.* Let $f_{t,i} = (I - L/2)^t \mathbb{1}_i$ and $T = \lfloor(\beta/2\eta)\ln n\rfloor$. Since $I - L/2$ is symmetric stochastic, the $L_1$-norm of $f_{t,i}$ is decreasing in $t$, so that $\|f_{t,i}\|_1 \leqslant \|f_{0,i}\|_1 = 1/n$ (in fact, equality holds). The lower bound on the soft-threshold rank of $I - L/2$ implies for every $t \in \mathbb{N}$ (in particular, for $t = T$),

$$(1 - \eta/2)^{2t} \cdot n^\beta/\delta \leqslant \mathrm{Tr}(I - L/2)^{2t} = n\sum_i \langle \mathbb{1}_i, (I - L/2)^{2t}\mathbb{1}_i\rangle = n\sum_i \|f_{t,i}\|^2 .$$

By our choice of $T$, we have $(1 - \eta/2)^{2T} n^\beta \geqslant 1$ (using that $\eta$ is small enough). Therefore, there exists a vertex $i \in V$ such that $\|f_{T,i}\|^2 \geqslant 1/\delta n^2$. The following claims imply the current lemma:

(1) for every $t \in \{0, \ldots, T\}$, it holds that $\|f_{t,i}\|_1^2 \leqslant \delta\|f_{t,i}\|^2$, and

(2) there exists $t \in \{0, \ldots, T - 1\}$ such that $\langle f_{t,i}, Lf_{t,i}\rangle \leqslant \varepsilon$ for $\varepsilon = 5\eta/\beta$.

The first claim holds because on the one hand, $\|f_{t,i}\|_1 \leqslant \|f_{0,i}\|_1 = 1/n^2$ and on the other hand, $\|f_{t,i}\|^2 \geqslant \|f_{T,i}\|^2 \geqslant 1/\delta n^2$. (Here, we use that the matrix $I - L/2$, being symmetric stochastic, decreases both $L_1$-norms and $L_2$-norms.) It remains to verify the second claim. Since $\|f_{T,i}\|^2 \geqslant 1/\delta n^2 = \|f_{0,i}\|^2/\delta n$, there exists $t \in \{0, \ldots, T - 1\}$ such that $\|f_{t+1,i}\|^2 \geqslant (\delta n)^{-1/T}\|f_{t,i}\|^2$. By our choice of $T$, we have $(\delta n)^{-1/T} \geqslant 1 - \varepsilon/2$ for $\varepsilon = 5\eta/\beta$ (using that $n$ is large enough). It follows that

$$(1 - \varepsilon/2)\|f_{t,i}\|^2 \leqslant \|f_{t+1,i}\|^2 = \langle f_{t,i}, (I - L/2)^2, f_{t,i}\rangle \leqslant \langle f_{t,i}, (I - L/2)f_{t,i}\rangle$$

(Here, we use that $(I - L/2)^2 - (I - L/2) = L/2 - L^2/4$ is positive semidefinite, because the eigenvalues of $L/2$ lies in $[0, 1]$.) □

## 5.4. Low Threshold Rank Decomposition

The goal of this section is to prove Theorem 5.6.

**Theorem** (Restatement of Theorem 5.6). *There exists a polynomial-time algorithm that given parameters $\eta, \beta > 0$ and a regular graph $G$ with $n$ vertices, computes a partition $\mathscr{P}$ of the vertices of $G$ such that $\mathrm{rank}^*_{1-\eta}(G[A]) \leqslant n^\beta$ for all $A \in \mathscr{P}$ and at most an $O(\sqrt{\eta/\beta^3})$ fraction of the edges of $G$ does not respect the partition $\mathscr{P}$.*

Recall that $G[A]$ denote the "regularized subgraph" induced by $A$, i.e., we restrict $G$ to the vertex set $A$ and add self-loops to each vertices in $A$ so as to restore its original degree.

The main ingredient of the proof of Theorem 5.6 is the following consequence of Lemma 5.16 (and Lemma 5.15, the local Cheeger bound).

**Lemma 5.17.** *There exists a polynomial-time algorithm that given parameters $\eta, \beta > 0$, a regular graph $G$, and a vertex set $K$ of $G$ such that $\mathrm{rank}^*_{1-\eta}(G[A]) > n^\beta$, computes a vertex set $S \subseteq K$ with volume $\mu(S) \leqslant \mu(A)/n^{\Omega(\beta)}$ and expansion $\Phi_{G[K]}(S) \leqslant O(\sqrt{\eta/\beta})$ in the graph $G[K]$.*

We remark that the condition $\mathrm{rank}^*_{1-\eta}(G[A]) > n^\beta$ in general does not imply the condition $\mathrm{rank}^*_{1-\eta}(\frac{1}{2}I + \frac{1}{2}G[A]) > n^\beta$ required for Lemma 5.16 (in case $G[A]$ has negative eigenvalues). However, it holds that $\mathrm{rank}^*_{1-\eta}(G[A]) \leqslant \mathrm{rank}^*_{1-\eta}(\frac{1}{2}I + \frac{1}{2}G[A]) + \mathrm{rank}^*_{1-\eta}(\frac{1}{2}I - \frac{1}{2}G[A])$. Furthermore, Lemma 5.16 also works under the condition that $\mathrm{rank}^*_{1-\eta}(\frac{1}{2}I - \frac{1}{2}G[A])$ is large (same proof).

The proof of Theorem 5.6 follows by iterating the previous lemma.

*Proof of Theorem 5.6.* We decompose the given graph $G$ using the following iterative algorithm. As long as $G$ contains a connected component $K \subseteq V(G)$ with $\mathrm{rank}^*_{1-\eta}(G[A]) > n^\beta$, do the following steps:

1. Apply Lemma 5.17 to the component $K$ to obtain a vertex set $S \subseteq K$ with volume $\mu(S) \leqslant \mu(A)/n^{\Omega(\beta)}$ and expansion $\Phi(S) \leqslant O(\sqrt{\eta/\beta})$. (Since $K$ is a connected component of $G$, the vertex set $S \subseteq K$ has the same expansion in the graphs $G$ and $G[K]$.)

2. Replace every edge leaving $S$ by two self-loops on the endpoints of the edge. (Each self-loop has weight half.)

The algorithm terminates after at most $n$ iterations (in each iteration the number of connected components increases by at least one). When the algorithm terminates, every connected component satisfies the desired bound on the soft-threshold rank.

It remains to bound the fraction of edges changed by the algorithm. Consider the following charging argument. When we replace the edge leaving a vertex set $S$, we charge cost $\Phi(S)$ to every vertex in $S$. Let $\mathsf{cost}(i)$ be total cost charged to vertex $i$ until the algorithm terminates. The fraction of edges changed by the algorithm is no more than $\mathbb{E}_{i\sim V(G)}\,\mathsf{cost}(i)$. How often can we charge the same vertex? If we charge a vertex $t$ times, then the volume of its connected component $K$ is at most $\mu(K) \leqslant n^{-\Omega(t\beta)}$. It follows that every vertex is charge at most $O(1/\beta)$ times. Hence, we can bound the total cost charged to a vertex $i$ by $\mathsf{cost}(i) \leqslant O(\sqrt{\eta/\beta^3})$, which implies the desired bound on the fraction of edges changed by the algorithm. □

## 5.5. Putting things together

In this section, we prove the general case of Theorem 5.2 (restated below).

**Theorem** (Restatement of Theorem 5.2). *There exists an algorithm that given a unique game $\mathcal{U}$ with $n$ vertices, alphabet size $k$, and optimal value* $\mathrm{opt}(\mathcal{U}) \geqslant 1-\varepsilon$, *computes in time* $\exp(kn^{\beta})$ *an assignment $x$ of value* $\mathcal{U}(x) \geqslant 1 - O(\sqrt{\varepsilon/\beta^3})$. *Here, $\beta$ is a parameter of the algorithm that can be chosen as small as* $\log\log n/\log n$.

*Proof.* To prove the theorem it is enough to show how to compute in the desired running a *partial* assignment $x$ with value $\mathcal{U}(x) \geqslant (1-O(\sqrt{\varepsilon/\beta^3}))\alpha$, where $\alpha > 0$ is the fraction of vertices labeled by $x$. (We can then extend this partial assignment to a total one by iterating the algorithm for the remaining unlabeled vertices. After at most $n$ iterations, all vertices are labeled. It is straight-forward to verify that the approximation guarantee does not suffer computing an assignment in this iterative fashion.)

To compute the desired partial assignment, we first decompose the constraint graph $G$ of the unique game $\mathcal{U}$ using Theorem 5.6. In this way, we obtain a partition $\mathscr{P} = \{A_1,\ldots,A_r\}$ of the vertex set such that $\mathrm{rank}^*_{1-\eta}(G[A_i]) \leqslant n^{\beta}$ for all $i \in [r]$ and at most a $1-O(\sqrt{\eta/\beta^3})$ fraction of the constraints of $\mathcal{U}$ do not respect the partition $\mathscr{P}$. Here, we choose the parameter $\eta > 0$ as $\eta = O(\varepsilon)$ (for a sufficiently large constant in the $O(\cdot)$-notation).

From the unique game $\mathfrak{U}$, we construct unique games $\mathfrak{U}_1,\ldots,\mathfrak{U}_r$ with vertex sets $A_1,\ldots,A_r$, respectively. (We replace every constraint of $\mathfrak{U}$ that does not respect the partition by two identity constraints on the endpoints of weight half.) In this way, the constraint graph of $\mathfrak{U}_i$ is the graph $G[A_i]$ (the regularized subgraph of $G$ induced by $A_i$).

An averaging argument shows that there exists $i \in [r]$ such that $\mathrm{opt}(\mathfrak{U}_i) \geqslant 1 - 2\varepsilon$ and $\Phi_G(A_i) = O(\sqrt{\eta/\beta^3})$.

Using Lemma 5.14, we can compute partial assignments $x^{(i)}$ for the unique game $\mathfrak{U}_i$ that labels an $\alpha$ fraction of the vertices in $A_i$ such that $\mathfrak{U}_i(x^{(i)}) \geqslant (1 - O(\sqrt{\varepsilon}))\alpha$ (using that $\varepsilon/\eta$ is sufficiently small). Furthermore, $\alpha_i \geqslant 1/4$. How good a partial assignment is $x^{(i)}$ for the original unique game $\mathfrak{U}$? Since $x^{(i)}$ is a partial assignment for the vertex set $A_i$, we have

$$\mathfrak{U}(x^{(i)}) \geqslant \mu(A_i)\mathfrak{U}_i(x^{(i)}) - \Phi(A_i)\mu(A_i)$$

(using the fact we replaced the constraints leaving $A_i$ by identity constraints). Since $\alpha \geqslant 1/4$, it holds that

$$\mathfrak{U}(x^{(i)}) \geqslant (1 - O(\sqrt{\varepsilon}) - 4\Phi(A_i))\alpha\mu(A_i)\,.$$

Since $x^{(i)}$ labels an $\alpha\mu(A_i)$ fraction of the vertices of $\mathfrak{U}$ and $\Phi(A_i) \leqslant O(\sqrt{\varepsilon/\beta^3})$, the partial assignment $x^{(i)}$ has the required properties,

$$\mathfrak{U}(x^{(i)}) \geqslant (1 - O(\sqrt{\varepsilon/\beta^3})) \mathop{\mathbb{P}}_{u\sim\mathfrak{U}}\left\{x_u^{(i)} \neq \bot\right\}. \qquad \square$$

## 5.6. Notes

The material presented in this chapter is based on the paper "Subexponential Algorithms for Unique Games and Related Problems" [ABS10], joint with Sanjeev Arora and Boaz Barak. A preliminary version of the paper appeared at FOCS 2010.

### Subspace Enumeration

Some of the material presented in Section 5.2 (Subspace Enumeration) is inspired by the works of Kolla and Tulsiani [KT07] and Kolla [Kol10]. In [KT07], the authors present an alternative proof of a quantitatively slightly weaker version of Theorem 3.1, the main theorem of Chapter 3 (Unique Games with Expanding Constraint Graph). To find a good assignment for a unique game, their algorithm enumerates a suitably discretized subspace of real-valued functions on the label-extended graph of the unique game (as in Section 5.2.2). They show that if a unique game with alphabet size $k$ has an assignment with value close to 1 and its constraint graph is an expander, then the label-extended graph has at most $k$ eigenvalues close to 1. (We omit the precise quantitative trade-offs for this discussion.) Furthermore, they show that any good assignment for the unique game corresponds to a function close to the subspace spanned by the eigenfunctions of the label-extended graph with eigenvalue close to 1. It follows that by enumerating this subspace (suitably discretized) one finds a function from which one can recover a good assignment for the unique game. (The running time is exponential in the dimension of this subspace, which is bounded by $k$ in this case.) Their recovering procedure roughly corresponds to the construction in the proof of Theorem 5.7 (though we establish this theorem in the more general context of graph expansion). In (the proof of) Lemma 5.14 we give a more sophisticated recovering procedure which has quantitatively better approximation guarantees.

Kolla [Kol10] noted that several unique games considered in the literature (in particular, the ones in [KV05]) have label-extended graphs with only few (say $\operatorname{poly}\log n$) eigenvalues close to 1 (even though their constraint graphs are not good enough expanders). Since the algorithm in [KT07] runs in time exponential in the number of eigenvalues of the label-extended graph that are close to 1, it follows that for the unique games in [KV05] the algorithm of [KT07] runs in quasi-polynomial time. (Since the unique games in [KV05] have no good assignments, the algorithm of [KT07] will certify that indeed no good assignments exists.) Furthermore, Kolla [Kol10] showed that if the constraint graph of a unique game has only few large eigenvalues, then also the label-extended graph has few large eigenvalues assuming that the eigenfunctions of the constraint graph are "well-spread" (in the sense that the $L_\infty$ and the $\ell_2$ norm are proportional). It follows that the algorithm of [KT07] applies to all unique games whose constraint graphs satisfy these properties (which includes the unique games in [KV05]). In this chapter, we use Lemma 5.12 to related the eigenvalues of the constraint graph to the eigenvalues of the label-extended graph. The advantage of Lemma 5.12 is that it does not assume additional properties of eigenfunctions of the constraint graphs (which are hard to control in general). The proof of Lemma 5.12 is not very involved. It follows by comparing the behavior of random walks on the label-extended graph to random walks on the constraint graph.

## Decomposition Approach for Unique Games

Our subexponential algorithm for Unique Games works by decomposing a general instance into a collection of (independent) "easy instances". In the context of Unique Games, this divide-and-conquer strategy has been investigated by two previous works, which we discuss in the following.

One of Trevisan's algorithms [Tre05] removes a constant fraction of the constraints in order to decompose a general unique game into a collection of disjoint unique games whose constraint graphs have spectral gap at least $1/\operatorname{poly}\log n$. His algorithm gives a non-trivial approximation guarantee if the value of the unique game is at least $1 - 1/\operatorname{poly}\log n$.

Arora et al. [AIMS10] studied a variation of Trevisan's approach (combined with the results in Chapter 3). Here, the idea is to remove a constant fraction of constraints in order to decompose a general unique game with $n$ vertices into disjoint instances that are either small (containing at most $n/2^{\Omega(1/\varepsilon)}$ vertices) or their constraint graph has spectral gap at least $\varepsilon$. (A decomposition with precisely these guarantees is not achieved in [AIMS10]. However, the algorithms works *as if* these guarantees were achieved. We refer to [AIMS10] for details.) For small instances, one can find good assignments in time $2^{n/2^{\Omega(1/\varepsilon)}}$ (using brute-force enumeration) and for instances whose constraint graph has spectral gap $\varepsilon$, one can use the algorithm for Unique Games presented in Chapter 3. The algorithm of [AIMS10], given a unique game with value $1-\varepsilon$, find an assignment with value at least $1/2$ in time

# Part II.

# Reductions

# 6. Graph Expansion and the Unique Games Conjecture

A major shortcoming of the current knowledge about the Unique Games Conjecture (compared to many other conjectures in complexity theory) is that only few consequences of a refutation of the conjecture are known. For example, the following scenario is not ruled out: There is a polynomial time algorithm refuting the Unique Games Conjecture and, at same time, no polynomial time algorithm — or even subexponential time algorithm — has a better approximation ratio for Max Cut than the Goemans–Williamson algorithm. (In contrast, a refutation of the conjecture that 3-Sat has no polynomial time algorithms implies that *every* problem in NP has a polynomial time algorithm. Similarly, a refutation of the Exponential Time Hypothesis [IPZ01] — 3-Sat has no $2^{o(n)}$ time algorithm — implies surprising algorithms for a host of other problems.) The lack of surprising consequences of a refutation of the UGC is one of main reasons why no consensus regarding the truth of the conjecture has been reached among researchers (whereas many other complexity conjectures have a strong consensus among researchers).

Prior to this work, the only known non-trivial consequence of an algorithmic refutation of the Unique Games Conjecture is an improved approximation algorithm for Unique Games itself. Rao [Rao08] showed that if there exists a polynomial time algorithm refuting the Unique Games Conjecture, then there exists an algorithm that given a unique game with value $1-\varepsilon$, finds in polynomial time an assignment of value $1-C\sqrt{\varepsilon}$. (The constant $C \geqslant 1$ and the degree of the polynomial for the running time of this algorithm depend on the guarantees of the algorithm assumed to refute the UGC.)

In this work, we demonstrate the first non-trivial consequence of an algorithmic refutation of the Unique Games Conjecture for a problem different that Unique Games. This problem, called Small-Set Expansion, is a natural generalization of Sparsest Cut, one of the most fundamental optimization problems on graphs (see e.g. [LR99, ARV09]). We give a reduction from Small-Set Expansion to Unique Games. An alternative formulation of our result: If Small-Set Expansion is NP-hard to approximate (in a certain range of parameters), then the Unique Games Conjecture holds true.

## 6.1. Main Results

To state our results we introduce the following hypothesis about the approximability of Small-Set Expansion (restated from §1.4).

**Hypothesis 6.1** (Small Set-Expansion (SSE) Hypothesis). *For every $\varepsilon > 0$, there exists $\delta > 0$ such that the following promise problem is* NP*-hard: Given a graph $G$, distinguish between the cases,*

***YES:*** *some vertex set of volume $\delta$ has expansion at most $\varepsilon$.*

***NO:*** *no vertex set of volume $\delta$ has expansion less than $1-\varepsilon$.*

We remark the best known polynomial-time approximation algorithms for Small-Set Expansion [RST10a] (based on a basic SDP relaxation) fail to refute this conjecture. The conjecture also holds in certain hierarchies of relaxations (see Chapter 8).

In this work, we show relations of this hypothesis to the approximability of Unique Games. First, we show that the Unique Games Conjecture is true if the Small-Set Expansion Hypothesis holds.

**Theorem 6.2.** *The Small-Set Expansion Hypothesis implies the Unique Games Conjecture.*

The proof of this theorem is based on a reduction from Small-Set Expansion to Unique Games (the composition of Reduction 6.5 and Reduction 6.14). We prove the theorem at the end of this section.

Second, we show that the Small-Set Expansion Hypothesis is true if the following stronger variant of the Unique Games Conjecture holds. (The results in the following chapter (Chapter 7) imply that this variant of the Unique Games Conjecture is in fact equivalent to the Small-Set Expansion Hypothesis.)

**Hypothesis 6.3** (Unique Games Conjecture on Small-Set Expanders). *For every $\eta > 0$, there exists $\delta_\eta > 0$ such that for every $\zeta > 0$, the following promise problem is* NP*-hard for some $R = R_{\eta,\zeta}$: Given a unique game $\mathfrak{U}$ with alphabet size $R$, distinguish between the cases,*

***YES:*** *the optimal value of* $\mathfrak{U}$ *satisfies* $\mathrm{opt}(\mathfrak{U}) \geqslant 1-\eta$,

***NO:*** *the optimal value of* $\mathfrak{U}$ *satisfies* $\mathrm{opt}(\mathfrak{U}) \leqslant \zeta$ *and every vertex set of volume* $\delta_\eta$ *in the constraint of* $\mathfrak{U}$ *has expansion at least* $1-\eta$.

We remark that this variant of the Unique Games Conjecture holds in certain hierarchies of SDP relaxation (using results from Chapter 7 and Chapter 8).

The following theorem is based on a reduction from Unique Games to Small-Set Expansion (Reduction 6.20 in §6.4). It follows directly from Theorem 6.21 in §6.4.

**Theorem 6.4.** *Hypothesis 6.3 (Unique Games Conjecture on Small-Set Expanders) implies the Small-Set Expansion Hypothesis.*

### 6.1.1. Proof of Theorem 6.2

Our reduction from Small-Set Expansion to Unique Games (used to prove Theorem 6.2) naturally decomposes into two parts. First, we reduce Small-Set Expansion to Partial Unique Games (partial assignments are allowed). See Reduction 6.5 and Theorem 6.6 in §6.2. Then, we show how to reduce Partial Unique Games to Unique Games. See Reduction 6.14 and Theorem 6.15 in §6.3. Theorem 6.2 follows by instantiating Theorem 6.6 and Theorem 6.15 with an appropriate choice of parameters.

## 6.2. From Small-Set Expansion to Partial Unique Games

**Reduction 6.5** (From Small-Set Expansion to Partial Unique Games)**.**

***Input:*** A regular graph $G$ with vertex set $V$ and parameters $\varepsilon > 0$ and $R \in \mathbb{N}$ (satisfying $\varepsilon R \in \mathbb{N}$).

***Output:*** A unique game $\mathfrak{U} = \mathfrak{U}_{R,\varepsilon}(G)$ with vertex set $V^{R'}$ and alphabet $\Sigma = [R']$ for $R' = (1+\varepsilon)R$.

The unique game $\mathfrak{U}$ corresponds to the following probabilistic verifier for an assignment $F\colon V^{R'} \to [R']$:

1. Sample $R$ random vertices $a_1, \ldots, a_R \sim G$.

2. Sample two random neighbors $b_i, b'_i \sim G(a_i)$ for every $i \in [R]$. (Here, the notation $b \sim G(a)$ means that $b$ is a random neighbor of $a$ in $G$.)

3. Sample $2\varepsilon R$ random vertices $b_{R+1}, b'_{R+1}, \ldots, b_{R+\varepsilon R}, b'_{R+\varepsilon R} \sim V$.

4. Let $A = (a_1, \ldots, a_R) \in V^R$ and $B = (b_1, \ldots, b_{R'})$, $B' = (b_1, \ldots, b'_{R'}) \in V^{R'}$.

5. Sample two random permutation $\pi, \pi' \in \mathbb{S}_{[R']}$.

6. Verify that $\pi^{-1}(F(\pi.B)) = (\pi')^{-1}F(\pi'.B')$. (Here, $\pi.B$ refers to the tuple obtained by permuting the coordinates of $B$ according to the permutation $\pi$. See for a formal definition.) (End of Reduction 6.5)

Reduction 6.5 has the following approximation guarantees.

**Theorem 6.6.** *Given a regular graph $G$ with $n$ vertices and parameters $R \in \mathbb{N}$ and $\varepsilon > 0$, Reduction 6.5 computes in time $\mathrm{poly}(n^R)$ a unique game $\mathcal{U} = \mathcal{U}_{R,\varepsilon}(G)$ such that the following assertions hold (for all $\varepsilon' > \varepsilon$):*

**Completeness:** *If the graph $G$ contains a vertex set with volume $\delta = 1/2R$ and expansion at most $\varepsilon$, then the unique game $\mathcal{U}$ has $\alpha$-partial value $\mathrm{opt}_\alpha(\mathcal{U}) \geqslant 1 - 5\varepsilon$, where $\alpha \geqslant 0.1$.*

**Soundness I:** *If every vertex set of $G$ with volume $\delta = 1/2R$ has expansion at least $1-\varepsilon$, then for all $\alpha \geqslant 0.1$, the unique game $\mathcal{U}$ has $\alpha$-partial value $\mathrm{opt}_\alpha(\mathcal{U}) \leqslant O(\varepsilon^{1/5})$.*

**Soundness II:** *If every vertex set of $G$ with volume between $\Omega(\varepsilon^2/R)$ and $O(1/\varepsilon^2 R)$ has expansion at least $\varepsilon'$ and half of the edges of every vertex of $G$ are self-loops, then for all $\alpha \geqslant 0.1$, the unique game $\mathcal{U}$ has $\alpha$-partial value at most $\mathrm{opt}_\alpha(\mathcal{U}) \leqslant 1 - \varepsilon'/4$.*

**Remark 6.7.** All properties of the reduction would essentially be preserved if we would modify step 3 of the reduction as follows: For every $i \in [R]$, with probability $\varepsilon$, replace $b_i$ by a random vertex in $V$ (do the same with $b'_1, \ldots, b'_R$, independently). In step 4, one would set $B = (b_1, \ldots, b_R)$ and $B' = (b_1, \ldots, b_R)$.

### 6.2.1. Completeness

Let $G$ be a regular graph with vertex set $V = \{1, \ldots, n\}$. Note that the size of the unique game $\mathcal{U} = \mathcal{U}_{\varepsilon,R}(G)$ produced by Reduction 6.5 is $n^{O(R)}$ and thus polynomial in the size of the graph $G$ for every $R = O(1)$. (Also the running time of the reduction is $n^{O(R)}$.)

The following lemma shows that Reduction 6.5 is complete, that is, if the graph $G$ contains a set with a certain volume and expansion close to 0, then the unique game obtained from Reduction 6.5 has a partial assignment that satisfies almost all constraints with labeled vertices.

**Lemma 6.8** (Completeness). *For every set $S \subseteq V$ with $\mu(S) = \delta$ and $\Phi(S) = \eta$, there exists a partial assignment $F = F_S$ for the unique game $\mathcal{U} = \mathcal{U}_{R,\varepsilon}(G)$ (as defined in Reduction 6.5) satisfying*

$$\mathcal{U}(F) \geqslant (1 - \varepsilon - 4\eta)\alpha\,,$$

*where $\alpha \geqslant (1 - R'\delta)R'\delta$ is the fraction of vertices of $\mathcal{U}$ labeled by the partial assignment $x$.*

*Proof.* We may assume that $R' \leqslant 1/\delta$, because the lemma is trivial otherwise. Consider the following partial assignment $F$ for the unique game $\mathcal{U}$,

$$F\colon V^{R'} \to [R'] \cup \{\bot\},$$
$$X = (x_1, \ldots, x_{R'}) \mapsto \begin{cases} i & \text{if } \{i\} = \{i' \in [R'] \mid x_i \in S\}, \\ \bot & \text{otherwise.} \end{cases}$$

We compute the fraction of vertices of $\mathcal{U}$ labeled by $F$,

$$\alpha = \mathop{\mathbb{P}}_{X \sim V^{R'}} \left\{F(X) \neq \bot\right\} = \binom{R'}{1}(1-\delta)^{R'-1}\delta = ((1-\delta)^{R'-1})\,R'\delta \geqslant (1 - R'\delta)R'\delta\,,$$

Next, we estimate the fraction of constraints satisfied by $F$ in $\mathcal{U}$. Sample tuples of vertices $A$, $B$, $B'$ as specified by Reduction 6.5. Note that the assignment $F$ behaves nicely when permuting coordinates, $F(\pi.X) = \pi(F(X))$ for every permutation $\pi$ of $[R']$ and every $X \in V^{R'}$ such that $F(X) \neq \bot$. Hence,

$$\mathcal{U}(F) = \mathop{\mathbb{P}}_{A,B,B'} \{F(B) = F(B') \in [R']\}\,.$$

(Here, we also use that a constraint is considered to be satisfied by a partial assignment only if both vertices of the constraint are labeled.) We bound the fraction of satisfied constraints from below

$$\begin{aligned} \mathop{\mathbb{P}}_{A,B,B'} \{F(B) = F(B') \in [R']\} &\geqslant \mathop{\mathbb{P}}_{A,B,B'} \{F(B) = F(B') \in [R]\} \\ &= \mathop{\mathbb{P}}_{X \sim V^{R'}} \{F(X) \in [R]\} \mathop{\mathbb{P}}_{A,B,B'} \{F(B') = F(B) \mid F(B) \in [R]\}\,. \end{aligned}$$

The inequality is almost tight, because the event $F(X) \in [R'] \setminus [R]$ is very unlikely (even relative to the event $F(X) \in [R']$),

$$\mathop{\mathbb{P}}_{X\sim V^{R'}} \{F(X) \in [R]\} = \mathop{\mathbb{P}}_{X\sim V^{R'}} \{F(X) \in [R']\}/(1+\varepsilon) \geqslant (1-\varepsilon)\alpha\,.$$

Next, we relate the probability of the event $F(B) = F(B')$ conditioned on $F(B) \in [R]$ to the expansion of the set $S$. It turns out that this probability is most directly related to the expansion of $S$ in the graph $G^2$ (instead of $G$). Let $\eta' = \mathbb{P}_{a\sim V,\ b,b'\sim G(a)} \{b' \notin S \mid b \in S\}$ be the expansion of $S$ in $G^2$. Then,

$$\begin{aligned}\mathop{\mathbb{P}}_{A,B,B'} \{F(B') = F(B) \mid F(B) \in [R]\} &= \mathop{\mathbb{P}}_{A,B,B'} \{F(B') = 1 \mid F(B) = 1\} \\ &= (1-\eta')(1-\eta'\delta/(1-\delta))^{R-1} \\ &\geqslant 1 - \left(1 + \tfrac{1-1/R}{1-\delta} R\delta\right)\eta' \geqslant 1 - 2\eta'\,.\end{aligned}$$

Here, we again use the symmetry of $F$ and we also use that $\mathbb{P}_{a\sim V,\ b,b'\sim G(a)} \{b' \in S \mid b \notin S\} = \eta'\delta/(1-\delta)$ (the expansion of $V \setminus S$ in $G^2$). The last inequality uses that $R' \leqslant 1/\delta$ and therefore $1 - 1/R \leqslant 1-\delta$ and $R\delta \leqslant 1$.

It remains to relate $\eta'$ (the expansion of $S$ in $G^2$) to $\eta$ (the expansion of $S$ in $G$).

$$\begin{aligned}\eta'\delta &= \mathop{\mathbb{P}}_{a\sim V,\ b,b'\sim G(a)} \{b \in S \wedge b' \notin S\} \\ &\leqslant \mathop{\mathbb{P}}_{a\sim V,\ b,b'\sim G(a)} \{(b \in S \wedge a \notin S) \vee (a \in S \wedge b' \notin S)\} \\ &\leqslant \mathop{\mathbb{P}}_{a\sim V,\ b\sim G(a)} \{b \in S \wedge a \notin S\} + \mathop{\mathbb{P}}_{a\sim V,\ b'\sim G(a)} \{a \in S \wedge b' \notin S\} = 2\delta\eta\,.\end{aligned}$$

Combining the previous bounds shows the desired lower bound on the fraction of constraints satisfied by $F$ in $\mathfrak{U}$,

$$\begin{aligned}\mathfrak{U}(F) &= \mathop{\mathbb{P}}_{A,B,B'} \{F(B) = F(B') \in [R']\} \\ &\geqslant \mathop{\mathbb{P}}_{X\sim V^{R'}} \{F(X) \in [R]\} \mathop{\mathbb{P}}_{A,B,B'} \{F(B') = 1 \mid F(B) = 1\} \\ &\geqslant (1-\varepsilon)\alpha \cdot (1-4\eta) \geqslant (1-\varepsilon-4\eta)\alpha\,.\end{aligned}$$

□

### 6.2.2. Soundness

Let $F\colon V^{R'} \to [R'] \cup \{\perp\}$ be a partial assignment for the unique game $\mathfrak{U} = \mathfrak{U}_{\mathbb{R},\varepsilon}(G)$ obtained by applying Reduction 6.5 to a regular graph $G$ with vertex set $V = \{1,\ldots,n\}$.

For a tuple $U \in V^{R'-1}$ and a vertex $x \in V$, let $f(U,x)$ be the probability that $F$ selects the coordinate of $x$ after we place it a random position of $U$ and permute the tuple randomly,

$$f(U,x) \stackrel{\text{def}}{=} \mathop{\mathbb{P}}_{i \in [R'],\ \pi \in \mathbb{S}_{R'}} \left\{ F(\pi.(U +_i x)) = \pi(i) \right\}.$$

Here, $U +_i x$ denotes the tuple obtained from $U$ by inserting $x$ as the $i$-th coordinate (and moving the original coordinates $i, \ldots, R'-1$ of $U$ by one to the right). (The above experiment wouldn't change if we fixed $i = R'$, because the permutation $\pi$ is random.)

For $U \in V^{R-1}$, define the function $f_U \colon V \to [0,1]$,

$$f_U(x) \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{W \sim G^{\otimes(R-1)}(U),\ Z \in V^{\varepsilon R}} f(W,Z,x). \tag{6.1}$$

(Here, $G^{\otimes(R-1)}$ refers to the $(R-1)$-fold tensor product of the graph $G$, and $W \sim G^{\otimes(R-1)}(U)$ means that $W$ is a random neighbor of the vertex $U$ in $G^{\otimes(R-1)}$. Note that for $U = (u_1, \ldots, u_{R-1})$, the distribution $G^{\otimes(R-1)}(U)$ is the product of the distributions $G(u_1), \ldots, G(u_{R-1})$.)

The following properties of the functions $\{f_U\}$ are straightforward to verify.

**Lemma 6.9.** *Let $\alpha = \mathbb{P}_{X \sim V^{R'}} \{F(X) \neq \perp\}$ be the fraction of vertices of $\mathfrak{U}$ labeled by the partial assignment $F$.*

1. *The typical $L_1$-norm of $f_U$ equals $\mathbb{E}_{U \sim V^{R-1}} \|f_U\|_1 = \frac{\alpha}{R'}$.*
2. *For every $U \in V^{R-1}$, the $L_1$-norm of $f_U$ satisfies $\|f_U\|_1 \leqslant \frac{1}{\varepsilon R}$.*
3. *The typical squared $L_2$-norm of $Gf_U$ relates to the fraction of constraints satisfied by $F$ in $\mathfrak{U}$ as follows,*

$$\mathop{\mathbb{E}}_{U \sim V^{R-1}} \|Gf_U\|^2 \geqslant \tfrac{1}{R'}\left(\mathfrak{U}(F) - \tfrac{1}{\varepsilon R}\right).$$

*(Here, we identify the regular graph $G$ with its stochastic adjacency matrix.)*

*Proof.* *Item 1*: The typical $L_1$-norm of $f_U$ evaluates to

$$\mathop{\mathbb{E}}_{U \sim V^R} \|f_U\|_1 = \mathop{\mathbb{P}}_{\substack{U \sim V^R,\ W \sim G^{\otimes(R-1)}(U),\ Z \sim V^{\varepsilon R},\ x \sim V,\ \pi \in \mathbb{S}_{R'} \\ i \in \{1,\ldots,R'\}}} \left\{ F\Big(\pi.\big((W,Z) +_i x\big)\Big) = \pi(i) \right\}$$

$$= \mathop{\mathbb{P}}_{\substack{X \sim V^{R'},\ \pi \in \mathbb{S}_{R'} \\ i \in \{1,\ldots,R'\}}} \left\{ F(\pi.X) = \pi(i) \right\} = \frac{1}{R'} \mathop{\mathbb{P}}_{X \sim V^{R'}} \{F(X) \neq \perp\}.$$

The second step uses that the joint distribution $(W,Z) +_i x$ and $i$ is the same as the joint distribution of $X$ and $i$. The last step uses that the distribution of $\pi(i)$ is uniformly random in $[R']$ even after conditioning on $X$ and $\pi$ (and thus $F(\pi.X)$).

*Item 2:* For fixed $U \in V^{R-1}$, the $L_1$-norm of $f_U$ evaluates to

$$\begin{aligned}
\|f_U\|_1 &= \mathop{\mathbb{P}}_{\substack{W\sim G^{\otimes(R-1)}(U),\ Z\sim V^{\varepsilon R},\ x\sim V,\ \pi\in\mathbb{S}_{R'} \\ i\in\{R,R+1,\ldots,R'\}}} \left\{ F\Big(\pi.\big((W,Z) +_i x\big)\Big) = \pi(i) \right\} \\
&= \mathop{\mathbb{P}}_{\substack{W\sim G^{\otimes(R-1)}(U),\ Z'\sim V^{\varepsilon R+1},\ \pi\in\mathbb{S}_{R'} \\ i\in\{R,R+1,\ldots,R'\}}} \left\{ F\Big(\pi.\big(W,Z'\big)\Big) = \pi(i) \right\} \\
&= \tfrac{1}{\varepsilon R+1} \mathop{\mathbb{P}}_{W\sim G^{\otimes(R-1)}(U),\ Z'\sim V^{\varepsilon R+1},\ \pi\in\mathbb{S}_{R'}} \left\{ \pi^{-1}\Big(F\big(\pi.(W,Z')\big)\Big) \in \{R,R+1,\ldots,R'\} \right\} \\
&\leqslant \tfrac{1}{\varepsilon R+1}\,.
\end{aligned}$$

In contrast to the proof of item 1, we insert $x$ in a random coordinate among $\{R,R+1,\ldots,R'\}$ (as opposed to completely random coordinate). The experiment as a whole does not change because $\pi$ is a random permutation. The second step uses that $(i,\ (U,Z) +_i x)$ has the same distribution as $(i,\ (U,Z'))$.

*Item 3:* Sample tuples $A,B,B'$ as specified by Reduction 6.5. Note that $B$ and $B'$ are distributed independent and identically conditioned on $A$. We denote this conditional distribution by $B \mid A$. The fraction of constraints satisfied by $F$ in $\mathfrak{U}$ evaluates to

$$\begin{aligned}
\mathfrak{U}(F) &= \sum_{r\in[R']} \mathop{\mathbb{P}}_{\substack{A,B,B' \\ \pi,\pi'\in\mathbb{S}_{R'}}} \left\{ F(\pi.B) = \pi(r) \ \wedge F(\pi'.B') = \pi'(r) \right\} \\
&= \sum_{r\in[R']} \mathop{\mathbb{E}}_{A\sim V^R} \left( \mathop{\mathbb{P}}_{\substack{B|A \\ \pi\in\mathbb{S}_{R'}}} \left\{ F(\pi.B) = \pi(r) \right\} \right)^2 .
\end{aligned}$$

For every $A \in V^R$ and any $r,r' \in \{R+1,\ldots,R'\}$, it holds that

$$\mathop{\mathbb{P}}_{\substack{B|A, \\ \pi\in\mathbb{S}_{R'}}} \left\{ F(\pi.B) = \pi(r) \right\} = \mathop{\mathbb{P}}_{\substack{B|A, \\ \pi\in\mathbb{S}_{R'}}} \left\{ F(\pi.B) = \pi(r') \right\} .$$

Here, we use that all coordinates $b_r$ of $B$ with $r \in \{R+1,\ldots,R'\}$ are distributed identically (even conditioned on $A$). It follows that for every $r \in \{R+1,\ldots,R'\}$,

$$\mathop{\mathbb{P}}_{\substack{B|A \\ \pi\in\mathbb{S}_{R'}}} \left\{ F(\pi.B) = \pi(r) \right\} \leqslant \tfrac{1}{\varepsilon R} .$$

Next, consider $r \in [R]$. For $B = (b_1, \dots, b_R) \in V^{R'}$, let $B^{-r} \in V^{R'-1}$ denote the tuple obtained from $B$ by removing the $r^{\text{th}}$ coordinate $b_r$. Then,

$$\underset{\substack{B|A,\\ \pi \in \mathbb{S}_{R'}}}{\mathbb{P}} \left\{ F(\pi.B) = \pi(r) \right\} = \underset{\substack{B|A,\\ \pi \in \mathbb{S}_{R'}}}{\mathbb{P}} \left\{ F(\pi.(B^{-r} +_r b_r)) = \pi(r) \right\} = \underset{B|A}{\mathbb{E}} f(B^{-r}, b_r).$$

Since $r \in [R]$, the vertex $b_r$ is distributed as a random neighbor of $a_r$ (the corresponding coordinate in $A$). It follows that

$$\underset{B|A}{\mathbb{E}} f(B^{-r}, b_r) = \underset{b_r \sim G(a_r)}{\mathbb{E}} f_{A^{-r}}(b_r) = G f_{A^{-r}}(a_r).$$

Recall that we identify the graph $G$ with its stochastic adjacency matrix. Hence, $G f_{A^{r-1}}$ denotes the function on $V$ obtained by applying the linear operator $G$ to the function $f_{A^{-r}}$. (Here, $A^{-r} \in V^{R-1}$ is the tuple obtained from $A$ by removing the $r^{\text{th}}$ coordinate $a_r$.) Combining the previous bounds, we get

$$\mathfrak{U}(F) \leqslant \sum_{r \in [R]} \underset{A \sim V^R}{\mathbb{E}} \left( G f_{A^{-r}}(a_r) \right)^2 + \varepsilon R \cdot \left( \tfrac{1}{\varepsilon R} \right)^2 = R \underset{U \sim V^{R-1}}{\mathbb{E}} \|G f_U\|^2 + \tfrac{1}{\varepsilon R},$$

which implies the desired bound on the typical squared $L_2$-norm of $G f_U$. □

The following lemma is a consequence of the previous lemma (Lemma 6.9) and a simple Markov-type inequality. The lemma shows that given a good partial assignment for unique game $\mathfrak{U} = \mathfrak{U}_{R,\varepsilon}(G)$ (as specified in Reduction 6.5), one can extract a function $f : V \to [0,1]$ such that $\|f\|_1 \approx 1/R$ and, at the same time, the squared $L_2$-norm of $Gf$ is comparable to the $L_1$-norm of $f$. In a later lemma (Lemma 6.11), we will see that given such a function we can find a non-expanding set for the graph $G^2$ with volume roughly $1/R$.

**Lemma 6.10.** *Let $F$ be a partial assignment for the unique game $\mathfrak{U} = \mathfrak{U}_{R,\varepsilon}(G)$ (as specified in Reduction 6.5). Suppose $\alpha = \mathbb{P}_{X \sim V^{R'}} \{F(X) \neq \bot\}$ is the fraction of vertices of $\mathfrak{U}$ labeled by $F$. Then, for every $\beta > 0$, there exists $U \in V^{R-1}$ such that the function $f_U : V \to [0,1]$ (as defined in (6.1)) satisfies*

$$\|G f_U\|^2 \geqslant (\mathfrak{U}(F)/\alpha - \beta - \tfrac{1}{\alpha \varepsilon R}) \|f_U\|_1,$$
$$\tfrac{\alpha\beta}{R'} \leqslant \|f_U\|_1 \leqslant \tfrac{1}{\varepsilon R}.$$

*Proof.* Since $\|G f_U\|^2 \leqslant \|f_U\|_1$ for every $U \in V^{R-1}$ (using that $G$ is regular and $0 \leqslant f_U \leqslant 1$), we can lower bound the expected square $L_2$-norm of $G f_U$

conditioned on $\|f_U\|_1 \geqslant \alpha\beta/R'$,

$$\begin{aligned}
\mathop{\mathbb{E}}_{U\sim V^{R-1}} \|Gf_U\|^2 \mathbb{1}_{\{\|f_U\|_1 \geqslant \alpha\beta/R'\}} &= \mathop{\mathbb{E}}_{U\sim V^{R-1}} \|Gf_U\|^2 - \mathop{\mathbb{E}}_{U\sim V^{R-1}} \|Gf_U\|^2 \mathbb{1}_{\|f_U\|_1 \leqslant \alpha\beta/R'} \\
&\geqslant \mathop{\mathbb{E}}_{U\sim V^{R-1}} \|Gf_U\|^2 - \alpha\beta/R' \\
&\geqslant \tfrac{\alpha}{R'} \Big( \mathfrak{U}(F)/\alpha - \tfrac{1}{\alpha\varepsilon R} - \beta \Big) .
\end{aligned}$$

In the last step, we use that $\mathbb{E}_{U\sim V^{R-1}} \|Gf_U\|^2 \geqslant (\mathfrak{U}(F) - \frac{1}{\varepsilon R})/R'$ (Lemma 6.9, item 3). Since $\mathbb{E}_{U\sim V^{R-1}} \|f_U\|_1 = \alpha/R'$ (Lemma 6.9, item 1), there exists a tuple $U \in V^{R-1}$ such that

$$\|Gf_U\|^2 \mathbb{1}_{\{\|f_U\|_1 \geqslant \alpha\beta/R'\}} \geqslant \Big( \mathfrak{U}(F)/\alpha - \tfrac{1}{\alpha\varepsilon R} - \beta \Big) \|f_U\|_1 .$$

This function $f_U$ satisfies both $\|f_U\|_1 \geqslant \alpha\beta/R'$ and $\|Gf_U\|^2 \geqslant (\mathfrak{U}(F)/\alpha - \beta - \frac{1}{\alpha\varepsilon R})\|f_U\|_1$. (Note that we may assume $1 - \eta - \beta - \frac{1}{\alpha\varepsilon R}$ is nonnegative, because otherwise the lemma is trivial.) On the other hand, $f_U$ also satisfies $\|f_U\|_1 \leqslant 1/\varepsilon R$ (by Lemma 6.9, item 2). □

The spirit of the following lemma is similar to Cheeger's inequality. Given a function $f$ on the vertex set $V$ with values between 0 and 1, we can find a vertex set $S$ with volume roughly $\|f\|_1$ and expansion roughly $1 - \|Gf\|^2/\|f\|_1$ in the graph $G^2$. The proof of the lemma is much simpler than the construction for Cheeger's inequality. It is enough to analyze the distribution over vertex sets $S$ obtained by including $x$ in $S$ with probability $f(x)$ independently for every vertex $x \in V$. Later, we will apply this lemma to the function obtained by the previous lemma (Lemma 6.10).

**Lemma 6.11.** *Suppose $f\colon V \to \mathbb{R}$ satisfies $0 \leqslant f(x) \leqslant 1$ for every vertex $x \in V$. Then, for every $\beta > 0$, there exists a set $S \subseteq V$ such that*

$$\begin{aligned}
\beta\|f\|_1 \leqslant \mu(S) \leqslant \tfrac{1}{\beta}\|f\|_1 , \\
\Phi_{G^2}(S) \leqslant 1 - \tfrac{\|Gf\|^2}{\|f\|_1} + 2\beta + \beta/(n\|f\|_1).
\end{aligned}$$

*Proof.* Consider the following distribution over level sets $S \subseteq V$ of $f$: For every vertex $x \in V$, include $x$ in $S$ with probability $f(x)$ (independently for every vertex).

The expected volume of $S$ is $\mathbb{E}_S\, \mu(S) = \mathbb{E}_S \|\mathbb{1}_S\|_1 = \|f\|_1$. The expectation of the square of the volume satisfies $\mathbb{E}_S\, \mu(S)^2 \leqslant \|f\|_1^2 + 1/n$. On the other hand, we can lower bound typical squared $L_2$-norm of $G\mathbb{1}_S$ by $\mathbb{E}_S \|G\mathbb{1}_S\|^2 \geqslant \|Gf\|^2$.

Next, we lower bound the expected squared $L_2$-norm of $G\mathbb{1}_S$ conditioned on the event $\beta \leqslant \mu(S)/\|f\|_1 \leqslant 1/\beta$,

$$\begin{aligned}
\mathop{\mathbb{E}}_S \|G\mathbb{1}_S\|^2 \mathbb{1}_{\{\beta\leqslant\mu(S)/\|f\|_1\leqslant 1/\beta\}} &\geqslant \mathop{\mathbb{E}}_S \|G\mathbb{1}_S\|^2 - \mathop{\mathbb{E}}_S \mu(S)\mathbb{1}_{\mu(S)>\|f\|_1/\beta} - \mathop{\mathbb{E}}_S \mu(S)\mathbb{1}_{\mu(S)<\beta\|f\|_1} \\
&\geqslant \|Gf\|^2 - \mathop{\mathbb{E}}_S \beta\mu(S)^2/\|f\|_1 - \beta\|f\|_1 \\
&\geqslant \|Gf\|^2 - 2\beta\|f\|_1 - \beta/(n\|f\|_1).
\end{aligned}$$

It follows that there exists a set $S^* \subseteq V$ with $\beta \leqslant \mu(S^*)/\|f\|_1 \leqslant 1/\beta$ such that

$$\frac{\|G\mathbb{1}_{S^*}\|^2}{\|\mathbb{1}_{S^*}\|_1} \geqslant \frac{\mathbb{E}_S \|G\mathbb{1}_S\|^2 \mathbb{1}_{\{\beta\leqslant\mu(S)/\|f\|_1\leqslant 1/\beta\}}}{\mathbb{E}_S \|\mathbb{1}_S\|_1} \geqslant \frac{\|Gf\|^2 - 2\beta\|f\|_1 - \beta/(n\|f\|_1)}{\|f\|_1}.$$

The quantity $1 - \|G\mathbb{1}_{S^*}\|^2/\|\mathbb{1}_{S^*}\|_1$ is the expansion of $S^*$ in the graph $G^2$. □

If we combine the two previous lemmas (Lemma 6.10 and Lemma 6.11), we can show that a good partial assignment for the unique game $\mathfrak{U} = \mathfrak{U}_{R,\varepsilon}(G)$ (obtained by applying Reduction 6.5 to the graph $G$) implies that the graph $G^2$ contains a set with low expansion and volume roughly $1/R$.

Eventually, we want a vertex set with low expansion in the graph $G$. The following lemmas describe how to reduce this problem to the problem of finding vertex sets with low expansion in $G^2$.

We first consider the case that there exists a partial assignment for the unique game $\mathfrak{U}$ that satisfies a fraction of constraints bounded away from 0. In this case, we show how to construct a set with volume $1/R$ and expansion bounded away from 1 in $G$.

**Lemma 6.12** (Soundness close to 0). *Suppose there exists a partial assignment $F$ for the unique game $\mathfrak{U} = \mathfrak{U}_{R,\varepsilon}(G)$ (as defined in Reduction 6.5) with $\mathfrak{U}(F) \geqslant \eta\alpha$, where $\alpha$ is the fraction of vertices of $\mathfrak{U}$ labeled by $F$. Then, there exists a set $S^* \subseteq V$ with $\mu(S^*) = \alpha/R$ and $\Phi(S^*) \leqslant 1 - O(\varepsilon\eta^4)$. Here, we make the (mild) assumptions that $1/\alpha\varepsilon\eta \ll R \ll \alpha\eta n$.*

*Proof.* Let $\beta > 0$. (We determine the best choice for $\beta$ later in the proof.) By Lemma 6.10, there exists $U \in V^{R-1}$ such that the function $f_U$ (as defined in (6.1)) satisfies the following conditions

$$\|Gf_U\|^2 \geqslant (\eta - \beta - \tfrac{1}{\alpha\varepsilon R})\|f_U\|_1\,,$$
$$\tfrac{\alpha\beta}{R'} \leqslant \|f_U\|_1 \leqslant \tfrac{1}{\varepsilon R}\,.$$

Using Lemma 6.11 we round the function $f_U$ to a vertex set $S \subseteq V$ with the following properties,

$$\tfrac{\alpha\beta^2}{R'} \leqslant \mu(S) \leqslant \tfrac{1}{\beta\varepsilon R},$$
$$\Phi_{G^2}(S) \leqslant 1 - \eta + \beta + \tfrac{1}{\alpha\varepsilon R} + 2\beta + \tfrac{R'}{n\alpha}.$$

If we choose $\beta$ significantly smaller than $\eta$ (say $\beta = \eta/10$) and use our assumptions on $R$, the expansion of $S$ in $G^2$ is at most $1 - \eta/2$.

Next, we want to construct a set with expansion bounded away from 1 in the graph $G$ and with roughly the same volume as $S$. Consider the set $S'$ of all vertices $x$ with $\mathbb{P}_{y\sim G(x)}\{y \in S\} \geqslant \gamma$. (We determine a good choice for $\gamma > 0$ later.) The volume of $S'$ cannot be much larger than the volume of $S$. Concretely, $\mu(S') \leqslant \mu(S)/\gamma$. On the other hand, we can relate the fraction of edges between $S$ and $S'$ to the expansion of $S$ in $G^2$,

$$\begin{aligned}
(1 - \Phi_{G^2}(S))\mu(S) &= \mathop{\mathbb{P}}_{x\sim V,\ y,y'\sim G(x)} \{y \in S,\ y' \in S\} \\
&= \mu(S') \mathop{\mathbb{P}}_{x\sim V,\ y,y'\sim G(x)} \{y \in S,\ y' \in S \mid x \in S'\} \\
&\quad + \mu(V \setminus S') \mathop{\mathbb{P}}_{x\sim V,\ y,y'\sim G(x)} \{y \in S,\ y' \in S \mid x \notin S'\} \\
&\leqslant \mu(S') \mathop{\mathbb{P}}_{x\sim V,\ y\sim G(x)} \{y \in S \mid x \in S'\} \\
&\quad + \gamma \cdot \mu(V \setminus S') \mathop{\mathbb{P}}_{x\sim V,\ y\sim G(x)} \{y \in S \mid x \notin S'\} \\
&= (1-\gamma)\mu(S') \mathop{\mathbb{P}}_{x\sim V,\ y\sim G(x)} \{y \in S \mid x \in S'\} \\
&\quad + \gamma \cdot \mu(S).
\end{aligned}$$

It follows that $G(S, S') \geqslant (1 - \Phi_{G^2}(S) - \gamma)\mu(S)$. Therefore, for $S'' = S \cup S'$,

$$1 - \Phi_G(S'') \geqslant \frac{G(S,S')}{\mu(S \cup S')} \geqslant \left(\tfrac{\eta}{2} - \gamma\right) \cdot \frac{\mu(S)}{\mu(S \cup S')} \geqslant \left(\tfrac{\eta}{2} - \gamma\right) \cdot \tfrac{\gamma}{2}.$$

Choosing $\gamma = \eta/4$, we obtain $\Phi_G(S'') \leqslant 1 - \eta^2/32$. On the other hand, the volume of $S''$ satisfies

$$\Omega\left(\tfrac{\alpha\eta^2}{R}\right) \leqslant \mu(S'') \leqslant O\left(\tfrac{1}{\eta^2\varepsilon R}\right)$$

To obtain a set $S^*$ with the desired volume $\alpha/R$, we either pad the set $S''$ with the desired number of vertices or we take a random subset of $S''$ with the desired cardinality. In either case, we obtain a set $S^*$ with volume $\alpha/R$ and expansion at most $1 - O(\varepsilon\eta^4)$. □

In the next lemma, the goal is to find vertex sets in $G$ with certain volume and expansion close to 0. It turns out that in this case it is convenient to apply Reduction 6.5 not to $G$ itself but to the graph $\frac{1}{2}I + \frac{1}{2}G$ (obtained by adding a self-loop of weight half to every vertex). A random step in the graph $\frac{1}{2}I + \frac{1}{2}G$ stays at the same vertex with probability $1/2$ and goes to a random neighbor in $G$ with the remaining probability. The following lemma shows that if the unique game $\mathfrak{U} = \mathfrak{U}_{R,\varepsilon}(\frac{1}{2}I + \frac{1}{2}G)$ (obtained by applying Reduction 6.5 to the graph $\frac{1}{2} + \frac{1}{2}G$) has a partial assignment that satisfies almost as many constraints as possible, then we can find a vertex set in $G$ with volume roughly $1/R$ and expansion close to 0. The proof is similar to the proof of the previous lemma (Lemma 6.12).

**Lemma 6.13** (Soundness close to 1). *Suppose there exists a partial assignment $x$ for the unique game $\mathfrak{U} = \mathfrak{U}_{R,\varepsilon}(\frac{1}{2}I + \frac{1}{2}G)$ (as defined in Reduction 6.5) with $\mathfrak{U}(x) \geqslant (1-\eta)\alpha$, where $\alpha$ is the fraction of vertices of $\mathfrak{U}$ labeled by $x$. Then, there exists a set $S \subseteq V$ with $\Omega(\alpha\eta^2/R) \leqslant \mu(S) \leqslant O(1/\varepsilon\eta R)$ and $\Phi(S) \leqslant 4\eta$. Here, we make the (mild) assumptions that $1/\alpha\varepsilon\eta \ll R \ll \alpha\eta n$.*

*Proof.* Let $\beta > 0$. (We choose this parameter later in the proof.) Let $G_{\circlearrowleft} = \frac{1}{2}I + \frac{1}{2}G$. Combining Lemma 6.10 and Lemma 6.11, we obtain a set $S$ with the following properties,

$$\frac{\alpha\beta^2}{R'} \leqslant \mu(S) \leqslant \frac{1}{\beta\varepsilon R},$$
$$\Phi_{G_{\circlearrowleft}^2}(S) \leqslant \eta + \beta + \frac{1}{\alpha\varepsilon R} + 2\beta + \frac{R'}{n\alpha}.$$

If we choose $\beta$ significantly smaller than $\eta$ and use our assumptions on $R$, the expansion of $S$ in $G_{\circlearrowleft}^2$ is at most $2\eta$. We compare the expansion of $S$ in $G_{\circlearrowleft}^2$ to its expansion in $G$,

$$\langle \mathbb{1}_S, G_{\circlearrowleft}^2 \mathbb{1}_S\rangle = \tfrac{1}{4}\langle \mathbb{1}_S, I\mathbb{1}_S\rangle + \tfrac{1}{2}\langle \mathbb{1}_S, G\mathbb{1}_S\rangle + \tfrac{1}{4}\langle \mathbb{1}_S, G^2\mathbb{1}_S\rangle \leqslant \tfrac{1}{2}\|\mathbb{1}_S\|^2 + \tfrac{1}{2}\langle \mathbb{1}_S, G\mathbb{1}_S\rangle.$$

It follows that $\Phi_{G_{\circlearrowleft}^2}(S) \geqslant \Phi_G(S)/2$, as desired. □

## 6.3. From Partial Unique Games to Unique Games

**Reduction 6.14** (From Partial Unique Games to Unique Games).

***Input:*** A unique game $\mathfrak{U}$ with vertex set $V$ and alphabet $\Sigma$, and a parameter $c \in \mathbb{N}$.

**Output:** A unique game $\mathfrak{U}' = \Psi_c(\mathfrak{U})$ with vertex set $V' = V^c$ and alphabet $\Sigma' = [c] \times \Sigma$.

The unique game $\mathfrak{U}'$ corresponds to the following probabilistic verifier for an assignment $F\colon V' \to \Sigma'$:

1. Sample $c$ random vertices $u_1, \ldots, u_c \sim V$.
2. Sample two random constraints $(u_r, v_r, \pi_r), (u_r, v'_r, \pi'_r) \sim \mathfrak{U} \mid u_r$ for every $r \in [c]$. (Recall that $\mathfrak{U} \mid u_r$ denotes the uniform distribution over constraints of $\mathfrak{U}$ that contain vertex $u_r$.)
3. Let $(r, j) = F(v_1, \ldots, v_c)$ and $(r', j') = F(v'_1, \ldots, v'_c)$.
4. Verify that $r = r'$ and that $j = \pi_r(i)$ and $j' = \pi'_{r'}(i)$ for some label $i \in \Sigma$. (Note that there can be at most one label $i \in \Sigma$ satisfying this condition.) (End of Reduction 6.14)

Reduction 6.14 has the following approximation guarantees.

**Theorem 6.15.** *Given a parameter $c \in \mathbb{N}$ and a unique game $\mathfrak{U}$ with $n$ vertices, Reduction 6.14 computes in time $\operatorname{poly}(n^c)$ a unique game $\mathfrak{U}' = \Psi_c(\mathfrak{U})$ such that the following assertions hold (for all $\alpha, \eta, \eta', \zeta > 0$):*

**Completeness:** *If the unique game $\mathfrak{U}$ has $\alpha$-partial value at least $\operatorname{opt}_\alpha(\mathfrak{U}) \geqslant 1-\eta$, then the unique game $\mathfrak{U}'$ has value at least $\operatorname{opt}(\mathfrak{U}') \geqslant 1 - 4\eta - 2e^{-\alpha c}$.*

**Soundness I:** *If the unique game $\mathfrak{U}$ has $1/2c$-partial value less than $\operatorname{opt}_{1/2c}(\mathfrak{U}) < \zeta$, then the unique game $\mathfrak{U}'$ has value less than $\operatorname{opt}(\mathfrak{U}') < 8\zeta$.*

**Soundness II:** *If the unique game $\mathfrak{U}$ has $1/2c$-partial value less than $\operatorname{opt}_{1/2c}(\mathfrak{U}) < 1-\eta'$ and half of the constraints of every vertex are trivial identity constraints, then the unique game $\mathfrak{U}'$ has value less than $\operatorname{opt}(\mathfrak{U}') < 1 - \eta'/32$.*

The $\alpha$-partial value of a unique game is the maximum value of an $\alpha$-partial assignment normalized by the fraction of labeled vertices. See (2.1) in Section 2.3 for the formal definition. We restate (2.1) below (out of context).

$$\operatorname{opt}_\alpha(\mathfrak{U}) \stackrel{\text{def}}{=} \max\left\{ \frac{1}{\mathbb{P}_{u\sim V}\{x_u \neq \perp\}} \mathfrak{U}(x) \;\middle|\; x \in (\Sigma \cup \{\perp\})^V,\ \mathop{\mathbb{P}}_{(u,v,\pi)\sim\mathfrak{U}}\{x_u \neq \perp\} \geqslant \alpha \right\}.$$

### 6.3.1. Completeness

Let $\mathfrak{U}$ be a unique game with vertex set $V = \{1,\dots,n\}$ and alphabet $\Sigma = \{1,\dots,k\}$.

Recall that $\mathrm{opt}_\alpha(\mathfrak{U})$ is the optimal value of an $\alpha$-partial assignment for the unique game $\mathfrak{U}$. (See (2.1) in Section 2.3 for more details.)

The following lemma shows that Reduction 6.14 is complete, that is, given a unique game $\mathfrak{U}$ with $\mathrm{opt}_\alpha(\mathfrak{U}) \geqslant 1-\eta$ for some constant $\alpha$, then for $c = O(\log(1/\eta))$, the unique game $\mathfrak{U}' = \Psi_c(\mathfrak{U})$ obtained by applying Reduction 6.14 has value $\mathrm{opt}(\mathfrak{U}') \geqslant 1 - O(\eta)$.

**Lemma 6.16.** *If the unique game $\mathfrak{U}$ has $\alpha$-partial value at least $\mathrm{opt}_\alpha(\mathfrak{U}) \geqslant 1-\eta$, then the unique game $\mathfrak{U}'$ has value at least $\mathrm{opt}(\mathfrak{U}') \geqslant 1 - 4\eta - 2e^{-\alpha c}$.*

*Proof.* Let $f\colon V \to \Sigma$ be an optimal $\alpha$-partial assignment for $\mathfrak{U}$. We may assume $\mathbb{P}_{u\sim V}\{f(u) \neq \bot\} = \alpha$ and thus, $\mathfrak{U}(f) \geqslant (1-\eta)\alpha$. To lower bound the value of the unique game $\mathfrak{U}'$, we consider the following partial assignment $F\colon V^c \to [c]\times\Sigma\cup\{\bot\}$,

$$F(u_1,\dots,u_c) = \begin{cases} (r,i) & \text{if } f(u_1) = \ldots = f(u_{r-1}) = \bot \text{ and } f(u_r) = i \text{ for } r \in [c], \\ (1,1) & \text{if } f(u_1) = \ldots = f(u_c) = \bot. \end{cases}$$

In words, we determine the label of the vertex tuple $(u_1,\dots,u_c)$ according to the first vertex $u_r$ that is labeled in the assignment $f$. If none of the vertices $u_r$ are labeled, we assign the dummy label $\bot$ to the tuple.

We claim that this partial assignment $F$ satisfies at least $1 - 4\eta - 2e^{-\alpha c}$ of the constraints of $\mathfrak{U}'$, which proves the lemma. (We could make this partial assignment into a total assignment by replacing $\bot$ by an arbitrary label, say $(1,1)$. This change can only increase the number of satisfied constraints.) To establish this claim, it is enough to show that

$$\mathop{\mathbb{P}}_{\substack{(u_1,v_1,\pi_1)\sim\mathfrak{U},\\ \dots,\\ (u_c,v_c,\pi_c)\sim\mathfrak{U}}} \Big\{\exists r\in[c],\, i\in\Sigma.\ F(u_1,\dots,u_c) = (r,i),\ F(v_1,\dots,v_c) = (r,\pi_r(i))\Big\} \geqslant 1 - 2\eta - e^{-\alpha c}. \quad (6.2)$$

(In fact, the unique game $\mathfrak{U}'$ is the square of the unique game corresponding to the verifier above.)

For our assignment $F$, the probability in (6.2) simplifies to the product of two probabilities $p_1$ and $p_2$,

$$\begin{aligned}
&\mathop{\mathbb{P}}_{\substack{(u_1,v_1,\pi_1)\sim\mathfrak{U},\\ \dots,\\ (u_c,v_c,\pi_c)\sim\mathfrak{U}}} \left\{\exists r \in [c],\, i \in \Sigma.\ F(u_1,\dots,u_c) = (r,i),\ F(v_1,\dots,v_c) = (r,\pi_r(i))\right\} \\
&= \mathop{\mathbb{P}}_{\substack{(u_1,v_1,\pi_1)\sim\mathfrak{U},\\ \dots,\\ (u_c,v_c,\pi_c)\sim\mathfrak{U}}} \{\exists r \in [c].\ f(u_r) \neq \bot \vee f(v_r) \neq \bot\} \\
&\qquad \cdot \mathop{\mathbb{P}}_{(u,v,\pi)\sim\mathfrak{U}} \{f(v) = \pi(f(u)) \mid f(u) \neq \bot \vee f(v) \neq \bot\} \\
&= p_1 \cdot p_2\,.
\end{aligned}$$

Since $f$ is $\alpha$-partial, the event $f(u_1) = \ldots = f(u_c) = \bot$ has probability $(1-\alpha)^c \leqslant e^{-\alpha c}$. Hence, with probability at least $1 - e^{-\alpha c}$, one of the vertices $u_1, \dots, u_c$ is labeled. Therefore, $p_1 \geqslant 1 - e^{-\alpha c}$.

Furthermore, we can lower bound the second probability $p_2$ as follows

$$\begin{aligned}
p_2 &= \mathop{\mathbb{P}}_{(u,v,\pi)\sim\mathfrak{U}} \left\{f(v) = \pi(f(u)) \,\middle|\, f(u) \neq \bot \vee f(v) \neq \bot\right\} \\
&= \frac{\mathbb{P}_{(u,v,\pi)\sim\mathfrak{U}}\{f(v) = \pi(f(u))\}}{\mathbb{P}_{(u,v,\pi)\sim\mathfrak{U}}\{f(u) \neq \bot \vee f(v) \neq \bot\}} \\
&= \frac{\mathbb{P}_{(u,v,\pi)\sim\mathfrak{U}}\{f(v) = \pi(f(u))\}}{2\,\mathbb{P}_{u\sim V}\{f(u) \neq \bot\} - \mathbb{P}_{(u,v,\pi)\sim\mathfrak{U}}\{f(u) \in \Sigma,\ f(v) \in \Sigma\}} \\
&\geqslant \frac{(1-\eta)\alpha}{2\alpha - (1-\eta)\alpha} \geqslant 1 - 2\eta\,.
\end{aligned}$$

We conclude that the probability in (6.2) is at least $p_1 \cdot p_2 \geqslant 1 - 2\eta - e^{-\alpha c}$, as claimed. □

## 6.3.2. Soundness

In the following, we show that Reduction 6.14 (from Partial Unique Games to Unique Games) is sound, that is, given a good assignment for the unique game $\mathfrak{U}' = \Psi_c(\mathfrak{U})$ (obtained by applying Reduction 6.14 to the unique game $\mathfrak{U}$), one can construct a good partial assignment for the original unique game $\mathfrak{U}$.

Let $F\colon V^c \to [c] \times \Sigma$ be an assignment for the unique game $\mathfrak{U}' = \Psi_c(\mathfrak{U})$. (Recall that $\mathfrak{U}$ is a unique game with vertex set $V = \{1,\dots,n\}$ and alphabet $\Sigma = \{1,\dots,k\}$.)

Based on the assignment $F$, we construct a collection $\{f_{U,r}\}$ of partial assignments for $\mathfrak{U}$. For $U \in V^{c-1}$ and $r \in \{1,\ldots,c\}$, define $f_{U,r}\colon V \to \Sigma \cup \{\perp\}$ by

$$f_{U,r}(x) \stackrel{\text{def}}{=} \begin{cases} i & \text{if } F(U +_r x) = (r,i), \\ \perp & \text{otherwise.} \end{cases}$$

(Here, $U +_r x$ denotes the tuple $(u_1,\ldots,u_r,x,u_{r+1},\ldots,u_{c-1})$ in $V^c$ obtained by inserting $u$ in the $r$-th coordinate of $U = (u_1,\ldots,u_{c-1})$.)

We first show that one of the partial assignments $f_{U,r}$ has good value in the unique game $\mathfrak{U}^2$ (the square of $\mathfrak{U}$, obtained by sampling two constraints of $\mathfrak{U}$ with a common vertex and composing the permutations).

**Lemma 6.17.** *Let $\beta > 0$. Then, there exists a partial assignment $f_{U,r}$ that labels at least $\beta/c$ of the vertices and has value at least*

$$\mathfrak{U}^2(f_{U,r}) \geqslant \frac{\mathfrak{U}'(F) - \beta}{1-\beta} \cdot \mathop{\mathbb{P}}_{x\sim V}\{f_{U,r}(x) \neq \perp\},$$

*where $\mathfrak{U}^2$ denotes the square of the unique game $\mathfrak{U}$ (obtained by sampling two constraints of $\mathfrak{U}$ with a common vertex and composing the permutations). In particular,* $\mathrm{opt}_{\beta/c}(\mathfrak{U}^2) \geqslant (\mathfrak{U}(F) - \beta)/(1-\beta)$.

We compute the typical value of $f_{U,r}$ for the square of the unique game $\mathfrak{U}$ (obtained by sampling two constraints from $\mathfrak{U}$ with a common vertex and composing the permutations),

$$\begin{aligned} &\mathop{\mathbb{E}}_{U\sim V^{c-1},\ r\in[c]} \mathfrak{U}^2(f_{U,r}) \\ &= \mathop{\mathbb{P}}_{U\sim V^{c-1},\ r\in[c],\ u\sim V,\ (u,v,\pi),(u,v',\pi')\sim \mathfrak{U}\mid u} \left\{\pi^{-1}(f_{U,r}(v)) = (\pi')^{-1}(f_{U,r}(v'))\right\} \\ &= \tfrac{1}{c}\,\mathfrak{U}'(F). \end{aligned}$$

For $U \in V^{c-1}$ and $r \in [c]$, define $\alpha_{U,r}$ as the fraction of vertices labeled by the partial assignment $f_{U,r}$,

$$\alpha_{U,r} \stackrel{\text{def}}{=} \mathop{\mathbb{P}}_{x\sim V}\{f_{U,r}(x) \neq \perp\}.$$

The typical value of $\alpha_{U,r}$ is $1/c$,

$$\mathop{\mathbb{E}}_{U\sim V^{c-1},\ r\in[c]} \alpha_{U,r} = \mathop{\mathbb{P}}_{U\sim V^{c-1},\ r\in[c],\ x\sim V}\{f_{U,r}(x) \neq \perp\} = \tfrac{1}{c}.$$

Since $\mathcal{U}^2(f_{U,r}) \leqslant \alpha_{U,r}$, the contribution of assignments $f_{U,r}$ with $\alpha_{U,r} \leqslant \beta/c$ to the expected value $\mathcal{U}(f_{U,r})$ is at most

$$\mathop{\mathbb{E}}_{U \in V^{c-1},\ r \in [c]} \mathcal{U}^2(f_{U,r}) \cdot \mathbb{1}_{\alpha_{U,r} \leqslant \beta/c} \leqslant \beta/c\,.$$

Therefore, we can lower bound the $\beta/c$-partial value of the unique game $\mathcal{U}^2$ as follows

$$\mathrm{opt}_{\beta/c}(\mathcal{U}) \geqslant \frac{\mathbb{E}_{U \sim V^{c-1},\ r \in [c]}\ \mathcal{U}^2(f_{U,r}) \cdot \mathbb{1}_{\alpha_{U,r} \geqslant \beta/c}}{\mathbb{E}_{U \sim V^{c-1},\ r \in [c]}\ \alpha_{U,r} \cdot \mathbb{1}_{\alpha_{U,r} \geqslant \beta/c}} \geqslant \frac{\mathcal{U}'(F)/c - \beta/c}{1/c - \beta/c} = \frac{\mathcal{U}'(F) - \beta}{1 - \beta}\,.$$

Using the previous lemma (Lemma 6.17) and basic relations between the optimal value of a unique game and its square, we can show the following lemmas.

**Lemma 6.18** (Soundness close to 0). *Let $\mathcal{U}' = \Psi_c(\mathcal{U})$ be the unique game obtained by applying Reduction 6.14 with parameter $c \in \mathbb{N}$ to the unique game $\mathcal{U}$. Suppose the value of $\mathcal{U}'$ is at least $\mathrm{opt}(\mathcal{U}') \geqslant \zeta$. Then, the $1/2c$-partial value of $\mathcal{U}$ satisfies $\mathrm{opt}_{1/2c}(\mathcal{U}) \geqslant \zeta/8$.*

Let $\mathcal{U}_{\circlearrowleft}$ be the unique game obtained by sampling with probability $1/2$ a random constraint from $\mathcal{U}$ and sampling with the remaining probability a trivial constraint $(u, u, \mathrm{id})$. Note that $\mathcal{U}_{\circlearrowleft}(f) = 1/2(\alpha + \mathcal{U}(f))$ for every assignment $f\colon V \to \Sigma \cup \{\bot\}$ that labels an $\alpha$ fraction of the vertices.

**Lemma 6.19** (Soundness close to 1). *Let $\mathcal{U}' = \Psi_c(\mathcal{U}_{\circlearrowleft})$ be the unique game obtained by applying Reduction 6.14 with parameter $c \in \mathbb{N}$ to the unique game $\mathcal{U}_{\circlearrowleft}$. Suppose the value of $\mathcal{U}'$ is at least $\mathrm{opt}(\mathcal{U}') \geqslant 1 - \eta$. Then, the $1/2c$-partial value of $\mathcal{U}$ satisfies $\mathrm{opt}_{1/2c}(\mathcal{U}) \geqslant 1 - \eta/16$.*

## 6.4. From Unique Games to Small-Set Expansion

**Reduction 6.20** (Reduction from Unique Games to Small-Set Expansion).

***Input:*** A unique game $\mathcal{U}$ with vertex set $V$ and alphabet $\Sigma = \{1, \ldots, R\}$, and parameters $q \in \mathbb{N}$ and $\varepsilon > 0$.

***Output:*** A graph $H = H_{q,\varepsilon}(\mathcal{U})$ with vertex set $V \times [q]^R$.

The edge distribution of $H$ is obtained from the unique game $\mathcal{U}$ in the following way:

1. Sample a random vertex $u \sim V$.
2. Sample a random point $x \in [q]^R$.
3. Sample two random constraints $(u, v, \pi), (u, v', \pi') \sim \mathfrak{U} \mid u$. (Here, $\mathfrak{U} \mid u$ denotes the uniform distribution over constraints of $\mathfrak{U}$ containing vertex $u$.)
4. Sample two random points $y, y' \sim T_{1-\varepsilon}(x)$. (Here, $T_{1-\varepsilon}$ is the usual noise graph on $[q]^R$ with noise parameter $1-\varepsilon$. See for a detailed definition.).
5. Output an edge between $(v, y)$ and $(v', y')$.

(End of Reduction 6.20)

The running time of Reduction 6.20 is $\operatorname{poly}(|V|, q^R)$. In particular, the running time is polynomial if $q = O(1)$ and $R = O(\log n)$.

The reduction Reduction 6.20 has the following approximation guarantee.

**Theorem 6.21.** *For every $\varepsilon > 0$ and $q \in \mathbb{N}$, there exist $\eta_\varepsilon > 0$ and $\zeta_{q,\varepsilon} > 0$ such that if $H = H_{q,\varepsilon}(\mathfrak{U})$ is the graph obtained by applying Reduction 6.20 to a unique game $\mathfrak{U}$, then the following two assertions hold:*

***Completeness:*** *If the unique game $\mathfrak{U}$ has value $\operatorname{opt}(\mathfrak{U}) \geqslant 1-\varepsilon$, then the graph $H$ contains a vertex set with volume $1/q$ and expansion at most $4\varepsilon$.*

***Soundness:*** *If the unique game $\mathfrak{U}$ has value $\operatorname{opt}(\mathfrak{U}) \leqslant \zeta_{q,\varepsilon}$ and every vertex set of volume $1/q$ in the constraint graph of $\mathfrak{U}$ has expansion at least $1-\eta_\varepsilon$, then every vertex set of volume $1/q$ has expansion $1-\varepsilon$ in $H$.*

**Remark 6.22.** The completeness assertion of Reduction 6.20 is in fact stronger than stated in Theorem 6.21. If the unique game $\mathfrak{U}$ has value $\operatorname{opt}(\mathfrak{U}) \geqslant 1-\varepsilon$, then the graph $H = H_{q,\varepsilon}$ allows a partition into $q$ vertex sets, each of volume $1/q$ and with expansion at most $4\varepsilon$.

### 6.4.1. Completeness

In the following, we show that Reduction 6.20 is complete, that is, we show that if the unique game $\mathfrak{U}$ has a good assignment (value $1-\eta$), then the graph $H = H_{q,\varepsilon}(\mathfrak{U})$ obtained by Reduction 6.20 contains a set with volume $1/q$ and expansion $O(\varepsilon + \eta)$.

**Lemma 6.23** (Completeness of Reduction 6.20). *Suppose the unique game $\mathfrak{U}$ has optimal value $\mathrm{opt}(\mathfrak{U}) \geqslant 1-\eta$. Then, the graph $H = H_{q,\varepsilon}(\mathfrak{U})$ obtained from Reduction 6.20 contains a vertex set $S$ with volume $\mu(S) = 1/q$ and expansion $\Phi(S) \leqslant 2(\eta+\varepsilon)$.*

*Proof.* Let $F\colon V \to [R]$ be an assignment for the unique game $\mathfrak{U}$. Suppose the assignment $F$ has value $\mathfrak{U}(F) = 1-\eta$. Consider the vertex subset $S \subseteq V \times [q]^R$ in the graph $H = H_{q,\varepsilon}(\mathfrak{U})$ (as specified in Reduction 6.20),

$$S \stackrel{\text{def}}{=} \left\{(u,x) \in V \times [q]^R \mid x_{F(u)} = 1\right\}.$$

(We remark that the choice of 1 is arbitrary. The following proof would work for any element of $[q]$.)

The volume of $S$ in the graph $H$ equals $\mu(S) = 1/q$. To determine the expansion of $S$, we compute $H(S,S)$, the fraction of edge of $H$ inside of $S$. Let $(u,x)$, $(v,y)$, $\pi$, $(v',y')$, $\pi'$ be distributed as in Reduction 6.20. Then,

$$\begin{aligned}
H(S,S) &= \mathop{\mathbb{P}}_{(u,x),\ (v,y),\ \pi,\ (v',y'),\ \pi'} \left\{(\pi.y)_{F(v)} = 1 \wedge (\pi'.y')_{F(v')} = 1\right\} \\
&\geqslant (1-\varepsilon)^2 \mathop{\mathbb{P}}_{(u,x),\ (v,y),\ \pi,\ (v',y'),\ \pi'} \left\{(\pi.x)_{F(v)} = 1 \wedge (\pi'.x)_{F(v')} = 1\right\} \\
&\geqslant (1-\varepsilon)^2 \mathop{\mathbb{P}}_{(u,x),\ (v,y),\ \pi,\ (v',y'),\ \pi'} \left\{x_{F(u)} = 1 \wedge F(v) = \pi(F(u)) \wedge F(v') = \pi'(F(u))\right\} \\
&= (1-\varepsilon)^2 \mathfrak{U}^2(F)/q \geqslant (1-\varepsilon)^2(1-2\eta)/q\,.
\end{aligned}$$

The first inequality uses that $\mathbb{P}_{y\sim T_\rho(x)}\{y_r = x_r\} \geqslant \rho$ for every coordinate $r \in [R]$. For the second inequality, we use that one event logically implies the other event. The notation $\mathfrak{U}^2(F)$ denotes the value of the assignment in the square of the unique game $\mathfrak{U}$. (The square of a unique game is obtained by sampling two random constraints with a common vertex and composing the permutations.)

We conclude that, as desired, $S$ satisfies

$$\begin{aligned}
\mu(S) &= \tfrac{1}{q}\,, \\
\Phi(S) &\leqslant 2\varepsilon + 2\eta\,.
\end{aligned}$$

□

### 6.4.2. Soundness

In the following, we show that Reduction 6.20 is sound, that is, if the graph $H = H_{q,\varepsilon}(\mathfrak{U})$ contains a vertex set $S$ with small volume and expansion bounded away from 1, then either there exists a good assignment for the unique game $\mathfrak{U}$ or the constraint graph of $\mathfrak{U}$ contains a vertex with small volume and expansion bounded away from 1. (See Lemma 6.24 at the end of this section for the formal statement.)

Let $S \subseteq V \times [q]^R$ be a vertex subset of the graph $H = H_{q,\varepsilon}(\mathfrak{U})$ obtained by applying Reduction 6.20 to a unique game $\mathfrak{U}$ with vertex set $V$ and alphabet $[R]$. Let $f = \mathbb{1}_S$ be the indicator function of $S$.

For a vertex $u \in V$, we define $f_u \colon [q]^R \to \{0,1\}$ by $f_u(x) := f(u,x)$. We define $g_u \colon [q]^R \to [0,1]$ by averaging the functions $f_v$ over the constraints $(u,v,\pi) \sim \mathfrak{U} \mid u$,

$$g_u(x) := \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}|u} f_v(\pi.x).$$

(For a constraint $(u,v,\pi)$, we identify the $i^{\text{th}}$ input coordinate of $g_u$ with the $\pi(i)^{\text{th}}$ input coordinate of $f_v$.)

The typical noise stability of the functions $g_u$ equals $H(S,S)$, the fraction of edges staying inside of $S$,

$$\begin{aligned} H(S,S) &= \mathop{\mathbb{E}}_{u\sim V,\ x\in[q]^R} \Bigg( \mathop{\mathbb{P}}_{\substack{(u,v,\pi)\sim\mathfrak{U}|u \\ y\sim T_{1-\varepsilon}(x)}} \Big\{ (v,y) \in S \Big\} \Bigg)^2 \\ &= \mathop{\mathbb{E}}_{u\sim V,\ x\in[q]^R} T_{1-\varepsilon} g_u(x)^2 = \mathop{\mathbb{E}}_{u\sim V} \|T_{1-\varepsilon} g_u\|^2 . \end{aligned}$$

The invariance principle (Theorem 2.10) allows us to estimate the noise stability $\|T_{1-\varepsilon} g_u\|^2 = \langle g_u, T_{(1-\varepsilon)^2} g_u \rangle$ in terms of Gaussian noise stability bounds, whenever the function $g_u$ has no influential coordinates. For the reader's convenience, we restate the invariance principle.

**Theorem** (Restatement of Theorem 2.10). *For every finite probability space $\Omega$ and constants $\rho \in [0,1)$, $\eta > 0$, there exists constants $\tau, \gamma > 0$ such that for every function $f \in L_2(\Omega^R)$ with $0 \leqslant f \leqslant 1$, either*

$$\langle f, T_\rho f \rangle \leqslant \Gamma_\rho(\mathbb{E} f) + \eta,$$

*or $\mathrm{Inf}_i\, T_{1-\gamma} f > \tau$ for some coordinate $i \in [R]$.*

Let $\eta > 0$ be sufficiently small (we determine this parameter later more precisely). By the invariance principle, there exists constants $\tau, \gamma > 0$ (depending on $q$, $\varepsilon$, and $\eta$) such that

$$H(S,S) = \mathop{\mathbb{E}}_{u\sim V} \|T_{1-\varepsilon} g_u\|^2 \leqslant \mathop{\mathbb{E}}_{u\sim V} \Gamma_{(1-\varepsilon)^2}(\mathbb{E} g_u) + \eta + \mathop{\mathbb{P}}_{v\sim V} \left\{ \max_{i\in[R]} \mathrm{Inf}_i\, T_{1-\gamma} g_u > \tau \right\} .$$

(Here, we also use that $\|T_{1-\varepsilon} g_u\|^2 \leqslant 1$ since $0 \leqslant g_u \leqslant 1$.)

We use the following basic bound for the Gaussian noise stability. Let $\delta > 0$ be sufficiently small (we determine this parameter later). Then,

$$\Gamma_{(1-\varepsilon)^2}(\mathbb{E} g_u) \leqslant \mathbb{E} g_u \cdot (\delta^{\varepsilon/2} + \mathbb{1}_{\mathbb{E} g_u > \delta}).$$

(If $\mathbb{E} g_u \leqslant \delta$, then we can bound the noise stability by $\delta^{\varepsilon/2} \mathbb{E} g_u$. Otherwise, we can use the trivial bound $\Gamma_{(1-\varepsilon)^2}(\mathbb{E} g_u) \leqslant \mathbb{E} g_u$.)

Influence decoding (Lemma 2.11) allows us to estimate the fraction of vertices $u$ such that $T_{1-\gamma} g_u$ has a coordinate with influence larger than $\tau$.

**Lemma** (Restatement of Lemma 2.11). *Let $\mathcal{U}$ be a unique game with vertex set $V$ and alphabet $[R]$. For some probability space $\Omega$, let $\{f_u\}_{u \in V}$ be a collections of normalized functions in $L_2(\Omega^R)$. Consider functions $g_u$ in $L_2(\Omega^R)$ defined by*

$$g_u(x) = \mathop{\mathbb{E}}_{(u,v,\pi) \sim \mathcal{U} | u} f_v(\pi.x).$$

*Then, for all $\gamma, \tau > 0$, there exists $c_{\gamma,\tau} > 0$ (in fact, $c_{\gamma,\tau} = \operatorname{poly}(\gamma, \tau)$) such that*

$$\operatorname{opt}(\mathcal{U}) \geqslant c_{\gamma,\tau} \cdot \mathop{\mathbb{P}}_{u \sim V} \left\{ \max_{i \in [R]} \operatorname{Inf}_i T_{1-\gamma} g_u > \tau \right\}.$$

Combining Gaussian noise stability bounds and influence decoding, our previous upper bound on $H(S,S)$ simplifies to

$$H(S,S) \leqslant \delta^{\varepsilon/2} \cdot \mu(S) + \eta + \operatorname{opt}(\mathcal{U})/c_{\gamma,\tau} + \mathop{\mathbb{E}}_{u \sim V} \mathbb{1}_{\mathbb{E} g_u > \delta} \mathbb{E} g_u.$$

This bound asserts the following: If $H(S,S)$ is bounded away from $\delta^{\varepsilon/2} \cdot \mu(S) + \eta$ (meaning the expansion $S$ is bounded away from 1), then either $\operatorname{opt}(\mathcal{U})$ is bounded away from 0 or $\mathbb{E}_{u \sim V} \mathbb{1}_{\mathbb{E} g_u > \delta} \mathbb{E} g_u$ is bounded away from 0. To show the soundness of Reduction 6.20, it remains to argue that the latter condition implies that the constraint graph of the unique game $\mathcal{U}$ contains a vertex set with small volume and expansion bounded away from 1.

To this end, we consider functions $\bar{f}, \bar{g} \colon V \to [0,1]$, obtained by averaging the functions $f_u$ and $g_u$ over the hypercube $[q]^R$,

$$\bar{f}(u) = \mathop{\mathbb{E}}_{x} f_u(x),$$
$$\bar{g}(u) = \mathop{\mathbb{E}}_{x} g_u(x).$$

The functions $\bar{f}$ and $\bar{g}$ satisfy the relation $\bar{g} = G\bar{f}$, where $G$ is the constraint graph of the unique game $\mathfrak{U}$, because

$$\begin{aligned} G\bar{f}(u) &= \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}|u} \bar{f}(v) \\ &= \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}|u} \mathop{\mathbb{E}}_{x\in[q]^R} f_v(x) \\ &= \mathop{\mathbb{E}}_{x\in[q]^R} \mathop{\mathbb{E}}_{(u,v,\pi)\sim\mathfrak{U}|u} f_v(\pi.x) \\ &= \mathop{\mathbb{E}}_{x\in[q]^R} g_u(x) = \bar{g}(u). \end{aligned}$$

We can estimate $\mathbb{E}_{u\sim V} \mathbb{1}_{\mathbb{E} g_u>\delta} \mathbb{E} g_u$ in terms of the $L_2$-norm of $\bar{g} = G\bar{f}$,

$$\mathop{\mathbb{E}}_{u\sim V} \mathbb{1}_{\mathbb{E} g_u>\delta} \mathbb{E} g_u \leqslant \mathop{\mathbb{E}}_{u} \bar{g}(u)\mathbb{1}_{\bar{g}(u)>\delta} \leqslant \mathop{\mathbb{E}}_{u} \bar{g}(u)^2/\delta = \|G\bar{f}\|^2/\delta.$$

At this point, we can use the following lemma, which we used in the soundness analysis of Reduction 6.20 earlier in this chapter.

**Lemma** (Restatement of Lemma 6.11). *Suppose $f\colon V \to \mathbb{R}$ satisfies $0 \leqslant f(x) \leqslant 1$ for every vertex $x \in V$. Then, for every $\beta > 0$, there exists a set $S \subseteq V$ such that*

$$\begin{aligned} \beta\|f\|_1 \leqslant \mu(S) \leqslant \tfrac{1}{\beta}\|f\|_1, \\ \Phi_{G^2}(S) \leqslant 1 - \tfrac{\|Gf\|^2}{\|f\|_1} + 2\beta + \beta/(n\|f\|_1). \end{aligned}$$

Let $\beta > 0$ be sufficiently small (we determine this parameter later) and let $S' \subseteq V$ be the vertex set obtained by applying Lemma 6.11 to the function $\bar{f}$. The set $S'$ satisfies

$$\begin{aligned} \beta\mu(S) \leqslant \mu_G(S') \leqslant \mu(S)/\beta \\ \Phi_{G^2}(S') \leqslant 1 - \|G\bar{f}\|^2/\mu(S) + 2\beta + \beta/(n\mu(S)). \end{aligned}$$

We arrive at the following upper bound on the fraction of edges staying inside the vertex set $S$ (normalized by its volume),

$$\frac{H(S,S)}{\mu(S)} \leqslant \delta^{\varepsilon/2} + \tfrac{\eta}{\mu(S)} + \tfrac{1}{c_{\gamma,\tau}\mu(S)} \operatorname{opt}(\mathfrak{U}) + \tfrac{1}{\delta} \cdot \frac{G^2(S',S')}{\mu(S')} + 2\beta + \tfrac{\beta}{n\mu(S)}$$

To simplify this bound further, we eliminate a few parameters. Choose $\delta$ such that $\delta^{\varepsilon/2} \ll \beta$ and choose $\eta$ such that $\eta \ll \beta\mu(S)$. Furthermore, we assume that $\mu(S) \gg 1/n$ (a very mild assumption). With these parameter choices, the bound on $H(S,S)$ simplifies to

$$\frac{H(S,S)}{\mu(S)} \leqslant 4\beta + C_{\beta,q,\varepsilon,\mu(S)}\,\mathrm{opt}(\mathfrak{U}) + \beta^{-2/\varepsilon} \cdot \frac{G^2(S',S')}{\mu(S')},$$

where $C_{\beta,q,\varepsilon}$ is some constant depending only on $\beta$, $q$, $\mu(S)$, and $\varepsilon$.

The last bound on $H(S,S)$ implies the following lemma asserting the soundness of Reduction 6.20. (Here, we also use the fact that a vertex set with expansion bounded away from 1 in $G^2$ implies a vertex set with roughly the same volume and expansion bounded away from 1 in $G$.)

**Lemma 6.24.** *For every $q \in \mathbb{N}$ and $\varepsilon, \eta > 0$, there exists $\zeta_{q,\varepsilon,\eta}, \zeta'_{\varepsilon,\eta} > 0$ such that the following holds: Let $H = H_{q,\varepsilon}(\mathfrak{U})$ be the graph obtained by applying Reduction 6.20 with parameters $q$ and $\varepsilon$ to the unique game $\mathfrak{U}$. Then, if $H$ contains a vertex set $S$ with volume $\mu_H(S) = 1/q$ and expansion at most $\Phi_H(S) \leqslant 1 - \eta$, then either* $\mathrm{opt}(\mathfrak{U}) > \zeta_{q,\varepsilon,\eta}$ *or the constraint graph $G$ of $\mathfrak{U}$ contains a vertex set $S'$ with volume $\mu_G(S') = 1/q$ and expansion at most $\Phi_G(S') \leqslant 1 - \zeta'_{\varepsilon,\eta}$.*

## 6.5. Notes

The material presented in this chapter is based on the paper "Graph Expansion and the Unique Games Conjecture" [RS10] joint with Prasad Raghavendra. A preliminary version appeared at STOC 2010.

# 7. Reductions between Expansion Problems

We show that the SSE hypothesis implies quantitatively tight inapproximability results for many graph expansion problems, in particular Sparsest Cut and Balanced Separator. Concretely, our results imply that it is SSE-hard to beat Cheeger's inequality and achieve an $(\varepsilon, o(\sqrt{\varepsilon}))$-approximation for Sparsest Cut. (We say a problem is SSE-hard if an efficient algorithm for the problem implies that the SSE hypothesis is false.)

These are the first strong hardness of approximation results for Sparsest Cut and Balanced Separator. (Even assuming the Unique Games Conjecture, no strong hardness result for these problems were known. The best known result was that one cannot achieve approximation ratios arbitrarily close to 1 unless 3-Sat has subexponential algorithms [AMS07].)

There is a strong parallel between this result and the known consequences of the Unique Games Conjecture. The Unique Games Conjecture asserts a qualitative inapproximability for Unique Games, which is generalization of Max Cut. In turn, the conjecture implies tight quantitative inapproximability results for Max Cut and similar basic problems. Similarly, the SSE hypothesis asserts a qualitative inapproximability for Small-Set Expansion, which generalizes Sparsest Cut. In turn, the hypothesis implies tight quantitative inapproximability results for Sparsest Cut and other graph expansion problems. The value of results of this kind is that they unify the question of improving known algorithms for a class of problems to a question about the qualitative approximability of a single problem.

These significant consequences of a confirmation of the SSE hypothesis make it an interesting open question to prove or refute the hypothesis.

## 7.1. Main Results

We show that following general graph expansion problem is SSE-hard.

**Theorem 7.1.** *The following problem is SSE-hard for all constants $\varepsilon, \delta, \eta > 0$: Given a graph $G$ distinguish between the cases,*

**YES:** *the graph $G$ contains a vertex set with volume at most $\delta$ and expansion at most $\varepsilon + \eta$,*

**NO:** *every vertex set $S$ in $G$ satisfies*

$$G(S,S) \leqslant \Gamma_{1-\varepsilon}\Big(\mu(S)\Big) + \eta .$$

A corollary of this theorem is an essentially tight SSE-hardness for Sparsest Cut. Concretely, given a graph $G$, it is SSE-hard for every $\varepsilon > 0$ to distinguish between the cases $\Phi_G \leqslant \varepsilon$ and $\Phi_G > \Omega(\sqrt{\varepsilon})$. This hardness is essentially tight because Cheeger's inequality allows us to distinguish between $\Phi_G \leqslant \varepsilon$ and $\Phi_G > O(\sqrt{\varepsilon})$. (However, the constants hidden in the $\Omega(\cdot)$- and $O(\cdot)$-notations are not the same.)

The YES case in the above theorem can be strengthened in various ways (leading to stronger hardness results). For example, we can assert that there exists a distribution over vertex sets $S$ such that

- with probability 1, the vertex set $S$ has volume at most $\delta$ and expansion at most $\varepsilon$,
- the distribution over vertex sets $S$ uniformly covers essentially the whole graph, in the sense that,

$$\mathop{\mathbb{E}}_{v\sim G}\left|\mathop{\mathbb{P}}_{S}\{v \in S\} - \mathop{\mathbb{E}}_{S}\mu(S)\right| \leqslant \eta \mathop{\mathbb{E}}_{S}\mu(S).$$

  (The volume of $S$ is also lower bounded by a function of $\eta$ and $\delta$.)

Such a distribution over vertex sets implies the existence of SDP solutions with certain symmetry properties, leading to quantitatively tight integrality gaps for Balanced Separator. (Using techniques in Chapter 8.)

Another consequence of such distributions is that one can extract vertex sets with desired volume while maintaining the expansion up to a factor of 2. (To extract the vertex set with the desired volume we use correlated sampling, similar to §4.2 in Chapter 4.) This observation implies the corollary of the previous theorem.

**Theorem 7.2.** *The following problem is SSE-hard for all constants $\varepsilon, \delta, \eta > 0$: Given a graph $G$ distinguish between the cases,*

***YES:*** *the graph $G$ allows a partition of the vertex set into sets with volume equal to $\delta$ and expansion at most $2\varepsilon + \eta$,*

***NO:*** *every vertex set $S$ in $G$ satisfies*

$$G(S,S) \leqslant \Gamma_{1-\varepsilon}\Big(\mu(S)\Big) + \eta\,.$$

## 7.2. Reduction

The reduction in this section composes the input graph with certain other graphs, in particular, noise graphs (see Section 2.4 for details about noise graphs). Beyond noise graphs, the reduction uses the following kind of graphs (which are related to noise graphs).

Let $\Omega$ be a finite probability space and let $R \in \mathbb{N}$. For $w \in \{\bot,\top\}^R$, let $M_w$ be the Markov operator on $L_2(\Omega^R)$ corresponding to the graph on $\Omega^R$ with the following edge distribution:

1. Sample $x \sim \Omega^R$.
2. For every coordinate $r \in [R]$, put $y_r := x_r$ if $w_r = \top$ and sample $y_r \sim \Omega$ if $w_r = \bot$.
3. Output the pair $xy$ as a random edge of $M_w$.

(The graph $M_w$ also has a simple combinatorial description. It is a disjoint union of cliques and two vertices $x,y \in \Omega^R$ are in the same clique if they agree on the coordinates $r$ with $w_r = \top$.)

Let $\{\bot,\top\}^R_\beta$ be $R$-fold product of the $\beta$-biased distribution on $\{\bot,\top\}$. In other words, $\{\bot,\top\}^R_\beta$ is the distribution over $\{\bot,\top\}^R$ in which each coordinate equals $\top$ with probability $\beta$ (independently for each coordinate).

Averaging the graph $M_w$ over $w \sim \{\bot,\top\}^R_\beta$ yields the noise graph $T_\beta$ on $\Omega^R$,

$$\mathop{\mathbb{E}}_{w\sim\{\bot,\top\}^R_\beta} M_w = T_\beta\,. \tag{7.1}$$

**Reduction 7.3.**

***Input:*** A regular graph $G$ with vertex set $V$, and parameters $R,q \in \mathbb{N}$ and $\varepsilon,\rho,\beta \in (0,1)$.

***Output:*** A regular graph $H = H_{R,q,\varepsilon,\rho,\beta}(G)$ with vertex set $V(H) = V^R \times [q]^R \times \{\perp, \top\}_\beta^R$.

The edge distribution of the graph $H$ is obtained from $G$ in the following way:

1. Sample a random $R$-tuple $A \sim V^R$ of vertices.
2. Sample two random neighbors $B, B' \sim (T_{1-\varepsilon}G)^{\otimes R}(A)$ of $A$ in the graph $(T_{1-\varepsilon}G)^{\otimes R}$. (Here, the graph $T_{1-\varepsilon}G$ corresponds to first taking a random step in $G$ and then taking a random step in $T_{1-\varepsilon}$, the usual noise graph on $V$ with parameter $1-\varepsilon$. See for further details.).)
3. Sample two random permutations $\pi, \pi' \in \mathbb{S}_R$.
4. Sample a random point $(x,z) \in [q]^R \times \{\perp, \top\}_\beta^R$.
5. Sample a random neighbors $(y,w),(y',w') \sim T_\rho^{\otimes R}(x,z)$ of $x$ in the graph $T_\rho^{\otimes R}$. (Here, $T_\rho$ is the usual noise graph on graph on $[q] \times \{\perp, \top\}_\beta$. See .).
6. Sample random neighbors $(\tilde{B}, \tilde{y}) \sim M_w(B,y)$ and $(\tilde{B}', \tilde{y}') \sim M_{w'}(B', y')$. (Here, $M_w$ is the operator that corresponds to resampling the coordinates $i \in [R]$ with $w_i = \perp$. See .)
7. Output an edge between $\pi.(\tilde{B}, \tilde{y}, w)$ and $\pi'.(\tilde{B}', \tilde{y}', w')$.

(End of Reduction 7.3)

Reduction 7.3 has the following approximation guarantees.

**Theorem 7.4.** *Given a graph $G$ with $n$ vertices and parameters $R \in \mathbb{N}$, $q \in \mathbb{N}$, $\varepsilon, \rho, \beta \in (0,1)$ with $\rho \geqslant 0.1$, Reduction 7.3 computes in time $\mathrm{poly}(n^R, q^R)$ a graph $H = H_{R,q,\varepsilon,\rho,\beta}(G)$ such that the following assertions hold (for some constants $\zeta = \zeta(q,\varepsilon,\rho,\beta) > 0$ and $\alpha \geqslant (1-\varepsilon)\varepsilon$):*

***Completeness I:*** *If $G$ contains a vertex set with volume $\varepsilon/\beta R$ and expansion at most $\varepsilon$, then $H$ contains $q$ disjoint vertex sets $T_1, \ldots, T_q$, each with volume $\mu(T_i) = \alpha/q$ and expansion at most $\Phi(T_i) \leqslant 1 - \rho^2 + O(\varepsilon + \beta)$.*

***Completeness II:*** *If $G$ contains $\alpha\beta R/\varepsilon$ disjoint vertex sets, each with volume $\varepsilon/\beta R$ and expansion at most $\varepsilon$, then there exists a distribution over $q$ disjoint vertex sets $T_1, \ldots, T_q$ such that*

- *with probability 1, each set $T_i$ has volume and expansion as above,*

– *a typical vertex $v$ in $H$ satisfies* $\left|\mathbb{P}_{T_i}\{v \in T_i\} - \alpha/q\right| \leqslant O(\varepsilon)\alpha/q + 2^{-\Omega(\varepsilon^3 \beta R)}$ *for every $i \in [q]$ (which means that the sets $T_i$ uniformly cover all but a small fraction of the graph $H$).*

**Soundness:** *If every vertex set in $G$ with volume $\varepsilon/\beta R$ has expansion less than $1-\zeta$, then every vertex set $T$ in $H$ satisfies*

$$H(T,T) \leqslant \Gamma_{\rho^2}\big(\mu(T)\big) + 3\beta^{1/3}.$$

**Remark 7.5.** The graph $H$ obtained from Reduction 7.3 is the label-extended graph of unique game. The second completeness assertion shows that this unique game has a uniform cover of $\Omega(1)$-partial assignments with value close to 1 if $G$ contains a vertex set with volume roughly $\varepsilon/\beta R$ and small expansion. Using correlated sampling (similar to the rounding in Chapter 4, §4.2, such a distribution over partial assignments implies that the value of the unique game is close to 1 (only a factor of two is lost in the closeness to 1). This observation implies that the Small-Set Expansion Hypothesis implies Hypothesis 6.3 (Unique Games Conjecture on Small-Set Expanders), which establishes the equivalence of the two hypotheses (we showed the other direction already in Chapter 6, §6.4).

### 7.2.1. Completeness

Let $G$ be a regular graph with vertex set $V = \{1,\ldots,n\}$. Let $H = H_{R,q,\varepsilon,\rho,\beta}(G)$ be the graph with vertex set $V(H) = V^R \times [q]^R \times \{\perp,\top\}^R_\beta$ obtained by applying Reduction 7.3 to $G$. The number of vertices of $H$ is $n^R \cdot O(q)^R$ (and the running time of the reduction is polynomial in the size of $H$). In particular, Reduction 7.3 is polynomial-time for $R = O(1)$ and $q = O(1)$.

The completeness of Reduction 7.3 asserts that if $G$ contains a set of volume roughly $1/\beta R$ and with small expansion, then a constant fraction of the graph $H$ can be partitioned into $q$ sets, each with the same volume and expansion roughly $1-\rho^2$.

**Lemma 7.6** (Completeness I). *If the graph $G$ contains a vertex set of volume $\delta$ and expansion at most $\eta$, then the graph $H = H_{R,q,\varepsilon,\rho,\beta}(G)$ obtained by Reduction 7.3 contains disjoint vertex set $T_1,\ldots,T_q$, each of volume $\alpha/q$ and with expansion at most $1-\rho^2 + O(\varepsilon+\eta+\beta/\rho+\beta R\delta)$, where $\alpha \geqslant (1-\beta R\delta)\beta R\delta$.*

*Proof*. Let $S \subseteq V$ be a vertex subset of $G$ with $\mu(S) = \delta$ and $\Phi(S) = \eta$. Similar to the completeness proof of Reduction 6.5, we construct a partial assignment $F\colon V^R \times \{\bot,\top\}_\beta^R \to [R] \cup \{\bot\}$,

$$F(A,z) \stackrel{\text{def}}{=} \begin{cases} r & \text{if } \{r\} = \{r' \in [R] \mid a_{r'} \in S \wedge z_{r'} = \top\}, \\ \bot & \text{otherwise.} \end{cases}$$

(The difference to the construction in the completeness proof for Reduction 6.5 is the additional argument $z \in \{\bot,\top\}_\beta^R$ of $F$. Effectively, the partial assignment $F$ ignores all coordinates $a_r$ of $A$ with $z_r = \bot$.) We compute the fraction $\alpha$ of vertices labeled by $F$,

$$\alpha := \mathop{\mathbb{P}}_{(A,z)\sim V^R \times \{\bot,\top\}_\beta^R} \{F(A,z) \neq \bot\} = \binom{R}{1}(1-\beta\delta)^{R-1}\beta\delta \geqslant \beta R\delta(1-\beta R\delta).$$

Let $S' = S \times \{\top\} \subseteq V \times \{\bot,\top\}_\beta$. Consider the graph $G' = (T_{1-\varepsilon}G) \otimes T_\rho$ on $V \times \{\bot,\top\}_\beta$. The set $S'$ has volume $\beta\delta$ in the graph $G'$. Furthermore, the fraction of edges of $G'$ staying inside of $S'$ is

$$\begin{aligned} G'(S',S') &= \mu(S') \cdot \mathop{\mathbb{P}}_{a_r\sim V,\ b_r \sim T_{1-\varepsilon}G} \{b_r \in S \mid a_r \in S\} \cdot \mathop{\mathbb{P}}_{z_r\sim\{\bot,\top\}_\beta,\ w_r\sim T_\rho(z_r)} \{w_r = \top \mid z_r = \top\} \\ &= \beta\delta\Big((1-\varepsilon)(1-\eta) + \varepsilon\delta\Big) \cdot \Big(\rho + (1-\rho)\beta\Big). \end{aligned}$$

(Here, we use the notation $a_r, b_r, w_r, z_r$ only for consistency with the description of Reduction 7.3. The index $r$ plays no role at this point.) Hence, the expansion of $S'$ in the graph $G'$ equals

$$\phi' := 1 - \Big((1-\varepsilon)(1-\eta) + \varepsilon\delta\Big) \cdot \Big(\rho + (1-\rho)\beta\Big) \leqslant (1-\rho) + \varepsilon + \eta + \beta.$$

For every $r \in [R]$, we have

$$\begin{aligned} &\mathop{\mathbb{P}}_{A\sim V^R,\ z\sim\{\bot,\top\}_\beta^R,\ B\sim(T_{1-\varepsilon}G)^{\otimes R}(A),\ w\sim T_\rho^{\otimes R}(z)} \Big\{F(B,w) = r \mid F(A,z) = r\Big\} \\ &= (1-\phi')\Big(1 - \phi'\tfrac{\mu(S')}{1-\mu(S')}\Big)^{R-1} \\ &\geqslant 1 - \phi' - R\beta\delta. \end{aligned}$$

Furthermore, for every $r \in [R]$,

$$\begin{aligned} \mathop{\mathbb{P}}_{(y,w)\sim T_\rho^{\otimes R}(x,z)} \{x_r = y_r \mid w_r = \top,\ z_r = \top\} &\geqslant \rho / \mathop{\mathbb{P}}_{w_r\sim T_\rho(z_r)} \{w_r = \top \mid z_r = \top\} \\ &\geqslant \rho/(\rho + (1-\rho)\beta) \geqslant 1 - \tfrac{1-\rho}{\rho}\beta \end{aligned}$$

Let $A, B, w, x, y, z$ be distributed as before (and as in Reduction 7.3), i.e., $A \sim V^R$, $B \sim (T_{1-\varepsilon} G)^{\otimes R}$, $(x, z) \sim [q]^R \times \{\bot, \top\}^R_\beta$, and $(y, w) \sim T_\rho^{\otimes R}(x, z)$. In addition, sample $\tilde{B} \sim M_w(B)$ and $\tilde{y} \sim M_w(y)$ as in Reduction 7.3. Then, for every $r \in [R]$,

$$\begin{aligned}
\rho' := & \mathbb{P}\left\{x_{F(A,z)} = \tilde{y}_{F(\tilde{B},w)} \mid F(A,z) = r\right\} \\
& \geqslant \mathbb{P}\left\{x_r = \tilde{y}_r \wedge F(\tilde{B}, w) = r \mid F(A,z) = r\right\} \\
& = \mathbb{P}\{x_r = y_r \wedge F(B, w) = r \mid F(A,z) = r\} \\
& = \mathbb{P}\{F(B,w) = r \mid F(A,z) = r\} \cdot \mathbb{P}\{x_r = y_r \mid F(B,w) = r,\ F(A,z) = r\} \\
& \geqslant (1 - \phi' - R\beta\delta) \cdot \left(1 - \tfrac{1-\rho}{\rho}\beta\right) \geqslant 1 - \phi' - R\beta\delta - \tfrac{1-\rho}{\rho}\beta .
\end{aligned}$$

The second step uses that $(\tilde{B}, \tilde{y})$ and $(B, y)$ agree on all coordinates $r$ with $w_r = \top$ (and that $F(\cdot, w)$ ignores coordinates with $w_r \neq \top$).

For our final choice of parameters, $\rho'$ will be arbitrarily close to $\rho$ (also note that by symmetry of $F$, the expression used to define $\rho'$ does not depend on $r$.)

We define vertex subsets $T_1, \ldots, T_q \subseteq V(H)$ as follows

$$T_i = \left\{(A, x, z) \in V^R \times [q]^R \times \{\bot, \top\}^R_\beta \ \mid\ x_{F(A,z)} = i\right\}.$$

By symmetry, each set $T_i$ has volume $\mu(T_i) = \alpha/q$ in $H$. (Recall that $\alpha$ is the fraction of $V^R \times \{\bot, \top\}^R_\beta$ labeled by the partial assignment $F$.) We claim that the sets $T_i$ have expansion at most $1 - (\rho')^2$ in the graph $H$. Note that the sets $T_1, \ldots, T_i$ are invariant under permuting coordinates (using that $F(\pi.(A, z)) = \pi(F(A, z))$ whenever $F(A, z) \neq \bot$). Hence, in order to compute the expansion of the sets, we can ignore the permutations $\pi$ and $\pi'$ in the construction of $H$ (see steps 3 and 7 in Reduction 7.3). We can compute the fraction of edges of

$H$ staying in $T_i$ as

$$\begin{aligned}
H(T_i, T_i) &= \mathop{\mathbb{P}}_{\left((\tilde{B},\tilde{y},w),\,(\tilde{B}',\tilde{y}',w')\right)\sim H}\left\{y_{F(B,y)} = i \wedge y'_{F(B',w')} = i\right\} \\
&= \mathop{\mathbb{E}}_{(A,x,z)}\left(\mathop{\mathbb{P}}_{(\tilde{B},\tilde{y},w)\mid(A,x,z)}\left\{\tilde{y}_{F(\tilde{B},w)} = i\right\}\right)^2 \\
&\geqslant \mathop{\mathbb{E}}_{(A,x,z)}\left(\mathop{\mathbb{P}}_{(\tilde{B},\tilde{y},w)\mid(A,x,z)}\left\{x_{F(A,z)} = i \wedge \tilde{y}_{F(\tilde{B},w)} = x_{F(A,z)}\right\}\right)^2 \\
&= \frac{\alpha}{q}\mathop{\mathbb{E}}_{(A,x,z)\mid F(A,z)\neq\perp}\left(\mathop{\mathbb{P}}_{(\tilde{B},\tilde{y},w)\mid(A,x,z)}\left\{\tilde{y}_{F(\tilde{B},w)} = x_{F(A,z)}\right\}\right)^2 \\
&\geqslant \frac{\alpha}{q}\left(\mathop{\mathbb{E}}_{(A,x,z)\mid F(A,z)\neq\perp}\mathop{\mathbb{P}}_{(\tilde{B},\tilde{y},w)\mid(A,x,z)}\left\{\tilde{y}_{F(\tilde{B},w)} = x_{F(A,z)}\right\}\right)^2 \\
&= \frac{\alpha}{q}\cdot(\rho')^2 \geqslant \frac{\alpha}{q}\cdot\left(\rho^2 - O(\varepsilon + \eta + \beta/\rho + \beta R\delta)\right). \qquad \square
\end{aligned}$$

**Remark 7.7** (Completeness II). The second completeness assertion for Reduction 7.3 follows from the proof of the previous lemma (which establishes the first completeness assertion). If the graph $G$ contains $k = \beta R$ disjoint vertex sets $S_1, \ldots, S_k$, each with volume $\delta = \varepsilon/k$ and expansion $\eta \leqslant \varepsilon$, then we can sample a random vertex set $S \in \{S_1, \ldots, S_k\}$ and apply the construction in previous the proof. In this way, we obtain a distribution over disjoint vertex sets $T_1, \ldots, T_q$ satisfying the conditions of our second completeness assertion. To verify the condition that the sets $T_i$ uniformly cover all but a small fraction of the graph $H$, it is enough to show the following the claim: Consider $k$ random vertices $x_1, \ldots, x_k \sim G$. Then, in expectation, at least a $(1 - O(\varepsilon))\varepsilon - 2^{-\Omega(\varepsilon^2 k)}$ fraction of the vertex sets $S_j$ contain exactly one of the vertices $x_1, \ldots, x_k$. (Note that in expectation, at most an $\varepsilon$ fraction of the vertex sets $S_j$ contain *at least* one of the vertices $x_1, \ldots, x_k$.) To verify the claim, let us condition on the event that at least $k' = (1-\varepsilon)\varepsilon k$ of the vertices $x_1, \ldots, x_k$ fall into one of the vertex sets $S_j$. We may assume that the first $k'$ vertices $x_1, \ldots, x_{k'}$ fall into one of the vertex sets $S_j$. Each of these vertices $x_i$ (with $i \leqslant k'$) falls into a random set among $S_1, \ldots, S_k$. Let us imagine we distribute these vertices one-by-one among sets $S_j$ (in the order $x_1, \ldots, x_{k'}$). Then, the probability that $x_i$ falls into a set that already contains one of the vertices $x_1, \ldots, x_{i-1}$ is at most $k'/k \leqslant \varepsilon$. Hence, we expect that a $1 - O(\varepsilon)$ fraction of the vertices $x_1, \ldots, x_{k'}$ fall into a unique set $S_j$, which demonstrates the claim. (End of Remark 7.7)

### 7.2.2. Soundness

Let $G$ be a regular graph with vertex set $V = \{1, \dots, n\}$. Let $H = H_{R,q,\varepsilon,\rho,\beta}(G)$ be the graph with vertex set $V(H) = V^R \times [q]^R \times \{\perp, \top\}^R_\beta$ obtained by applying Reduction 7.3 to $G$.

In the following, we show that Reduction 7.3 is sound, that is, if $H$ contains a vertex set with volume $\delta = \Omega(1/q)$ and expansion bounded away from $1 - \Gamma_{\rho^2}(\delta)/\delta$ (the minimum expansion of vertex sets of volume $\delta$ in the Gaussian noise graph $U_{\rho^2}$), then the graph $H$ contains a vertex set of volume $\varepsilon/\beta R$ and expansion bounded away from 1.

Let $\Omega$ denote the product probability space $[q] \times \{\perp, \top\}_\beta$ of the uniform distribution on $[q]$ and the $\beta$-biased distribution on $\{\perp, \top\}$. We can identify the vertex set of $H$ with $V^R \times \Omega^R$.

Let $T \subseteq V(H)$ be a vertex set in $H$. Let $f \colon V^R \times \Omega^R$ be its $\{0,1\}$-indicator function. We consider the following two symmetrizations of $f$,

$$\bar{f}(A, x, z) := \mathop{\mathbb{E}}_{\pi \in \mathbb{S}_R} f\Big(\pi.(A, x, z)\Big), \tag{7.2}$$

$$\tilde{f}(A, x, z) := \mathop{\mathbb{E}}_{(\tilde{A}, \tilde{x}) \sim M_z(A, x)} \bar{f}(\tilde{A}, \tilde{x}, z). \tag{7.3}$$

By the symmetries of the graph $H$, we have

$$\langle f, Hf \rangle = \langle \bar{f}, H\bar{f} \rangle = \langle \tilde{f}, H\tilde{f} \rangle.$$

We write $\tilde{f}_A(x, z) = \tilde{f}(A, x, z)$. For all $A \in V^R$, we define functions $g_A \colon \Omega^R \to [0, 1]$ as the average of $\tilde{f}_B$ over the neighbors $B$ of the vertex $A$ in the product graph $G^{\otimes R}$ (with additional noise),

$$g_A := \mathop{\mathbb{E}}_{B \sim (T_{1-\varepsilon} G)^{\otimes R}(A)} \tilde{f}_B.$$

The fraction of edges staying inside of $T$ equals the typical noise stability of $g_A$ in the graph $T_\rho$ on $\Omega^R$,

$$\begin{aligned} H(T, T) &= \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{(x,z) \sim \Omega^R} \left( \mathop{\mathbb{E}}_{B \sim (T_{1-\varepsilon} G)^{\otimes R}(A)} \mathop{\mathbb{E}}_{(y,w) \sim T_\rho(x,z)} \tilde{f}_B(y, w) \right)^2 \\ &= \mathop{\mathbb{E}}_{A \sim V^R} \mathop{\mathbb{E}}_{(x,z) \sim \Omega^R} \Big( T_\rho g_A(x, z) \Big)^2 = \mathop{\mathbb{E}}_{A \sim V^R} \| T_\rho g_A \|^2 . \end{aligned}$$

We can further upper bound $H(T, T)$ by combining the soundness analysis of Reduction 6.20 (in particular, the Invariance Principle and Influence Decoding) with the soundness analysis of Reduction 6.5. Combining these tools (which is non-trivial, but standard), establishes the following claim.

**Claim 7.8.** *There exists a constant $\zeta = \zeta(q, \varepsilon, \rho, \beta) > 0$ such that either*

$$H(T,T) \leqslant \mathop{\mathbb{E}}_{A\sim V^R} \Gamma_{\rho^2}(\mathbb{E}\, g_A) + \beta\,.$$

*or $G$ contains a vertex set with volume $\varepsilon/\beta R$ and expansion less than $1-\zeta$.*

Next, we show that for most vertex tuples $A \in V^R$, the expectation of the function $g_A$ (over $\Omega^R$) roughly equals the expectation of the function $f$. (Since $f$ indicates the vertex set $T$, its expectation is the volume of $T$.) To this end, we compare $\mathbb{E}_A(\mathbb{E}\, g_A)^2$ to $(\mathbb{E}\, f)^2$ in the following claim. (We remark that one could derive upper bounds on $\mathbb{E}_A(\mathbb{E}\, g_A)^2$ using properties of the graph $G^{\otimes R}$. In fact, we followed this approach in the soundness proof of Reduction 6.20. The key novelty of Reduction 7.3 is that we can control $\mathbb{E}_A(\mathbb{E}\, g_A)^2$ without further assumptions on the graph $G^{\otimes R}$.)

**Claim 7.9.**

$$\mathop{\mathbb{E}}_{A}(\mathbb{E}\, g_A)^2 \leqslant (\mathbb{E}\, f)^2 + \beta \|f\|^2\,.$$

*Proof.* Recall that $g_A$ is defined by averaging the function $\bar{f}$,

$$g_A(x,z) = \mathop{\mathbb{E}}_{B\sim(T_{1-\varepsilon}G)^{\otimes R}(A)} \; \mathop{\mathbb{E}}_{(\tilde{B},\tilde{x})\sim M_z(B,x)} \bar{f}(\tilde{B},\tilde{x},z)\,.$$

Using Cauchy–Schwarz, the typical value of the square of $\mathbb{E}\, g_A$ is at most

$$\begin{aligned}\mathop{\mathbb{E}}_{A}\Big(\mathbb{E}\, g_A\Big)^2 &= \mathop{\mathbb{E}}_{A\sim V^R}\Bigg(\mathop{\mathbb{E}}_{(x,z)\sim\Omega^R} \; \mathop{\mathbb{E}}_{B\sim(T_{1-\varepsilon}G)^{\otimes R}(A)} \; \mathop{\mathbb{E}}_{(\tilde{A},\tilde{x})\sim M_z(A,x)} \bar{f}(A,x,z)\Bigg)^2 \\ &\leqslant \mathop{\mathbb{E}}_{A\sim V^R,\; x\sim[q]^R} \underbrace{\Bigg(\mathop{\mathbb{E}}_{z\sim\{\bot,\top\}^R_\beta} \; \mathop{\mathbb{E}}_{(\tilde{A},\tilde{x})\sim M_z(A,x)} \bar{f}(\tilde{A},\tilde{x},z)\Bigg)^2}_{=:M\bar{f}(A,x)} = \|M\bar{f}\|^2\,.\end{aligned}$$

Notice that $M$ (as defined above) is a linear operator from the space $L_2(V^R \times [q]^R \times \{\bot,\top\}^R_\beta)$ to the space $L_2(V^R \times [q]^R)$.

To upper bound the norm $\|M\bar{f}\|$, we consider the spectral decomposition of the linear operator $M^*M$ on $L_2(V^R \times [q]^R \times \{\bot,\top\}^R_\beta)$. Here, $M^*$ is the adjoint of $M$, so that $\langle Mh, g\rangle = \langle h, M^* g\rangle$ for any two functions $h \in L_2(V^R \times [q]^R \times \{\bot,\top\}^R_\beta)$ and $g \in L_2(V^R \times [q]^R)$. One can verify that $M^*M$ is a Markov operator. Its largest eigenvalue is 1 and the constant functions form the corresponding eigenspace.

To show the claim, it is enough to verify that the second largest eigenvalue of $M^*M$ equals $\beta$, because

$$\|M\bar{f}\|^2 = \langle \bar{f}, M^*M\bar{f}\rangle \leqslant (\mathbb{E}\bar{f})^2 + \beta\|\bar{f} - \mathbb{E}\bar{f}\|^2 \leqslant (\mathbb{E} f)^2 + \beta\|f\|^2 .$$

(The first inequality uses that $\bar{f} - \mathbb{E}\bar{f}$ is orthogonal to the first eigenspace and thus contributes at most $\beta\|\bar{f} - \mathbb{E}\bar{f}\|^2$ to the inner product $\langle \bar{f}, M^*M\bar{f}\rangle$. The second inequality uses that $\bar{f}$ is obtained from $f$ by averaging, which implies that $\mathbb{E}\bar{f} = \mathbb{E} f$ and $\|\bar{f}\|^2 \leqslant \|f\|^2$.)

To compute the second largest eigenvalue of $M^*M$, we note that this operator has the same non-zero eigenvalues as the operator $MM^*$ (which acts on $L_2(V^R \times [q]^R)$). One can verify that $MM^*$ corresponds to the noise graph $T_\beta$ on $V^R \times [q]^R$. (See also (7.1).) Hence, the second largest eigenvalue of $M^*M$ equals $\beta$. □

We conclude the soundness analysis of Reduction 7.3 by combining the previous claims. Let $\gamma > 0$ be sufficiently small. Claim 7.9 allows us to upper bound the fraction of vertex tuples $A \in V^R$ with $\mathbb{E} g_A > \mathbb{E} f + \gamma$. Concretely, Chebyshev's inequality implies

$$\mathop{\mathbb{P}}_{A\sim V^R}\left\{\mathbb{E} g_A > \mathbb{E} f + \gamma\right\} \leqslant \mathop{\mathbb{E}}_{A\sim V^R}\left(\mathbb{E} g_A - \mathbb{E} f\right)^2/\gamma^2 \leqslant \beta/\gamma^2 .$$

Let us assume that $G$ contains no vertex set with volume $\varepsilon/\beta R$ and expansion less than $1-\zeta$. Then, using Claim 7.8, the following upper bound on $H(T,T)$ holds

$$H(T,T) \leqslant \mathop{\mathbb{E}}_{A\sim V^R} \Gamma_{\rho^2}(\mathbb{E} g_A) + \beta \leqslant \Gamma_{\rho^2}(\mathbb{E} f + \gamma) + \beta/\gamma^2 + \beta \leqslant \Gamma_{\rho^2}(\mathbb{E} f) + 2\gamma + 2\beta/\gamma^2 .$$

(In the last step, we use that the function $\Gamma_{\rho^2}$ is 2-Lipschitz.) To achieve the best upper bound on $H(T,T)$, we choose $\gamma = \beta^{1/3}$, which shows that $H(T,T) \leqslant \Gamma_{\rho^2}(\mathbb{E} f) + 3\beta^{1/3}$.

To summarize, we established the following soundness guarantee of Reduction 7.3.

**Lemma 7.10** (Soundness). *There exists a constant $\zeta = \zeta(q,\varepsilon,\rho,\beta) > 0$ such that if all vertex sets with volume $\varepsilon/\beta R$ in $G$ have expansion at least $1-\zeta$, then every vertex set $T$ in $H = H_{R,q,\varepsilon,\rho,\beta}(G)$ satisfies*

$$H(T,T) \leqslant \Gamma_{\rho^2}(\mu(T)) + 3\beta^{1/3}$$

## 7.3. Notes

The material of this chapter is based on an unpublished manuscript with Prasad Raghavendra and Madhur Tulsiani, titled "Reductions between Expansion Problems" [RST10b].

# Part III.

# Lower Bounds

# 8. Limits of Semidefinite Relaxations

Non-trivial consequences of a refutation of the Unique Games Conjecture (as we showed in the Chapter 6) are some form of evidence for the truth of the conjecture. Another form of evidence for the truth of the conjecture are lower bounds for Unique Games in restricted models of computation.

In the context of approximation problems, it is natural to study models of computations defined by hierarchies of relaxations (typically linear or semidefinite relaxations). Such hierarchies contain relaxations with gradually increasing complexity — from linear complexity to exponential complexity. Relaxations with higher complexity provide better approximations for the optimal value of optimization problems. (The relaxations with highest, i.e., exponential, complexity typically compute the optimal value exactly.)

Starting with the seminal work of Arora, Bollobás, and Lovász [ABL02], lower bounds on the complexity of many approximation problems were obtained in various hierarchies. In this chapter, we show the first super-polynomial lower bound for Unique Games in a hierarchy that captures the best known algorithms for all constraint satisfaction problems (see Chapter 8). Previously known lower bounds for Unique Games, considered only relaxations with fixed polynomial complexity [KV05], or considered hierarchies that do not capture current algorithms (in particular, the hierarchies provide only trivial approximation for Max Cut) [CMM09].

Our lower bounds translate to corresponding lower bounds for all UG-hard problems (via the known UG-hardness reductions). Even for specific UG-hard problem like Max Cut such lower bounds were not known before.

## 8.1. Overview

In this chapter, we exhibit an integrality gap for certain strong SDP relaxations of Unique Games. More precisely, we consider two strong hierarchies of SDP relaxations $\{\mathsf{LH}_r\}_{r\in\mathbb{N}}$ and $\{\mathsf{SA}_r\}_{r\in\mathbb{N}}$ (see Section 8.3 for the definitions). We give rough definitions of these relaxations here for the convenience of the reader. The $r^{\text{th}}$ level relaxation $\mathsf{LH}_r$ consists of the following: 1) SDP vectors for every vertex of the unique game, 2) All valid constraints on vectors corresponding to at most $r$ vertices. Equivalently, the $\mathsf{LH}_r$ relaxation consists of SDP vectors and local distributions $\mu_S$ over integral assignments to sets $S$ of at most $r$ variables, such that the second moments of local distributions $\mu_S$ match the corresponding inner products of SDP vectors.

The $\mathsf{SA}_r$ relaxation is a strengthening of $\mathsf{LH}_r$ with the additional constraint that for two sets $S, T$ of size at most $r$, the corresponding local distribution over integral assignments $\mu_S, \mu_T$ must have the same marginal distribution over $S \cap T$. The $\mathsf{SA}_r$ relaxation corresponds to simple SDP relaxation strengthened by $r^{\text{th}}$ round of Sherali-Adams hierarchy [SA90]. Let $\mathsf{LH}_r(\Phi)$ and $\mathsf{SA}_r(\Phi)$ denote the optimum value of the corresponding SDP relaxations on the instance $\Phi$. Further, let $\mathrm{opt}(\Phi)$ denote the value of the optimum labeling for $\Phi$. For the $\mathsf{LH}$ and $\mathsf{SA}$ hierarchies, we show:

**Theorem 8.1.** *For all constants $\eta > 0$, there exists a* Unique Games *instance $\Phi$ on $N$ vertices such that $\mathsf{LH}_r(\Phi) \geqslant 1 - \eta$ and $\mathrm{opt}(\Phi) \leqslant \eta$ for $r = O(2^{(\log\log N)^{\frac{1}{4}}})$*

**Theorem 8.2.** *For all constants $\eta > 0$, there exists a* Unique Games *instance $\Phi$ on $N$ vertices such that $\mathsf{SA}_r(\Phi) \geqslant 1 - \eta$ and $\mathrm{opt}(\Phi) \leqslant \eta$ for $r = O((\log\log N)^{\frac{1}{4}})$*

Demonstrated for the first time in [KV05], and used in numerous later works [CMM09, STT07b, Tul09, Rag08, GMR08, MNRS08], it is by now well known that integrality gaps can be composed with hardness reductions.

In particular, given a reduction $\mathsf{Red}$ from Unique Games to a certain problem $\Lambda$, on starting the reduction with a integrality gap instance $\Phi$ for Unique Games, the resulting instance $\mathsf{Red}(\Phi)$ is a corresponding integrality gap for $\Lambda$. Composing the integrality gap instance for $\mathsf{LH}_r$ or $\mathsf{SA}_r$ relaxation of Unique Games, along with UG reductions in [KKMO07, Aus07, Rag08, GMR08, MNRS08, RS09b], one can obtain integrality gaps for $\mathsf{LH}_r$ and $\mathsf{SA}_r$ relaxations for several important problems. For the sake of succinctness, we will state the following general theorem:

**Theorem 8.3.** *Let $\Lambda$ denote a problem in one of the following classes:*

- *A Generalized Constraint Satisfaction Problem*
- *An Ordering Constraint Satisfaction Problem*

*Let SDP denote the SDP relaxation that yields the optimal approximation ratio for $\Lambda$ under UGC. Then the following holds: Given an instance $\mathfrak{I}$ of the problem $\Lambda$, with SDP$(\mathfrak{I}) \geqslant c$ and $\mathrm{opt}(\mathfrak{I}) \leqslant s$, for every constant $\eta > 0$, there exists an instance $\Psi_\eta$ over $N$ variables such that:*

- $\mathsf{LH}_r(\Psi_\eta) \geqslant c - \eta$ *and* $\mathrm{opt}(\Psi_\eta) \leqslant s + \eta$ *with* $r = O(2^{(\log\log N)^{1/4}})$.
- $\mathsf{SA}_r(\Psi_\eta) \geqslant c - \eta$ *and* $\mathrm{opt}(\Psi_\eta) \leqslant s + \eta$ *with* $r = O((\log\log N)^{1/4})$.

The $O$ notation in the number of rounds hides a constant depending on $\eta$.

The classes of problems for which the above result holds include Max Cut [KKMO07], Max 2-Sat [Aus07], Grothendieck Problem [RS09b] $k$-way Cut [MNRS08] and Maximum Acyclic Subgraph [GMR08]. Notable exceptions that do not directly fall under this framework are Vertex Cover and Sparsest Cut.

Reductions from Unique Games to Sparsest Cut have been exhibited in [KV05] and [CKK+06]. With the integrality gap for $\mathsf{LH}_r$ relaxation of Unique Games (Theorem 8.1), these reductions imply a corresponding $\mathsf{LH}_r$ integrality gap for Sparsest Cut. Integrality gaps for Sparsest Cut are directly related to lower bounds for distortion required to embed a given metric into $L_1$ metric. Here the $L_1$ metric consists of points in $\mathbb{R}^d$ for arbitrarily large $d$, and the distance between two points $(x, y)$ is $\|x - y\|_1$. An $L_2^2$ metric consists of a set of points in $\mathbb{R}^d$ such that the squares of the distances between them also form a metric (satisfy triangle inequality). Restated in this language, the SDP vectors of the Sparsest Cut integrality gap that we construct, yield the following result:

**Theorem 8.4.** *For some absolute constants $\gamma, \delta > 0$, there exists an $N$-point $L_2^2$ metric that requires distortion at least $\Omega(\log\log N)^\delta$ to embedd into $L_1$, while every set of size at most $O(2^{(\log\log N)^\gamma})$ embedds isometrically into $L_1$.*

The Uniform Sparsest Cut problem is among the many important problems for which no Unique Games reduction is known. In [DKSV06], the techniques of [KV05] were extended to obtain an integrality gap for Uniform Sparsest Cut for the SDP with triangle inequalities. Roughly speaking, the SDP gap construction in [DKSV06] consists of the hypercube with its vertices identified by certain symmetries such as cyclic shift of the coordinates. Using the techniques from this chapter, the following SDP integrality gap for the Balanced Separator problem can be exhibited. The details of the proof of this theorem are omitted from the thesis.

**Theorem 8.5.** *For some absolute constants $\gamma, \delta > 0$, there exists an instance $G$ on $N$ vertices of* Balanced Separator *such that the ratio* $\mathrm{opt}(G)/\mathsf{LH}_r(G) \geqslant \Omega(\log\log N)^{\delta}$ *for* $r = O(\log\log N)^{\gamma}$.

### 8.1.1. Techniques

In this section, we will present a brief overview of the techniques and a roadmap for the rest of the chapter.

The overall strategy in this work to construct SDP integrality gaps is along the lines of Khot–Vishnoi [KV05]. Let us suppose we wish to construct a SDP integrality gap for a problem $\Lambda$ (say Max Cut). Let $\mathsf{Red}_\Lambda$ be a reduction from Unique Games to the problem $\Lambda$. The idea is to construct a SDP integrality gap $\Phi$ for Unique Games, and then execute the reduction $\mathsf{Red}_\Lambda$ on the instance $\Phi$, to obtain the SDP gap construction $\mathsf{Red}_\Lambda(\Phi)$. Surprisingly, as demonstrated in [KV05], the SDP vector solution for $\Phi$ can be transformed through the reduction to obtain the SDP solution for $\mathsf{Red}_\Lambda(\Phi)$.

Although this technique has been used extensively in numerous works [CMM09, STT07a, Tul09, Rag08, GMR08, MNRS08] since [KV05], there is a crucial distinction between [KV05] and later works. In all other works, starting with an SDP gap $\Phi$ for Unique Games, one obtains an integrality gap for an SDP relaxation that is no stronger. For instance, starting with a integrality gap for 10-rounds of a SDP hierarchy, the resulting SDP gap instance satisfies at most 10 rounds of the same hierarchy.

The surprising aspect of [KV05], is that it harnesses the UG reduction $\mathsf{Red}_\Lambda$ to obtain an integrality gap for a "stronger" SDP relaxation than the one which it stared with. Specifically, starting with an integrality gap $\Phi$ for a simple SDP relaxation of Unique Games, [KV05] exhibit an SDP gap for Max Cut which obeys all valid constraints on 3 variables. The proof of this fact (the triangle inequality) is perhaps the most technical and least understood aspect about [KV05]. One of the main contributions of this chapter is to conceptualize and simplify this aspect of [KV05]. Armed with the understanding of [KV05], we then develop the requisite machinery to extend it to a strong SDP integrality gap for Unique Games.

To obtain strong SDP gaps for Unique Games, we will apply the above strategy on the reduction from Unique Games to $\mathsf{E2Lin}_q$ obtained in [KKMO07]. Note that $\mathsf{E2Lin}_q$ is a special case of Unique Games. Formally, we show the following reduction from a *weak gap* instance for Unique Games over a large alphabet to a integrality gap for a strong SDP relaxation of $\mathsf{E2Lin}_q$.

**Theorem 8.6.** *(Weak Gaps for* Unique Games $\implies$ *Strong gaps for* $\mathsf{E2Lin}_q$*)*
*For a positive integer* $q$*, let* $\mathsf{Red}_{\mathsf{E2Lin}_q}$ *denote the reduction from* Unique Games *to* $\mathsf{E2Lin}_q$*. Given a* $(1-\eta,\delta)$-weak gap *instance* $\Phi$ *for* Unique Games*, the* $\mathsf{E2Lin}_q$ *instance* $\mathsf{Red}_{\mathsf{E2Lin}_q}(\Phi)$ *is a* $(1-2\gamma, 1/q^{\gamma/2}+o_\delta(1))$ *SDP gap for the relaxation* $\mathsf{LH}_r$ *for* $r = 2^{O(1/\eta^{1/4})}$*. Further,* $\mathsf{Red}_{\mathsf{E2Lin}_q}(\Phi)$ *is a* $(1-\gamma,\delta)$ *SDP gap for the relaxation* $\mathsf{SA}_r$ *for* $r = O(1/\eta^{1/4})$.

Using the weak gap for Unique Games constructed in [KV05], along with the above theorem, implies Theorems 8.1 and 8.2. As already pointed out, by now it is fairly straightforward to compose an $r$-round integrality gap for Unique Games, with reductions to obtain a $r$ round integrality gaps for other problems. Hence, Theorem 8.3 is a fairly straightforward consequence of Theorems 8.1 and 8.2.

### 8.1.2. Organization

In the next section, we present a detailed proof overview that describes the entire integrality gap construction restricted to the case of Max Cut. The formal definitions of the SDP hierarchies $\mathsf{LH}_r$,$\mathsf{SA}_r$ and their robustness are presented in Section 8.3. We formally define weak gap instances for Unique Games in Section 8.4. We also outline an alternate integrality gap for a very minimal SDP relaxation of Unique Games in the same section. This section is followed by the description of the integrality gap instance for $\mathsf{E2Lin}_q$ obtained by reduction of Khot et al. [KKMO07]. In the rest of the chapter, we construct SDP vectors and local distributions to show that this is an integrality gap for the strong SDP relaxations – $\mathsf{LH}_r$ and $\mathsf{SA}_r$. The two subsequent sections are devoted to developing the requisite machinery of integral vectors, their tensor products and local distributions for Unique Games. The SDP vectors and local distributions for the integrality gap instance described in Section 8.5 are exhibited in Sections Section 8.8 and Section 8.8.2.

## 8.2. Proof Sketch

For the sake of exposition, we will describe the construction of an SDP integrality gap for Max Cut. To further simplify matters, we will exhibit an integrality gap for the basic Goemans-Williamson relaxation, augmented with the triangle inequalities on every three vectors. While an integrality gap of this nature is already part of the work of Khot–Vishnoi [KV05], our proof will be conceptual and amenable to generalization.

Let $\Phi$ be a SDP integrality gap for UNIQUE GAMES on an alphabet $[R]$. For each vertex $B$ in $\Phi$, the SDP solution associates $R$ orthogonal unit vectors $B = \{b_1, \cdots, b_R\}$. For the sake of clarity, we will refer to a vertex $B$ in $\Phi$ and the set of vectors $B = \{b_1, \ldots, b_R\}$ associated with it as a "cloud". The clouds satisfy the following properties:

- (Matching Property) For every two clouds $A, B$, there is a unique matching $\pi_{B \leftarrow A}$ along which the inner product of vectors between $A$ and $B$ is maximized. Specifically, if $\rho(A,B) = \max_{a \in A, b \in B} \langle a, b \rangle$, then for each vector $a$ in $A$, we have $\langle a, \pi_{B \leftarrow A}(a) \rangle = \rho(A,B)$.

- (High objective value) For most edges $e = (A,B)$ in the UNIQUE GAMES instance $\Phi$, the maximal matching $\pi_{A \leftarrow B}$ is the same as the permutation $\pi_e$ corresponding to the edge, and $\rho(A,B) \approx 1$.

Let $\mathsf{Red}_{\text{MAX CUT}}(\Phi)$ be the MAX CUT instance obtained by executing the reduction in [KKMO07] on $\Phi$. The reduction $\mathsf{Red}_{\text{MAX CUT}}$ in [KKMO07] introduces a long code ($2^R$ vertices indexed by $\{-1,1\}^R$) for every cloud in $\Phi$. Hence the vertices of $\mathsf{Red}_{\text{MAX CUT}}(\Phi)$ are given by pairs $(B, x)$ where $B$ is a cloud in $\Phi$ and $x \in \{-1,1\}^R$.

The SDP vectors we construct for the integrality gap instance resemble (somewhat simpler in this work) the vectors in [KV05]. Roughly speaking, for a vertex $(B,x)$, we associate an SDP vector $\boldsymbol{V}^{B,x}$ defined as follows:

$$\boldsymbol{V}^{B,x} = \frac{1}{\sqrt{R}} \sum_{i \in [R]} x_i b_i^{\otimes t}$$

The point of departure from [KV05] is the proof that the vectors form a feasible solution for the stronger SDP. Instead of directly showing that the inequalities hold for the vectors, we exhibit a distribution over integral assignments whose second moments match the inner products. Specifically, to show that triangle inequality holds for three vertices $S = \{(A,x),(B,y),(C,z)\}$, we will exhibit a $\mu_S$ distribution over $\{\pm 1\}$ assignments to the three vertices, such that

$$\mathop{\mathbb{E}}_{\{Y^{A,x}, Y^{B,y}, Y^{C,z}\} \sim \mu_S} [Y^{A,x} Y^{B,y}] = \langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle$$

The existence of an integral distribution matching the inner products shows that the vectors satisfy all valid inequalities on the three variables, including the triangle inequality. We shall construct the distribution $\mu_S$ over local assignments in three steps,

**Local Distributions over Labelings for Unique Games.** For a subset of clouds $\mathcal{S}$ within the Unique Games instance $\Phi$, we will construct a distribution $\mu_{\mathcal{S}}$ over labelings to the set $\mathcal{S}$. The distribution $\mu_{\mathcal{S}}$ over $[R]^{\mathcal{S}}$ will be "consistent" with the SDP solution to $\Phi$. More precisely, if two clouds $A$ and $B$ are *highly correlated* ($\rho(A,B) \approx 1$), then when the distribution $\mu_{\mathcal{S}}$ assigns label $\ell$ to $A$, with high probability it assigns the corresponding label $\pi_{B\leftarrow A}(\ell)$ to $B$. Recall that $\rho(A,B)$ was defined as $\max_{a\in A, b\in B}\langle a,b\rangle$.

Consider a set $\mathcal{S}$ where every pair of clouds $A, B$ are *highly correlated* ($\rho(A,B) \geqslant 0.9$). We will refer to such a set of clouds as *Consistent*. For a *Consistent* set $\mathcal{S}$,assigning a label $\ell$ for a cloud $A$ in $\mathcal{S}$, forces the label of every other cloud $B$ to $\pi_{B\leftarrow A}(\ell)$. Furthermore, it is easy to check that the resulting labeling satisfies consistency for every pair of clouds in $\mathcal{S}$. (see Lemma 8.39 for details) Hence, in this case, the distribution $\mu_{\mathcal{S}}$ could be simply obtained by picking the label $\ell$ for an arbitrary cloud in $\mathcal{S}$ uniformly at random, and assigning every other cloud the induced label.

Now consider a set $\mathcal{S}$ which is not consistent. Here the idea is to decompose the set of clouds $\mathcal{S}$ into clusters, such that each cluster is consistent. Given a decomposition, for each cluster the labeling can be independently generated as described earlier. In this chapter, we will use a geometric decomposition to decompose the set of clouds $\mathcal{S}$ into clusters. The crucial observation is that the correlations $\rho(A,B)$ for clouds $A,B \in \mathcal{S}$, can be approximated well by a certain $L_2^2$ metric. More precisely, for each cloud $A$, we can associate a unit vector $\boldsymbol{v}_A = \sum_{a\in A} a^{\otimes s}$ such that the $L_2^2$ distance between $\boldsymbol{v}_A, \boldsymbol{v}_B$ is a good approximation of the quantity $1-\rho(A,B)$.

By using $t$ random halfspace cuts on this geometric representation, we obtain a partition into $2^t$ clusters. A pair of clouds $A, B$ that are not highly correlated ($\rho(A,B) < 1 - 1/16$), are separated by the halfspaces with probability at least $1-(3/4)^t$. Hence for a large enough $t$, all resulting clusters are consistent with high probability. (see Lemma 8.42).

A useful feature of the geometric clustering is that for two subsets $\mathcal{T} \subset \mathcal{S}$, the distribution over labelings $\mu_{\mathcal{T}}$ is equal to the marginal of the distribution $\mu_{\mathcal{S}}$ on $\mathcal{T}$. To see this, observe that the distribution over clusterings depends solely on the geometry of the associated vectors. On the downside, the geometric clustering produces inconsistent clusters with a very small but non-zero probability. (see Corollary 8.44).

The details of the construction of local distributions to Unique Games are presented in 8.7.

**Constructing Approximate Distributions.** Fix a set $S \subseteq \mathcal{S} \times \{\pm 1\}^R$ of vertices in the Max Cut instance $\mathsf{Red}_{\text{Max Cut}}(\Phi)$. We will now describe the construction of the local integral distribution $\mu_S$.

In the reduction $\mathsf{Red}_{\text{Max Cut}}$, the labeling $\ell$ to a cloud $B$ in the Unique Games instance is encoded as choosing the $\ell$th dictator cut in the long code corresponding to cloud $B$. Specifically, assigning the label $\ell$ to a cloud $B$ should translate into assigning $x_\ell$ for every vertex $(B, x)$ in the long code of $B$. Hence, a straightforward approach to define the distribution $\mu_S$ would be the following:

- Sample a labeling $\ell : \mathcal{S} \to [R]$ from the distribution $\mu_{\mathcal{S}}$,
- For every vertex $(B, x) \in \mathcal{S}$, assign $x_{\ell(B)}$.

Although inspired by this, our actual construction of $\mu_S$ is slightly more involved. First, we make the following additional assumption regarding the Unique Games instance $\Phi$:

> *Assumption*: All the SDP vectors for the integrality gap instance $\Phi$ are $\{\pm 1\}$-vectors (have all their coordinates from $\{\pm 1\}$).

The SDP gap instance for Unique Games constructed in [KV05] satisfies this additional requirement. Furthermore, we outline a generic transformation to convert an arbitrary Unique Games SDP gap into one that satisfies the above property (see Observation 8.25). A $\{\pm 1\}$-vector is to be thought of as a distribution over $\{\pm 1\}$ assignments. It is easy to see that tensored powers of $\{\pm 1\}$-vectors yield $\{\pm 1\}$-vectors. Let $T$ denote the number of coordinates in the vectors $\boldsymbol{V}^{B,x}$. The distribution $\mu_S$ is defined as follows,

- Sample a labeling $\ell : \mathcal{S} \to [R]$ from the distribution $\mu_{\mathcal{S}}$, and a coordinate $i \in [T]$ uniformly at random.
- For every vertex $(B, x) \in S$, assign $Y^{B,x}$ to be the $i$th coordinate of the vector $x_{\ell(B)} b_{\ell(B)}^{\otimes t}$.

We will now argue that the first two moments of the local distributions $\mu_S$ defined above, approximately match the corresponding inner products between SDP vectors.

Consider the inner product $\langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle$ of the SDP vectors corresponding to some pair of vertices $(A,x)$ and $(B,y)$ in $S$. The inner product consists of $R^2$ terms of the form $\langle x_i a_i^{\otimes t}, y_j b_j^{\otimes t} \rangle$. The crucial observation we will utilize is that the inner product $\langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle$ is approximately determined by the $R$ terms corresponding to the matching $\pi_{B \leftarrow A}$. In other words, we have

$$\langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle \approx \frac{1}{n} \sum_{\ell \in [R]} x_\ell y_{\pi_{B\leftarrow A}(\ell)} \langle a_\ell^{\otimes t}, b_{\pi_{B\leftarrow A}(\ell)}^{\otimes t} \rangle \qquad \leqslant \quad \rho(A,B)^t$$

(see Section 8.4.1 for details)

If $\rho(A,B) < 0.9$, then with high probability the clustering would place the clouds $A,B$ in different clusters. Hence the labels assigned to $A,B$ would be completely independent of each other, and so would the assignments to $(A,x)$ and $(B,y)$. Hence, we would have $\mathbb{E}[Y^{A,x} Y^{B,y}] = 0$. On the other hand, by the above inequality the inner product $\langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle \leqslant 0.9^t \approx 0$. Therefore, for clouds $A,B$ that are not *highly correlated*, the inner product of vectors $\boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y}$ agree approximately with the distribution over local assignments.

At the other extreme, if $\rho(A,B) \approx 1$, then with high probability the clustering would not separate $A$ from $B$. If $A,B$ are not separated, then the distribution $\mu_S$ over labelings will respect the matching between $A$ and $B$. Specifically, whenever $A$ is assigned label $\ell$ by $\mu_S$, with high probability $B$ is assigned the label $\pi_{B\leftarrow A}(\ell)$. Consequently, in this case we have

$$\mathop{\mathbb{E}}_{\mu_S}[Y^{A,x} Y^{B,y}] = \frac{1}{n} \sum_{\ell \in [R]} \langle x_\ell a_\ell^{\otimes t}, y_{\pi_{B\leftarrow A}(\ell)} b_{\pi_{B\leftarrow A}(\ell)}^{\otimes t} \rangle \approx \langle \boldsymbol{V}^{A,x}, \boldsymbol{V}^{B,y} \rangle$$

**Smoothing.** In Section 8.9, we show a robustness property for the $\mathsf{LH}_r$ and $\mathsf{SA}_r$ relaxations by which approximately feasible solutions to these hierarchies can be converted (*smoothed*) into perfectly feasible solutions with a small loss in the objective value.

To illustrate the idea behind the robustness, consider a set of unit vectors $\{v_i\}_{i=1}^R$ that satisfy all triangle inequalities up to an additive error of $\varepsilon$, i.e.,

$$\|v_i - v_j\|^2 + \|v_j - v_k\|^2 - \|v_i - v_k\|^2 \geqslant -\varepsilon$$

We include formal statements of the claims about the robustness of solutions to $\mathsf{LH}_r$ and $\mathsf{SA}_r$ (Theorems 8.11, 8.10) in Section 8.3. We refer the reader to Section 8.9 for the proofs of these claims.

**Extending to $\mathsf{E2Lin}_q$.** The above argument for Max Cut can be made precise. However, to obtain an SDP gap for larger number of rounds, we use a slightly more involved construction of SDP vectors.

$\{\pm 1\}$-vectors were natural in the above discussion, since Max Cut is a CSP over $\{0,1\}$. For $\mathsf{E2Lin}_q$, it is necessary to work with vectors whose coordinates are from $\mathbb{F}_q$, as opposed to $\{\pm 1\}$. The tensoring operation for $\mathbb{F}_q$-integral vectors is to be appropriately defined to ensure that while the behaviour of the inner products resemble traditional tensoring, the tensored vectors are $\mathbb{F}_q$-integral themselves (see Section 8.6 for details).

For the case of Max Cut, we used a gap instance $\Phi$ for Unique Games all of whose SDP vectors where $\{\pm 1\}$-vectors. In case of $\mathsf{E2Lin}_q$, the SDP vectors corresponding to the Unique Games instance $\Phi$ would have to be $\mathbb{F}_q$-integral vectors. We outline a generic transformation to convert an arbitrary Unique Games SDP gap into one that satisfies this property (see Observation 8.35).

## 8.3. Hierarchies of SDP Relaxations

In this section, we define the $\mathsf{LH}_r$ and $\mathsf{SA}_r$ hierarchies and formally state their robustness properties.

### 8.3.1. $\mathsf{LH}_r$-Relaxation

Let $\mathfrak{I}$ be a CSP (say Unique Games) instance over a set of variables $\mathcal{V}$, alphabet size $q$ and arity $k$. A feasible solution to the $\mathsf{LH}_r$ relaxation consists of the following:

1. A *collection of (local) distributions* $\{\mu_S\}_{S\subseteq\mathcal{V},|S|\leqslant r}$, where $\mu_S : [q]^S \to \mathbb{R}_+$ is a distribution over $[q]$-assignments to $S$, that is, $\mu_S \in \Delta([q]^S)$.

2. A *(global) vector solution* $\{\boldsymbol{b}_{i,a}\}_{i\in\mathcal{V},a\in[q]}$, where $\boldsymbol{b}_{i,a} \in \mathbb{R}^d$ for every $i \in \mathcal{V}$ and $a \in [q]$.

**$\mathsf{LH}_r$-Relaxation.**

$$\text{maximize} \quad \mathop{\mathbb{E}}_{P\sim\mathcal{P}} \mathop{\mathbb{E}}_{x\sim\mu_P} P(x) \tag{$\mathsf{LH}_r$}$$

$$\text{subject to} \quad \langle \boldsymbol{b}_{i,a}, \boldsymbol{b}_{j,b}\rangle = \mathop{\mathbb{P}}_{x\sim\mu_S}\left\{x_i = a, x_j = b\right\} \qquad S\subseteq\mathcal{V}, |S|\leqslant r,\ i,j\in S,\ a,b\in[q], \tag{8.1}$$

$$\mu_S \in \Delta\left([q]^S\right) \tag{8.2}$$

Here, $\Delta\left([q]^S\right)$ denotes probability distributions over $[q]^S$. As usual, we denote by $\mathsf{LH}_r(\mathfrak{I})$ the value of an optimal solution to this relaxation.

The above relaxation succinctly encodes all possible inequalities on up to $r$ vectors. The next remark makes this observation precise.

**Remark 8.7.** A linear inequality on the inner products of a subset of vectors $\{\boldsymbol{b}_{i,a}\}_{i\in S, a\in[q]}$ for $S\subseteq\mathcal{V}$ is *valid* if it inequality if it holds for all distributions over $[q]$-assignments to the variables $S$. A feasible solution to the $\mathsf{LH}_r$-relaxation satisfies all valid inequalities on sets of up to $r$ vectors.

### 8.3.2. $\mathsf{SA}_r$-Relaxation

Enforcing consistency between the marginals of the local distributions yields the $\mathsf{SA}_r$-relaxation.

**$\mathsf{SA}_r$-Relaxation:**

$$\text{maximize} \quad \mathop{\mathbb{E}}_{P\sim\mathcal{P}} \mathop{\mathbb{E}}_{x\sim\mu_P} P(x) \tag{$\mathsf{SA}_r$}$$

$$\text{subject to} \quad \langle \boldsymbol{b}_{i,a}, \boldsymbol{b}_{j,b}\rangle = \mathop{\mathbb{P}}_{x\sim\mu_S}\left\{x_i = a, x_j = b\right\} \qquad S\subseteq\mathcal{V},\ |S|\leqslant r,\ i,j\in S,\ a,b\in[q], \tag{8.3}$$

$$\left\|\operatorname{margin}_{A\cap B}\mu_A - \operatorname{margin}_{A\cap B}\mu_B\right\|_1 = 0 \qquad A,B\subseteq\mathcal{V},\ |A|,|B|\leqslant r. \tag{8.4}$$

$$\mu_S \in \Delta\left([q]^S\right) \tag{8.5}$$

**Remark 8.8.** The $\mathsf{SA}_r$ relaxation is closely related to the $r^{\text{th}}$ level of the Sherali–Adams hierarchy. In fact, $\mathsf{SA}_r$ is obtained from the basic SDP relaxation by $r$-rounds Sherali–Adams lift-and-project.

### 8.3.3. Robustness

Here we state the formal claims regarding robustness for the hierarchies $\mathsf{LH}_r$ and $\mathsf{SA}_r$. The proofs are in Section 8.9.

**Definition 8.9.** An SDP solution $\{\boldsymbol{v}_{i,a}\}_{i\in\mathcal{V},a\in\mathbb{F}_q}$, $\{\mu_S\}_{S\subseteq\mathcal{V},|S|\leqslant r}$ is said to be $\varepsilon$-*infeasible for* $\mathsf{LH}_r$ *(or* $\mathsf{SA}_r$*)* if it satisfies all the constraints of the program up to an additive error of $\varepsilon$.

**Theorem 8.10.** *Given an $\varepsilon$-infeasible solution $\{\boldsymbol{b}_{i,a}\}_{i\in\mathcal{V},a\in\mathbb{F}_q}$, $\{\mu_S\}_{S\subseteq\mathcal{V},|S|\leqslant r}$ to the* $\mathsf{LH}_r$ *relaxation, there exists a feasible solution $\{\boldsymbol{b}'_{i,a}\}, \{\mu'_S\}_{S\subset\mathcal{V},|S|\leqslant r}$ for* $\mathsf{LH}_r$ *such that for all subsets $S\subseteq\mathcal{V}, |S|\leqslant r$, $\|\mu_S-\mu'_S\|_1\leqslant \operatorname{poly}(q)\cdot r^2\varepsilon$.*

**Theorem 8.11.** *Given an $\varepsilon$-infeasible solution $\{\boldsymbol{b}_{i,a}\}_{i\in\mathcal{V},a\in\mathbb{F}_q}$, $\{\mu_S\}_{S\subseteq\mathcal{V},|S|\leqslant r}$ to the* $\mathsf{SA}_r$ *relaxation, there exists a feasible solution $\{\boldsymbol{b}'_{i,a}\}, \{\mu'_S\}_{S\subseteq\mathcal{V},|S|\leqslant r}$ for* $\mathsf{SA}_r$ *such that for all subsets $S\subseteq\mathcal{V}, |S|\leqslant r$, $\|\mu_S-\mu'_S\|_1\leqslant \operatorname{poly}(q)\cdot\varepsilon\cdot q^r$.*

## 8.4. Weak Gaps for Unique Games

We refer to an integrality gap instance for a fairly simple SDP relaxation of Unique Games as a *weak gap* instance. Formally, a *weak gap* instance for Unique games is defined as follows.

**Definition 8.12.** *(Weak SDP solutions and weak gap instances)* Let $\Upsilon = (V, E, \{\pi_e\colon [n]\to[n]\}_{e\in E})$. We say a collection $\mathcal{B}=\{B_u\}_{u\in V}$ is a *weak SDP solution of value* $1-\eta$ for $\Upsilon$ if the following conditions hold:

1. (**Orthonormality**) For every vertex $u\in V$, the collection $\mathcal{B}$ contains an ordered set $B_u=\{b_{u,1},\dots,b_{u,n}\}$ of $n$ orthonormal vectors in $\mathbb{R}^d$.
2. (**$\ell_2^2$-triangle inequality**) Any two vectors in $\bigcup\mathcal{B}$ have non-negative inner product and any three vectors in $\bigcup\mathcal{B}$ satisfy the $\ell_2^2$-triangle inequality ($\|x-y\|^2\leqslant\|x-z\|^2+\|z-y\|^2$).
3. (**Strong Matching Property**) For every pair of vertices $u,v\in V$, the sets $B_u$ and $B_v$ satisfy the following *strong matching property*: There exists $n$ disjoint matchings between $B_u, B_v$ given by bijections $\pi^{(1)},\dots,\pi^{(n)}\colon B_u\to B_v$ such that for all $i\in[n], b, b'\in B_u$, we have $\langle b,\pi^{(i)}(b)\rangle=\langle b',\pi^{(i)}(b')\rangle$.
4. (**High SDP value**) For every edge $e=(u,v)\in E$, the vector sets $B_u$ and $B_v$ have significant correlation under the permutation $\pi=\pi_e$. Specifically, $\langle b_{u,\ell}, b_{v,\pi(\ell)}\rangle^2\geqslant 0.99$ for all $\ell\in[n]$.

5. The collection $\mathcal{B}$ of orthonormal sets is a good SDP solution for $\Upsilon$, in the sense that

$$\mathop{\mathbb{E}}_{v\in V} \mathop{\mathbb{E}}_{\substack{w,w'\in N(v)\\ \pi=\pi_{w,v},\ \pi'=\pi_{w',v}}} \frac{1}{n}\sum_{\ell\in[n]} \langle b_{w,\pi(\ell)}, b_{w',\pi'(\ell)}\rangle \geqslant 1-\eta.$$

We say that $\Upsilon$ is a *weak* $(1-\eta,\delta)$*-gap instance* of Unique Games if $\Upsilon$ has a weak SDP solution of value $1-\eta$ and no labeling for $\Upsilon$ satisfies more than a $\delta$ fraction of the constraints.

**Remark 8.13.** The weak gap instances defined here are fairly natural objects. In fact, if $\mathfrak{I}$ is an instance of Γ-Max-2Lin($R$) with $\mathrm{sdp}(\mathfrak{I}) \geqslant 1-\eta$ and $\mathrm{opt}(\mathfrak{I}) \leqslant \delta$, it is easy to construct a corresponding weak gap instance $\mathfrak{I}'$. The idea is to start with an optimal SDP solution for $\mathfrak{I}$, symmetrize it (with respect to the group $\Phi$), and delete all edges of $\mathfrak{I}$ that contribute less than $\sqrt{3/4}$ to the SDP objective.

We observe the following consequence of Fact 8.18 and item 4 of Definition 8.12.

**Observation 8.14.** *If $\mathcal{B} = \{B_u\}_{u\in V}$ is a weak SDP solution for $\Phi = (V,E,\{\pi_e\}_{e\in E})$, then for any two edges $(w,v),(w',v)\in E$, the two bijections $\pi = \pi^{-1}_{(w',v)} \circ \pi_{(w,v)}$ and $\pi_{B_{w'}\leftarrow B_w}$ (see Def. 8.17) give rise to the same matching between the vector sets $B_w$ and $B_{w'}$,*

$$\pi(i) = j \quad\Longleftrightarrow\quad \pi_{B_{w'}\leftarrow B_w}(b_{w,i}) = b_{w',j}\,.$$

The previous observation implies that in a weak gap instance $\Phi$ the collection of permutations $\{\pi_e\}_{e\in E}$ is already determined by the geometry of the vector sets in a weak SDP solution $\mathcal{B}$.

There are a few explicit constructions of weak gap instances of Unique Games, most prominently the Khot–Vishnoi instance [KV05]. In particular, the following observation is a restatement of Theorem 9.2 and Theorem 9.3 in [KV05].

**Observation 8.15.** *For all $\eta,\delta > 0$, there exists a weak $(1-\eta,\delta)$-gap instance with $2^{2^{O(\log(1/\delta)/\eta)}}$ vertices.*

### 8.4.1. Properties of Weak Gap Instances

Observation 8.35 implies that without much loss we can assume that a weak SDP solution is $\mathbb{F}_q$*-integral*, that is, all vectors are $\mathbb{F}_q$-integral. Here we use again $\langle\cdot,\cdot\rangle_\psi := \langle\psi(\cdot),\psi(\cdot)\rangle$ as inner product for $\mathbb{F}_q$-integral vectors.

**Lemma 8.16.** *Let $\Phi = (V, E, \{\pi_e\}_{e\in E})$ be a weak $(1-\eta,\delta)$-gap instance. Then, for every $q \in \mathbb{N}$, we can find a weak $\mathbb{F}_q$-integral SDP solution of value $1 - O(\sqrt{\eta \log q})$ for a* Unique Games *instance $\Phi'$ which is obtained from $\Phi$ by deleting $O(\sqrt{\eta \log q})$ edges.*

*Proof.* Let $\mathcal{B}$ be a weak SDP solution for $\Phi$ of value $1-\eta$. By applying the transformation from Observation 8.35 to the vectors in $\mathcal{B}$, we obtain a collection $\mathcal{B}' = \{B'_u\}_{u\in V}$ of sets of $\mathbb{F}_q$-integral vectors. For every $u \in V$, the vectors in $B'_u$ are orthonormal. Furthermore, any two sets $B'_u, B'_v$ in $\mathcal{B}'$ satisfy the strong matching property (using the facts that the original sets $B_u, B_v$ satisfy this property and that $\langle b'_{u,i}, b'_{v,j}\rangle_\psi$ is a function of $\langle b_{u,i}, b_{v,j}\rangle$).

Let $\eta_{v,w,w',\ell} = 1 - \langle b_{w,\pi(\ell)}, b_{w',\pi'(\ell)}\rangle$. Using Jensen's inequality, we can verify that the value of the SDP solution $\mathcal{B}'$ is high,

$$\begin{aligned}
&\mathop{\mathbb{E}}_{v\in V} \mathop{\mathbb{E}}_{\substack{w,w'\in N(v)\\ \pi=\pi_{w,v},\ \pi'=\pi_{w',v}}} \frac{1}{R}\sum_{\ell\in[R]} \langle {b'}_{w,\pi(\ell)}, {b'}_{w',\pi'(\ell)}\rangle_\psi \\
&\qquad \geqslant \mathop{\mathbb{E}}_{v\in V} \mathop{\mathbb{E}}_{\substack{w,w'\in N(v)\\ \pi=\pi_{w,v},\ \pi'=\pi_{w',v}}} \frac{1}{R}\sum_{\ell\in[R]} 1 - O\left(\sqrt{\eta_{v,w,w',\ell}\log q}\right) \qquad \text{(by Obs. 8.35)}\\
&\qquad \geqslant 1 - O(\sqrt{\eta \log q}) \qquad \text{(using Jensen's inequality).}
\end{aligned}$$

So far, we verified that $\mathcal{B}'$ satisfies all requirements of a weak SDP solution besides item 4 of Definition 8.12. We can ensure that this condition is also satisfied by deleting all edges from $E$ where the condition is violated. Using standard averaging arguments, it is easy to see that the matching property and the high SDP value imply that this condition is satisfied for all but at most an $O(\sqrt{\eta \log q})$ fraction of edges. □

We will refer to the set of orthonormal vectors associated with a vertex $B$ as a *cloud*. In what follows, we identify the vertices $B$ in a *weak* gap instance with their corresponding clouds, and thus refer to vertices/clouds interchangeably.

**Definition 8.17.** For $A, B \in \mathcal{B}$, we denote

$$\rho(A,B) \stackrel{\text{def}}{=} \max_{a\in A, b\in B} |\langle a, b\rangle|.$$

We define $\pi_{B\leftarrow A}\colon A \to B$ to be any[1] bijection from $A$ to $B$ such that $|\langle a, \pi_{B\leftarrow A}(a)\rangle| = \rho(A,B)$ for all $a \in A$.

[1] The matching property asserts that such a matching exists. If it is not unique, we pick an arbitrary one. We will assume $\pi_{A\to B} = \pi^{-1}_{B\to A}$.

As a direct consequence of the orthogonality of the clouds in $\mathcal{B}$, we have the following fact about the uniqueness of $\pi_{B \leftarrow A}$ for highly correlated clouds $A, B \in \mathcal{B}$.

**Fact 8.18.** *Let $A, B \in \mathcal{B}$. If $\rho(A,B)^2 > 3/4$, then there exists exactly one bijection $\pi \colon A \to B$ such that $|\langle a, \pi(a)\rangle| = \rho(A,B)$ for all $a \in A$.*

**Remark 8.19.** The collection $\mathcal{B}$ succinctly encodes a Unique Games instance. For a graph $G = (\mathcal{B}, E)$ on $\mathcal{B}$, the goal is to find a labeling $\{\ell_A \in A\}_{A \in \mathcal{B}}$ (a labeling can be seen as a system of representatives for the clouds in $\mathcal{B}$) so as to maximize the probability

$$\mathop{\mathbb{P}}_{(A,B)\in E} \left\{ \ell_A = \pi_{A \leftarrow B}(\ell_B) \right\}.$$

### Tensoring

**Lemma 8.20.** *For $t \in \mathbb{N}$ and every pair of clouds $A, B \in \mathcal{B}$,*

$$\frac{1}{n} \sum_{\substack{a \in A, b \in B \\ a \neq \pi_{A \leftarrow B}(b)}} |\langle a, b\rangle|^t \leqslant 2 \cdot (3/4)^{t/2}.$$

*Proof.* By near-orthogonality, $\sum_{a \in B} \langle a, b\rangle^2 \leqslant 3/2$ for every $b \in B$. Hence, $\langle a, b\rangle^2 \leqslant 3/4$ for all $a \neq \pi_{A \leftarrow B}(b)$. Thus,

$$\frac{1}{n} \sum_{\substack{a \in A, b \in B \\ a \neq \pi_{A \leftarrow B}(b)}} |\langle a, b\rangle|^t \leqslant (3/4)^{\frac{t-2}{2}} \cdot \frac{1}{n} \sum_{a \in A, b \in B} |\langle a, b\rangle|^2 \leqslant (3/4)^{\frac{t-2}{2}} \cdot 3/2.$$

□

The notation $X = Y \pm Z$ means that $|X - Y| \leqslant Z$.

**Corollary 8.21.** *For $t \in \mathbb{N}$ and every pair of clouds $A, B \in \mathcal{B}$,*

$$\frac{1}{n} \sum_{a \in A, b \in B} \langle a, b\rangle^t = \frac{1}{n} \sum_{a \in A} \langle a, \pi_{B \leftarrow A}(a)\rangle^t \quad \pm 2 \cdot (3/4)^{t/2}.$$

**Remark 8.22.** The left-hand side in the corollary is the inner product of the vectors $1/\sqrt{R} \sum_{u \in A} u^{\otimes t}$ and $1/\sqrt{R} \sum_{v \in B} v^{\otimes t}$. If $t$ is even, then we can replace the right-hand side by $\rho(A,B)^t$. This fact that the functional $\rho(A,B)^t$ is closely approximated by inner products averaged-tensored vectors has implicitly been used in [KV05] and was explicitly noted in [AKK+08, Lemma 2.2].

## 8.5. Integrality Gap Instance for Unique Games

In this section, we will exhibit the construction strong SDP integrality gap for the $\mathsf{E2Lin}_q$ problem. Recall that the $\mathsf{E2Lin}_q$ problem is a special case of Unique Games. To this end, we follow the approach of Khot–Vishnoi [KV05] to construct the gap instance.

Khot et al. [KKMO07] show a UGC-based hardness result for the $\mathsf{E2Lin}_q$ problem. Specifically, they exhibit a reduction $\Phi_{\gamma,q}$ that maps a Unique Games instance $\Phi$ to an $\mathsf{E2Lin}_q$ instance $\Phi_{\gamma,q}(\Phi)$ such that the following holds: For every $\gamma > 0$ and all $q \geqslant q_0(\gamma)$,

**Completeness:** If $\Phi$ is $1-\eta$-satisfiable then $\Phi_{\gamma,q}(\Phi)$ is $1-\gamma-o_{\eta,\delta}(1)$ satisfiable.

**Soundness:** If $\Phi$ has no labeling satisfying more than $\delta$-fraction of the constraints, then no assignment satisfies more than $q^{-\eta/2}+o_{\eta,\delta}(1)$-fraction of equations in $\Phi_{\gamma,q}(\Phi)$.

Here the notation $o_{\eta,\delta}(1)$ refers to any function that tends to 0 whenever $\eta$ and $\delta$ go to naught. The details of the $\Phi_{\gamma,q}$ reduction are included in Figure 8.1 for the sake of completeness.

The rest of the chapter is devoted to the proof of the following theorem.

**Theorem 8.23.** *Let $\Phi$ be a weak $(1-\eta,\delta)$-gap instance of* Unique Games. *Then, for every $q$ of order unity, there exists an SDP solution for the $\mathsf{E2Lin}_q$ instance $\Phi_{\gamma,q}(\Phi)$ such that*

- *the SDP solution is feasible for $\mathsf{LH}_r$ with $r = 2^{\Omega(1/\eta^{1/4})}$,*
- *the SDP solution is feasible for $\mathsf{SA}_r$ with $r = \Omega(\eta^{1/4})$,*
- *the SDP solution has value $1-\gamma-o_{\eta,\delta}(1)$ for $\Phi_{\gamma,q}(\Phi)$.*

*In particular, the $\mathsf{E2Lin}_q$ instance $\Phi_{\gamma,q}(\Phi)$ is a $(1-\gamma-o_{\eta,\delta}(1), q^{-\eta/2}+o_{\eta,\delta}(1)))$ integrality gap instance for the relaxation $\mathsf{LH}_r$ for $r = 2^{\Omega(1/\eta^{1/4})}$. Further, $\Phi_{\gamma,q}(\Phi)$ is a $(1-\gamma-o_{\eta,\delta}(1), q^{-\eta/2}+o_{\eta,\delta}(1))$ integrality gap instance for the relaxation $\mathsf{SA}_r$ for $r = \Omega(1/\eta^{1/4})$.*

## 8.6. Integral Vectors

In this section, we will develop tools to create and manipulate vectors all of whose coordinates are "integral".

**$\mathsf{E2Lin}_q$ Hardness Reduction** [KKMO07]

**Input** A Unique Games instance $\Phi$ with vertex set $V$, edge set $E \subseteq V \times V$ (we assume the graph $(V, E)$ to be regular), and permutations $\{\pi_e : [R] \to [R]\}_{e \in E}$.

**Output** An $\mathsf{E2Lin}_q$ instance $\Phi_{\gamma,q}(\Phi)$ with vertex set $\mathcal{V} = V \times \mathbb{F}_q^R$. Let $\{\mathcal{F}_v : \mathbb{F}_q^R \to \mathbb{F}_q\}_{v \in V}$ denote an $\mathbb{F}_q$-assignment to $\mathcal{V}$. The constraints of $\Phi_{\gamma,q}(\Phi)$ are given by the tests performed by the following probabilistic verifier:

- Pick a random vertex $v \in V$. Choose two random neighbours $w, w' \in N(v) \subseteq V$. Let $\pi, \pi'$ denote the permutations on the edges $(w, v)$ and $(w', v)$.
- Sample $x \in \mathbb{F}_q^R$ uniformly at random. Generate $y \in \mathbb{F}_q^R$ as follows:

$$y_i = \begin{cases} x_i \text{ with probability } 1 - \gamma \\ \text{uniform random element from } \mathbb{F}_q \text{ with probability } \gamma \end{cases}$$

- Generate a uniform random element $c \in \mathbb{F}_q$.
- Test if $\mathcal{F}_w(y \circ \pi + c \cdot \mathbb{1}) = \mathcal{F}_{w'}(x \circ \pi') + c$. (Here, $x \circ \pi$ denotes the vector $(x_{\pi(i)})_{i \in [R]}$.)

Figure 8.1.: Reduction from Unique Games to $\mathsf{E2Lin}_q$

**$\{\pm 1\}$-integral vectors.** We begin by defining our notion of a $\{\pm 1\}$-integral vector.

**Definition 8.24.** Let $\mathcal{R} = (\Omega, \mu)$ be a probability space. A function $u\colon \mathcal{R} \to \{\pm 1\}$ is called an $\{\pm 1\}$-*integral vector*. In other words, $u$ is a $\{\pm 1\}$-valued random variable defined on the probability space $\mathcal{R}$. We define an inner product of functions $u, v\colon \mathcal{R} \to \{\pm 1\}$ by

$$\langle u, v\rangle = \mathop{\mathbb{E}}_{r \sim \mathcal{R}} u(r)v(r).$$

In our construction, we often start with $\{\pm 1\}$-integral vectors given by the hypercube $\{\pm 1\}^R$. In the terminology of $\{\pm 1\}$-integral vectors, we can think of the hypercube $\{\pm 1\}^R$ as the set of $\{\pm 1\}$-integral vectors where $\mathcal{R}$ is the uniform distribution over $\{1, \ldots, R\}$.

The following lemma shows how the Goemans–Williamson [GW95] rounding scheme can be thought of as a procedure to "round" arbitrary real vectors to $\{\pm 1\}$-integral vectors.

**Observation 8.25.** *Given a family of unit vectors $\{v_1, \ldots, v_R\} \in \mathbb{R}^d$, define the set of $\{\pm 1\}$-valued functions $v_1^*, \ldots, v_R^* : \mathcal{R} \to \{\pm 1\}$ with $\mathcal{R} = \mathcal{G}^d$ - the Gaussian space of appropriate dimension as follows:*

$$v_i^*(g) = \mathsf{sign}(\langle v_i, g\rangle)$$

*for $g \in \mathcal{G}^d$. The $\{\pm 1\}$-valued functions $\{v_i^*\}$ satisfy $\langle v_1^*, v_2^*\rangle = 2\arccos(\langle v_1, v_2\rangle)/\pi$. Specifically, this operation obeys the following properties:*

$$\langle u, v\rangle = 0 \Longleftrightarrow \langle u^*, v^*\rangle = 0 \qquad \langle u, v\rangle = 1 - \varepsilon \Longrightarrow \langle u^*, v^*\rangle \geqslant 1 - O(\sqrt{\varepsilon})$$

The tensor product operation on $\{\pm 1\}$-integral vectors, yields a $\{\pm 1\}$-integral vector.

**Definition 8.26.** Given two $\{\pm 1\}$-valued functions $u : \mathcal{R}_1 \to \{\pm 1\}$ and $v : \mathcal{R}_2 \to \{\pm 1\}$, the tensor product $u \otimes v : \mathcal{R}_1 \times \mathcal{R}_2 \to \{\pm 1\}$ is defined as $u \otimes v(r_1, r_2) = u(r_1)v(r_2)$.

**Observation 8.27.** *For $u, u'\colon \mathcal{R}_1 \to \{\pm 1\}$ and $v, v'\colon \mathcal{R}_2 \to \{\pm 1\}$, we have*

$$\begin{aligned}\langle u \otimes v, u' \otimes v'\rangle &= \mathop{\mathbb{E}}_{r_1, r_2}[u \otimes v(r_1, r_2) u' \otimes v'(r_1, r_2)] \\ &= \mathop{\mathbb{E}}_{r_1}[u(r_1)u'(r_1)] \mathop{\mathbb{E}}_{r_2}[v(r_2)v'(r_2)] = \langle u, u'\rangle\langle v, v'\rangle\end{aligned}$$

**$\mathbb{F}_q$-integral vectors.** Let $q$ be a prime. Now, we will define $\mathbb{F}_q$-integral vectors and their tensor products.

**Definition 8.28.** A $\mathbb{F}_q$-integral vector $v\colon \mathcal{R} \to \mathbb{F}_q$ is a function from a measure space $\mathcal{R}$ to $\mathbb{F}_q$. For a $\mathbb{F}_q$-integral vector $v\colon \mathcal{R} \to \mathbb{F}_q$, its symmetrization $\tilde{v}\colon \mathcal{R} \times \mathbb{F}_q^* \to \mathbb{F}_q$ is defined by $\tilde{v}(r,t) = t \cdot v(r)$.

Given a map $f\colon \mathbb{F}_q \to \mathbb{C}^d$, we denote by $f(v) := f \circ v$ the composition of functions $f$ and $v$. Here are few examples of functions that will be relevant to us:

1. The function $\chi\colon \mathbb{F}_q \to \mathbb{C}^{q-1}$ given by

$$\chi(i) \stackrel{\text{def}}{=} \tfrac{1}{\sqrt{q-1}}(\omega^{1\cdot i}, \ldots, \omega^{j \cdot i}, \ldots, \omega^{(q-1)i}),$$

   where $\omega$ is a primitive $q^{\text{th}}$ root of unity. The vector $\chi(i) \in \mathbb{C}^{q-1}$ is the restriction of the $i^{\text{th}}$ character function of the group $\mathbb{Z}_q$ to the set $\mathbb{F}_q^*$. It is easy to see that

$$\langle \chi(a), \chi(b) \rangle = \mathop{\mathbb{E}}_{t \in \mathbb{F}_q^*}\left[\omega^{ta} \cdot \omega^{-tb}\right] = \begin{cases} 1 & \text{if } a = b, \\ -\frac{1}{q-1} & \text{if } a \neq b. \end{cases}$$

2. Let $\psi_0, \psi_1, \ldots, \psi_{q-1}$ denote the corners of the $q$-ary simplex in $\mathbb{R}^{q-1}$, translated so that the origin is its geometric center. Define the function $\psi\colon \mathbb{F}_q \to \mathbb{R}^{q-1}$ as $\psi(i) := \psi_i$. Again, the vectors satisfy

$$\langle \psi(a), \psi(b) \rangle = \begin{cases} 1 & \text{if } a = b, \\ -\frac{1}{q-1} & \text{if } a \neq b. \end{cases}$$

**Remark 8.29.** A $\mathbb{F}_q$-integral vector $v \in \mathbb{F}_q^N$ can be thought of as a $\mathbb{F}_q$-valued function over the measure space $([N], \mu)$ where $\mu$ is the uniform distribution over $[N]$.

**Remark 8.30.** The following notions are equivalent: Collection of $\mathbb{F}_q$-valued functions on some measure space $\mathcal{R} \iff$ Collection of jointly-distributed, $\mathbb{F}_q$-valued random variables $\iff$ Distribution over $\mathbb{F}_q$-assignments.

For the case of $\mathbb{F}_q$-integral vector, the tensor product operation is to be defined carefully, in order to mimic the properties of the traditional tensor product. We will use the following definition for the tensor operation $\otimes_q$.

**Definition 8.31.** Given two $\mathbb{F}_q$-valued functions $u\colon \mathcal{R} \to \mathbb{F}_q$ and $u'\colon \mathcal{R}' \to \mathbb{F}_q$, define the symmetrized tensor product $u \otimes_q u'\colon (\mathcal{R} \times \mathbb{F}_q^*) \times (\mathcal{R}' \times \mathbb{F}_q^*) \to \mathbb{F}_q$ as

$$(u \otimes_q u')(r,t,r',t') \stackrel{\text{def}}{=} t \cdot u(r) + t' \cdot u'(r').$$

**Lemma 8.32.** *For any $\mathbb{F}_q$-valued functions $u,v\colon \mathcal{R} \to \mathbb{F}_q$ and $u',v'\colon \mathcal{R}' \to \mathbb{F}_q$,*

$$\langle \psi(u \otimes_q u'), \psi(v \otimes_q v')\rangle = \langle \psi(u), \psi(v)\rangle \langle \psi(u'), \psi(v')\rangle.$$

*Proof.*

$$\begin{aligned}
&\langle \psi(u \otimes_q u'), \psi(v' \otimes_q v')\rangle \\
&= \langle \chi(u \otimes_q u'), \chi(v' \otimes_q v')\rangle \qquad (\text{using } \langle \psi_a, \psi_b\rangle = \langle \chi(a), \chi(b)\rangle) \\
&= \mathop{\mathbb{E}}_{(r,t)} \mathop{\mathbb{E}}_{(r',t')} \mathop{\mathbb{E}}_{\ell \in \mathbb{F}_q^*} \omega^{\ell t u(r) + \ell t' u'(r')} \cdot \omega^{-\ell t v(r) - \ell t' v'(r')} \qquad (\text{by definitions of } \otimes_q \text{ and } \chi) \\
&= \mathop{\mathbb{E}}_{\ell \in \mathbb{F}_q^*} \left( \mathop{\mathbb{E}}_{(r,t)} \omega^{\ell t u(r) - \ell t v(r)} \right) \cdot \left( \mathop{\mathbb{E}}_{(r',t')} \omega^{\ell t' u'(r') - \ell t' v'(r')} \right) \\
&= \mathop{\mathbb{E}}_{\ell \in \mathbb{F}_q^*} \left( \mathop{\mathbb{E}}_{r} \langle \chi(\ell u(r)), \chi(\ell v(r))\rangle \right) \cdot \left( \mathop{\mathbb{E}}_{r'} \langle \chi(\ell u'(r')), \chi(\ell v'(r'))\rangle \right) \\
&= \mathop{\mathbb{E}}_{\ell \in \mathbb{F}_q^*} \langle \chi(\ell u), \chi(\ell v)\rangle \langle \chi(\ell u'), \chi(\ell v')\rangle \\
&= \langle \chi(u), \chi(v)\rangle \langle \chi(u'), \chi(v')\rangle \qquad (\text{using } \langle \chi(\ell a), \chi(\ell b)\rangle = \langle \chi(a), \chi(b)\rangle \text{ for } \ell \in \mathbb{F}_q^*) \\
&= \langle \psi(u), \psi(v)\rangle \langle \psi(u'), \psi(v')\rangle \qquad (\text{using } \langle \psi_a, \psi_b\rangle = \langle \chi(a), \chi(b)\rangle)
\end{aligned}$$

□

**Remark 8.33.** Unlike the ordinary tensor operation, the $q$-ary tensor operation we defined is not associative. Formally, we define the tensoring operation to be right-associative

$$u_1 \otimes_q u_2 \otimes_q \ldots \otimes_q u_{k-1} \otimes_q u_k \stackrel{\text{def}}{=} u_1 \otimes_q \Big( u_2 \otimes_q \big( \ldots (u_{k-1} \otimes_q u_k) \cdots \big) \Big).$$

The lack of associativity will never be an issue in our constructions.

We need the following simple technical observation in one of our proofs.

**Observation 8.34.** *Let $u,v\colon \mathcal{R} \to \mathbb{F}_q$ be two "symmetric" $\mathbb{F}_q$-integral vectors. that is, $\mathbb{P}_r\{u(r) - v(r) = a\} = \mathbb{P}_r\{u(r) - v(r) = b\}$ for all $a,b \in \mathbb{F}_q^*$. Then, for all $a,b \in \mathbb{F}_q$, we have $\mathbb{E}_r \langle \psi(a + u(r)), \psi(b + v(r))\rangle = \langle a \otimes u, b \otimes v\rangle$.*

We wish to point out that in our applications, the vectors $u$ and $v$ will be tensor powers. In this case, the symmetry condition is always satisfied.

*Proof.* Using the symmetry assumption, we see that

$$\begin{aligned}\mathop{\mathbb{P}}_{r\sim\mathcal{R},t,t'\in\mathbb{F}_q^*}&\Big\{ta+t'u(r)=tb+t'v(r)\Big\}\\ &=\mathop{\mathbb{P}}_{r\sim\mathcal{R},t\in\mathbb{F}_q^*}\Big\{a-b=t\cdot\Big(v(r)-u(r)\Big)\Big\}\\ &\qquad=\mathop{\mathbb{P}}_{r\sim\mathcal{R}}\Big\{a-b=v(r)-u(r)\Big\}\end{aligned} \tag{8.6}$$

If we let $\rho$ denote this probability, then we have $\langle a\otimes u, b\otimes v\rangle=\rho-(1-\rho)/(q-1)$ (using the left-hand side of Eq. (8.6) as well as $\mathbb{E}_r\langle\psi(a+u(r)),\psi(b+v(r))\rangle=\rho-(1-\rho)/(q-1)$ (using the right-hand side of Eq. (8.6)). ☐

The following procedure yields a way to generate $\mathbb{F}_q$-integral vectors from arbitrary vectors. The transformation is inspired by the rounding scheme for Unique Games in Charikar et al. [CMM06a].

**Observation 8.35.** *Define the function $\zeta\colon\mathcal{G}^q\to\mathbb{F}_q$ on the Gaussian domain as follows:*

$$\zeta(x_1,\ldots,x_q)=\operatorname{argmax}_{i\in[q]}x_i \tag{8.7}$$

*Given a family of unit vectors $\{v_1,\ldots,v_R\}\in\mathbb{R}^d$, define the set of $\mathbb{F}_q$-valued functions $v_1^*,\ldots,v_R^*\colon\mathcal{R}\to\mathbb{F}_q$ with $\mathcal{R}=(\mathcal{G}^d)^q$ —the Gaussian space of appropriate dimension— as follows:*

$$v_i^*(g_1,\ldots,g_q)=\zeta(\langle v_i,g_1\rangle,\ldots,\langle v_i,g_q\rangle)$$

*for $g_1,\ldots,g_q\in(\mathcal{G}^d)^q$. The $\mathbb{F}_q$-valued functions $\{v_i^*\}$ satisfy,*

1. *$\langle u,v\rangle=0\Longrightarrow\langle\psi(u^*),\psi(v^*)\rangle=0$,*
2. *$\langle u,v\rangle=1-\varepsilon\Longrightarrow\langle\psi(u^*),\psi(v^*)\rangle=1-f(\varepsilon,q)=1-O(\sqrt{\varepsilon\log q})$.*

*Proof.* To see (1), observe that if $\langle u,v\rangle=0$, then the sets of random variables $\{\langle u,g_1\rangle,\ldots,\langle u,g_q\rangle\}$ and $\{\langle v,g_1\rangle,\ldots,\langle v,g_q\rangle\}$ are completely independent of each other. Therefore,

$$\langle\psi(u^*),\psi(v^*)\rangle=\mathop{\mathbb{E}}_{r\in\mathcal{G}^{dq}}\Big[\psi(u^*(r))\Big]\cdot\mathop{\mathbb{E}}_{r\in\mathcal{G}^{dq}}\Big[\psi(u^*(r))\Big]=0\,.$$

Assertion 2 follows from Lemma C.8 in [CMM06a]. ☐

## 8.7. Local Distributions for Unique Games

In this section, we will construct local distribution over labelings to a Unique Games instance.

The following facts are direct consequences of the (symmetrized) $\ell_2^2$-triangle inequality.

**Fact 8.36.** *Let $a, b, c \in \bigcup \mathcal{B}$ with $|\langle a, b\rangle| = 1 - \eta_{ab}$ and $|\langle b, c\rangle| = 1 - \eta_{bc}$. Then, $|\langle a, c\rangle| \geqslant 1 - \eta_{ab} - \eta_{bc}$.*

**Fact 8.37.** *Let $A, B, C \in \mathcal{B}$ with $\rho(A, B) = 1 - \eta_{AB}$ and $\rho(B, C) = 1 - \eta_{BC}$. Then, $\rho(A, C) \geqslant 1 - \eta_{AB} - \eta_{BC}$.*

The construction in the proof of the next lemma is closely related to propagation-style UG algorithms [Tre05, AKK+08].

**Definition 8.38.** A set $\mathcal{S} \subseteq \mathcal{B}$ is *consistent* if

$$\forall A, B \in \mathcal{S}. \qquad \rho(A, B) \geqslant 1 - 1/16 .$$

**Lemma 8.39.** *If $\mathcal{S} \subseteq \mathcal{B}$ is consistent, there exists bijections $\{\pi_A \colon [R] \to A\}_{A \in \mathcal{S}}$ such that*

$$\forall A, B \in \mathcal{S}. \qquad \pi_B = \pi_{B \leftarrow A} \circ \pi_A .$$

*Proof.* We can construct the bijections in a greedy fashion: Start with an arbitrary cloud $C \in \mathcal{S}$ and choose an arbitrary bijection $\pi_C \colon [R] \to C$. For all other clouds $B \in \mathcal{S}$, choose $\pi_B := \pi_{B \leftarrow C} \circ \pi_C$.

Let $A, B$ be two arbitrary clouds in $\mathcal{S}$. Let $\sigma_{A \leftarrow B} := \pi_A \circ \pi_B^{-1}$. To prove the lemma, we have to verify that $\sigma_{A \leftarrow B} = \pi_{A \leftarrow B}$. By construction, $\sigma_{A \leftarrow B} = \pi_{A \leftarrow C} \circ \pi_{C \leftarrow B}$. Let $\eta = 1/16$. Since $\rho(A, C) \geqslant 1 - \eta$ and $\rho(B, C) \geqslant 1 - \eta$, we have $|\langle b, \sigma_{A \leftarrow B}(b)\rangle| \geqslant 1 - 2\eta$ for all $b \in B$ (using Fact 8.36). Since $(1 - 2\eta)^2 > 1 - 4\eta = 3/4$, Fact 8.18 (uniqueness of bijection) implies that $\sigma_{A \leftarrow B} = \pi_{A \leftarrow B}$. □

Hence, for a consistent set of clouds $\mathcal{S}$, the distribution over local Unique Games labelings $\mu_{\mathcal{S}}$ can be defined easily as follows:

> Sample $\ell \in [R]$ uniformly at random, and for every cloud $A \in \mathcal{S}$, assign $\pi_A(\ell)$ as label.

To construct a local distribution for a set $\mathcal{S}$ which is not consistent, we partition the set $\mathcal{S}$ into consistent clusters. To this end, we make the following definition:

**Definition 8.40.** A set $\mathcal{S} \subseteq \mathcal{B}$ is *consistent* with respect to a partition $P$ of $\mathcal{B}$ (denoted $\mathrm{Cons}(\mathcal{S}, P)$) if

$$\forall \mathcal{C} \in P. \quad \forall A, B \in \mathcal{C} \cap \mathcal{S}. \qquad \rho(A, B) \geqslant 1 - 1/16 .$$

We use $\mathrm{Incons}(\mathcal{S}, P)$ to denote the event that $\mathcal{S}$ is not consistent with $P$. The following is a corollary of Lemma 8.39.

**Corollary 8.41.** *Let $P$ be a partition of $\mathcal{B}$ and let $\mathcal{S} \subseteq \mathcal{B}$. If $\mathrm{Cons}(\mathcal{S}, P)$, then there exists bijections $\{\pi_A \colon [R] \to A \mid A \in \mathcal{S}\}$ such that*

$$\forall \mathcal{C} \in P. \quad \forall A, B \in \mathcal{C} \cap \mathcal{S}. \qquad \pi_B = \pi_{B \leftarrow A} \circ \pi_A .$$

The following lemma relies on the fact that the correlations $\rho(A, B)$ behave up to a small errors like inner products of real vectors. In other words, there is a geometric representation of the correlations $\rho(A, B)$ that can be used for the decomposition. This insight has also been used in UG algorithms[AKK+08].

**Lemma 8.42.** *For every $t \in \mathbb{N}$, there exists a distribution over partitions $P$ of $\mathcal{B}$ such that*

– *if $\rho(A, B) \geqslant 1 - \varepsilon$, then*

$$\mathbb{P}\{P(A) = P(B)\} \geqslant 1 - O(t\sqrt{\varepsilon}) .$$

– *if $\rho(A, B) \leqslant 1 - 1/16$, then*

$$\mathbb{P}\{P(A) = P(B)\} \leqslant (3/4)^t .$$

*Proof.* Let $s \in \mathbb{N}$ be even and large enough (we will determine the value of $s$ later). For every set $B \in \mathcal{B}$, define a vector $\boldsymbol{v}_B \in \mathbb{R}^D$ with $D := d^s$ as

$$\boldsymbol{v}_B := \tfrac{1}{\sqrt{R}} \sum_{v \in B} v^{\otimes s} .$$

We consider the following distribution over partitions $P$ of $\mathcal{B}$: Choose $t$ random hyperplanes $H_1, \ldots, H_t$ through the origin in $\mathbb{R}^D$. Consider the partition of $\mathbb{R}^D$ formed by these hyperplanes. Output the induced partition $P$ of $\mathcal{B}$ (two sets $A, B \in \mathcal{B}$ are in the same cluster of $P$ if and only if $\boldsymbol{v}_A$ and $\boldsymbol{v}_B$ are not separated by any of the hyperplanes $H_1, \ldots, H_t$).

Since $s$ is even, Corollary 8.21 shows that for any two sets $A, B \in \mathcal{B}$,

$$\langle \boldsymbol{v}_A, \boldsymbol{v}_B \rangle = \rho(A,B)^s \pm 2 \cdot (3/4)^{-s/2}.$$

Furthermore, if $\rho(A,B) = 1-\varepsilon$, then

$$\langle \boldsymbol{v}_A, \boldsymbol{v}_B \rangle \geqslant (1-\varepsilon)^s \geqslant 1 - s\varepsilon.$$

Let $\eta = 1/16$. We choose $s$ minimally such that $(1-\eta)^s + 2 \cdot (3/4)^{-s/2} \leqslant 1/\sqrt{2}$. (So $s$ is an absolute constant.) Then for any two sets $A, B \in \mathcal{B}$ with $\rho(A,B) \leqslant 1-\eta$, their vectors have inner product $\langle \boldsymbol{v}_A, \boldsymbol{v}_B \rangle \leqslant 1/\sqrt{2}$. Thus, a random hyperplane through the origin separates $\boldsymbol{v}_A$ and $\boldsymbol{v}_B$ with probability at least $1/4$. Therefore,

$$\mathbb{P}\{P(A) = P(B)\} \leqslant (3/4)^t.$$

On the other hand, if $\rho(A,B) = 1-\varepsilon$, then the vectors of $A$ and $B$ have inner product $\langle \boldsymbol{v}_A, \boldsymbol{v}_B \rangle \geqslant 1 - s\varepsilon$. Thus, a random hyperplane through the origins separates the vectors with probability at most $O(\sqrt{\varepsilon})$. Hence,

$$\mathbb{P}\{P(A) = P(B)\} \geqslant \left(1 - O(\sqrt{\varepsilon})\right)^t \geqslant 1 - O(t\sqrt{\varepsilon}).$$

□

**Remark 8.43.** Using a more sophisticated construction, we can improve the bound $1 - O(t\sqrt{\varepsilon})$ to $1 - O(\sqrt{t\varepsilon})$.

The previous lemma together with a simple union bound imply the next corollary.

**Corollary 8.44.** *The distribution over partitions from Lemma 8.42 satisfies the following property: For every set $\mathcal{S} \subseteq \mathcal{B}$,*

$$\mathbb{P}\left\{\mathrm{Incons}(\mathcal{S}, P)\right\} \leqslant |S|^2 \cdot (3/4)^t$$

**Remark 8.45.** Using a slightly more refined argument (triangle inequality), we could improve the bound $r^2 \cdot (3/4)^t$ to $r \cdot (3/4)^t$.

## 8.8. Construction of SDP Solutions for E2LIN(q)

In this section, we construct SDP vectors and local distributions for $\mathcal{B} \times \mathbb{F}_q^R$ that form the variables in the $\Phi_{\gamma,q}(\Phi)$ instance described in Section 8.5. The set $\mathcal{B} \times \mathbb{F}_q^R$ correspond to the set of vertices in the instance obtained by applying a $q$-ary long code based reduction on the Unique Games instance encoded by $\mathcal{B}$. For a vertex $(B, x) \in \mathcal{B} \times \mathbb{F}_q^R$, we index the coordinates of $x$ by the elements of $B$. Specifically, we have $x = (x_b)_{b \in B} \in \mathbb{F}_q^B$.

**Geometric Partitioning.** Apply Lemma 8.42 to the collection of sets of vectors $\mathcal{B}$. We obtain a distribution $\mathcal{P}$ over partitions $P$ of $\mathcal{B}$ into $T$ disjoint subsets $\{P_\alpha\}_{\alpha=1}^{T}$. For a subset $\mathcal{S} \subset \mathcal{B}$, let $\mathcal{S} = \{\mathcal{S}_\alpha\}_{\alpha=1}^{T}$ denote the partition induced on the set $\mathcal{S}$, that is, $\mathcal{S}_\alpha := P_\alpha \cap \mathcal{S}$. For a family $B \in \mathcal{B}$, let $\alpha_B$ denote the index of the set $P_{\alpha_B}$ in the partition $P$ that contains $B$.

### 8.8.1. Vector Solution

For a vertex $(B, x) \in \mathcal{B} \times \mathbb{F}_q^R$, the corresponding SDP vectors are given by functions $\boldsymbol{V}_j^{B,x} \colon \mathcal{P} \times [T] \times \mathcal{R} \to \mathbb{R}^q$ defined as follows:

$$\boldsymbol{W}_j^{B,x}(r) = \tfrac{1}{\sqrt{R}} \sum_{b \in B} \psi\Big(x_b - j + b^{\otimes t}(r)\Big) \tag{8.8}$$

$$\boldsymbol{U}_j^{B,x}(P, \alpha, r) = P_\alpha(B) \cdot \boldsymbol{W}_j^{B,x}(r) \tag{8.9}$$

$$\boldsymbol{V}_j^{B,x} = \tfrac{1}{q} \boldsymbol{V}_0 + \tfrac{\sqrt{q-1}}{q} \boldsymbol{U}_j^{B,x} \tag{8.10}$$

Here $\mathcal{R}$ is the measure space over which the tensored vectors $b^{\otimes t}$ are defined. The notation $P_\alpha(B)$ denotes the 0/1-indicator for the event $B \in P_\alpha$. Further, $\boldsymbol{V}_0$ is a unit vector orthogonal to all the vectors $\boldsymbol{U}_j^{B,x}$.

Let us evaluate the inner product between two vectors $\boldsymbol{V}_i^{A,x}$ and $\boldsymbol{V}_j^{B,y}$, (in this way, we also clarify the intended measure on the coordinate set)

$$\begin{aligned}
\langle \boldsymbol{V}_i^{A,x}, \boldsymbol{V}_j^{B,y} \rangle &= \tfrac{1}{q^2} + \tfrac{q-1}{q^2} \langle \boldsymbol{U}_i^{A,x}, \boldsymbol{U}_j^{B,y} \rangle \\
&= \tfrac{1}{q^2} + \tfrac{q-1}{q^2} \mathop{\mathbb{E}}_{P \sim \mathcal{P}} \textstyle\sum_{\alpha=1}^{T} P_\alpha(A) P_\alpha(B) \langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y} \rangle \\
&= \tfrac{1}{q^2} + \tfrac{q-1}{q^2} \mathop{\mathbb{P}}_{P \sim \mathcal{P}} \{P(A) = P(B)\} \langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y} \rangle
\end{aligned} \tag{8.11}$$

Let us also compute the inner product of $\boldsymbol{W}_i^{A,x}$ and $\boldsymbol{W}_j^{B,y}$. Recall the notation $\langle u, v \rangle_\psi := \langle \psi(u), \psi(v) \rangle$.

$$\begin{aligned}
\langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y} \rangle &= \tfrac{1}{n} \sum_{a \in A, b \in B} \mathop{\mathbb{E}}_{r \sim \mathcal{R}} \langle x_a - i + a^{\otimes t}(r), y_b - j + b^{\otimes t}(r) \rangle_\psi \\
&= \tfrac{1}{n} \sum_{a \in A, b \in B} \langle (x_a - i) \otimes a^{\otimes t}, (y_b - j) \otimes b^{\otimes t} \rangle_\psi \quad \text{(by Observation 8.34)} \\
&= \tfrac{1}{n} \sum_{a \in A, b \in B} \langle \psi(x_a - i), \psi(y_b - j) \rangle \langle a, b \rangle_\psi^t \quad \text{(by Lemma 8.32)}
\end{aligned} \tag{8.12}$$

## 8.8.2. Local Distributions

Fix a subset $\mathcal{S} \subset \mathcal{B}$ of size at most $r$. In this section, we will construct a local distribution over $\mathbb{F}_q$-assignments for the vertex set $S = \mathcal{S} \times \mathbb{F}_q^R$ (see Figure 8.2). Clearly, the same construction also yields a distribution for a general set of vertices $S' \subset \mathcal{B} \times \mathbb{F}_q^R$ of size at most $r$.

**Remark 8.46.** In the construction in Figure 8.2, the steps 6–7 are not strictly necessary, but they simplify some of the following calculations. Specifically, we could use the $\mathbb{F}_q$-assignment $\{F^{B,x}\}_{(B,x)\in S}$ to define the local distribution for the vertex set $S$. The resulting collection of local distributions could be extended to an approximately feasible SDP solution (albeit using a slightly different vector solution).

We need the following two simple observations.

**Observation 8.47.** *For all $a, b \in \mathbb{F}_q$, we have*

$$\mathop{\mathbb{P}}_{\kappa\in\mathbb{F}_q}[a+\kappa = i \wedge b+\kappa = j] = \tfrac{1}{q^2} + \tfrac{q-1}{q^2}\langle \psi(a-i), \psi(b-j)\rangle .$$

*Proof.* If $a-i = b-j$ then both LHS and RHS are equal to $1/q$, otherwise both are equal to 0. ☐

**Observation 8.48.** *Fix $a, b \in \mathbb{F}_q$, over a random choice of $h_1, h_2 \in \mathbb{F}_q$,*

$$\mathop{\mathbb{E}}_{h_1,h_2\in\mathbb{F}_q}[\langle \psi(a+h_1), \psi(b+h_2)\rangle] = 0 .$$

*Proof.* Follows easily from the fact that $\langle \psi(i), \psi(j)\rangle = 1$ if $i = j$ and $-1/q-1$ otherwise. ☐

The next lemma shows that the second-order correlations of the distribution $\mu_S$ approximately match the inner products of the vector solution $\{\boldsymbol{V}_i^{A,x}\}$.

**Lemma 8.49.** *For any two vertices $(A,x), (B,y) \in S$,*

$$\mathop{\mathbb{P}}_{Z\sim\mu_S}\left[Z^{A,x} = i \wedge Z^{B,y} = j\right] = \langle \boldsymbol{V}_i^{A,x}, \boldsymbol{V}_j^{B,y}\rangle \quad \pm 10|\mathcal{S}|^2 (3/4)^{t/2} .$$

*Proof.* Firstly, since $\mathbb{P}[\mathrm{Cons}(\mathcal{S}, P)] \geqslant 1 - |S|^2(3/4)^t$ (by Corollary 8.44),

$$\mathop{\mathbb{P}}_{\mu_S}\left[Z^{A,x} = i \wedge Z^{B,y} = j\right] = \mathop{\mathbb{P}}_{\mu_S}\left[Z^{A,x} = i \wedge Z^{B,y} = j \;\middle|\; \mathrm{Cons}(\mathcal{S}, P)\right] \quad \pm |S|^2(3/4)^t . \tag{8.13}$$

For $S = \mathcal{S} \times \mathbb{F}_q^R$, the local distribution $\mu_S$ over assignments $\mathbb{F}_q^S$ is defined by the following sampling procedure:

**Partitioning:**

1. Sample a partition $P = \{P_\alpha\}_{\alpha=1}^T$ of $\mathcal{B}$ from the distribution $\mathcal{P}$ obtained by Lemma 8.42. Let $\alpha_A, \alpha_B$ denote the indices of sets in the partition $P$ that contain $A, B \in \mathcal{S}$ respectively.

2. If Incons$(\mathcal{S}, P)$ then output a uniform random $\mathbb{F}_q$-assignment to $S = \mathcal{S} \times \mathbb{F}_q^R$. Specifically, set
$$Z^{(B,x)} = \text{ uniform random element from } \mathbb{F}_q \qquad \forall B \in \mathcal{S}, x \in \mathbb{F}_q^R.$$

**Choosing Consistent Representatives:**

4. If Cons$(\mathcal{S}, P)$ then by Corollary 8.41, for every part $\mathcal{S}_\alpha = P_\alpha \cap \mathcal{S}$, there exists bijections $\Pi_{\mathcal{S}_\alpha} = \{\pi_B : [R] \to B \mid B \in \mathcal{S}_\alpha\}$ such that for every $A, B \in \mathcal{S}_\alpha$,
$$\pi_A = \pi_{A \leftarrow B} \circ \pi_B.$$

5. Sample $L = \{\ell_\alpha\}_{\alpha=1}^T$ by choosing each $\ell_\alpha$ uniformly at random from $[R]$. For every cloud $B \in \mathcal{S}$, define $\ell_B = \ell_{\alpha_B}$. The choice of $L$ determines a set of representatives for each $B \in \mathcal{S}$. Specifically, the representative of $B$ is fixed to be $\pi_B(\ell_B)$.

**Sampling Assignments:**

5. Sample $r \in \mathcal{R}$ from the corresponding probability measure and assign
$$F^{B,x}(P, L, r) = x_{\pi_B(\ell_B)} + \pi_B(\ell_B)^{\otimes t}(r).$$

6. Sample $H = \{h_\alpha\}_{\alpha=1}^T$ by choosing each $h_\alpha$ uniformly at random from $[q]$. For every cloud $B \in \mathcal{B}$, define $h_B = h_{\alpha_B}$.

7. Sample $\kappa$ uniformly at random from $[q]$.

8. For each $B \in \mathcal{S}_\alpha$ and $x \in \mathbb{F}_q^R$, set
$$Z^{B,x}(P, L, r, H, \kappa) = F^{B,x}(P, L, r) + h_B + \kappa.$$

9. Output the $\mathbb{F}_q$-assignment $\{Z^{B,x}\}_{(B,x) \in S}$.

Figure 8.2.: Local distribution over $\mathbb{F}_q$-assignments

Using Observation 8.47, and the definition of $Z^{A,x}$ and $Z^{B,y}$ we can write

$$\mathop{\mathbb{P}}_{\mu_S}\left[Z^{A,x} = i \wedge Z^{B,y} = j \,\middle|\, \mathrm{Cons}(\mathcal{S}, P)\right]$$
$$= \frac{1}{q^2} + \frac{q-1}{q^2} \mathop{\mathbb{E}}_{P,H,L,r}\left[\langle \psi(F^{A,x} + h_A - i), \psi(F^{B,y} + h_B - j)\rangle \,\middle|\, \mathrm{Cons}(\mathcal{S}, P)\right]. \quad (8.14)$$

If $A, B$ fall in the same set in the partition $P$ (that is $\alpha_A = \alpha_B$), then we have $h_A = h_B$. If $A, B$ fall in different sets (that is $\alpha_A \neq \alpha_B$), then $h_A, h_B$ are independent random variables uniformly distributed over $\mathbb{F}_q$. Using Observation 8.48, we can write

$$\mathop{\mathbb{E}}_{P,H,L,r}\left[\langle \psi(F^{A,x} + h_A - i), \psi(F^{B,y} + h_B - j)\rangle \,\middle|\, \mathrm{Cons}(\mathcal{S}, P)\right]$$
$$= \mathop{\mathbb{E}}_{P,L,r}\left[\mathbf{1}(\alpha_A = \alpha_B)\langle \psi(F^{A,x} - i), \psi(F^{B,y} - j)\rangle \,\middle|\, \mathrm{Cons}(\mathcal{S}, P)\right]. \quad (8.15)$$

Let $P$ be a partition such that $\mathrm{Cons}(\mathcal{S}, P)$ and $\alpha_A = \alpha_B = \alpha$. The bijections $\pi_A, \pi_B$ (see step 4 Figure 8.2) satisfy $\pi_A = \pi_{A\leftarrow B} \circ \pi_B$. Note that therefore $a = \pi_{A\leftarrow B}(b)$ whenever $a = \pi_A(\ell)$ and $b = \pi_B(\ell)$ for some $\ell \in [R]$. Hence,

$$\mathop{\mathbb{E}}_{L}\mathop{\mathbb{E}}_{r}\left[\langle \psi(F^{A,x}(P,L,r) - i), \psi(F^{B,y}(P,L,r) - j)\rangle\right]$$
$$= \mathop{\mathbb{E}}_{\ell_\alpha}\mathop{\mathbb{E}}_{r}\left[\langle \psi(x_{\pi_A(\ell_\alpha)} - i + \pi_A(\ell_\alpha)^{\otimes t}(r)), \psi(y_{\pi_B(\ell_\alpha)} - j + \pi_B(\ell_\alpha)^{\otimes t}(r))\rangle\right]$$
$$= \frac{1}{R} \sum_{\substack{a\in A, b\in B\\ a = \pi_{A\leftarrow B}(b)}} \mathop{\mathbb{E}}_{r}\langle \psi(x_a - i + a^{\otimes t}(r)), \psi(y_b - j + b^{\otimes t}(r))\rangle \qquad (\text{using } \pi_A = \pi_{A\leftarrow B} \circ \pi_B)$$
$$= \frac{1}{R} \sum_{\substack{a\in A, b\in B\\ a = \pi_{A\leftarrow B}(b)}} \langle \psi(x_a - i), \psi(y_b - j)\rangle \cdot \langle a, b\rangle_\psi^t \quad (\text{using Observation 8.34 and Lemma 8.32})$$
$$= \langle \boldsymbol{W}_i^{A,x}, \boldsymbol{W}_j^{B,y}\rangle \pm 2 \cdot (3/4)^{t/2} \qquad (\text{using Eq. (8.12) and Lemma 8.20}).$$

Combining the last equation with the previous equations (8.13)–(8.15), we can finish the proof

$$\begin{aligned}
&\mathop{\mathbb{P}}_{\mu_S}\left[Z^{A,x}=i\wedge Z^{B,y}=j\right]\\
&\quad=\tfrac{1}{q^2}+\tfrac{q-1}{q^2}\mathop{\mathbb{E}}_{P}\left[\mathbf{1}(\alpha_A=\alpha_B)\,\middle|\,\mathrm{Cons}(\mathcal{S},P)\right]\langle \boldsymbol{W}_i^{A,x},\boldsymbol{W}_j^{B,y}\rangle \quad \pm(|S|^2(3/4)^t+2\cdot(3/4)^{t/2})\\
&\quad=\tfrac{1}{q^2}+\tfrac{q-1}{q^2}\mathop{\mathbb{P}}_{P}\left[P(A)=P(B)\right]\cdot\langle \boldsymbol{W}_i^{A,x},\boldsymbol{W}_j^{B,y}\rangle \quad \pm 10|\mathcal{S}|^2(3/4)^{t/2}\\
&\qquad\qquad\qquad(\text{using } \mathbb{P}\{\mathrm{Cons}(\mathcal{S},P)\}\geqslant 1-|S|^2(3/4)^t \text{ and } |\langle \boldsymbol{W}_i^{A,x},\boldsymbol{W}_j^{B,y}\rangle|\leqslant 1)\\
&=\langle \boldsymbol{V}_i^{A,x},\boldsymbol{V}_j^{B,y}\rangle \quad \pm 10|\mathcal{S}|^2(3/4)^{t/2} \qquad (\text{using Eq. (8.11)}).
\end{aligned}$$

□

**Lemma 8.50.** *Let $\mathcal{S}'\subset\mathcal{S}$ be two subsets of $\mathcal{B}$ and let $S'=\mathcal{S}'\times\mathbb{F}_q^R$ and $S=\mathcal{S}\times\mathbb{F}_q^R$. Then,*

$$\|\mu_{S'}-\mathrm{margin}_{S'}S\|_1\leqslant 2|\mathcal{S}|^2(3/4)^t\,.$$

*Proof.* For a partition $P\in\mathcal{P}$, let $\mu_{S|P}$ denote the distribution $\mu_S$ conditioned on the choice of partition $P$. Firstly, we will show the following claim:

**Claim 8.51.** *If* $\mathrm{Cons}(\mathcal{S}',P)$ *and* $\mathrm{Cons}(\mathcal{S},P)$*, then* $\mu_{S'|P}=\mathrm{margin}_{S'}\mu_{S|P}$.

*Proof.* Let $\{\mathcal{S}_\alpha\}$ and $\{\mathcal{S}'_\alpha\}$ denote the partitions induced by $P$ on the sets $\mathcal{S}$ and $\mathcal{S}'$ respectively. Since $\mathcal{S}'\subseteq\mathcal{S}$, we have $\mathcal{S}'_\alpha\subseteq\mathcal{S}_\alpha$ for all $\alpha\in[T]$. By our assumption, each of the sets $\mathcal{S}'_\alpha$ are *consistent* in that $\rho(A,B)\geqslant 1-1/16$ for all $A,B\in\mathcal{S}'_\alpha$. Similarly, the sets $\mathcal{S}_\alpha$ are also *consistent*.

Let us consider the pair of sets $\mathcal{S}'_\alpha\subset\mathcal{S}_\alpha$ for some $\alpha\in[T]$. Intuitively, the vectors within these sets fall into $R$ distinct clusters. Thus the distribution over the choice of consistent representatives are the same in $\mu_{S'|P}$ and $\mathrm{margin}_{S'}\mu_{S|P}$. Formally, we have two sets of bijections $\Pi_{\mathcal{S}'_\alpha}=\{\pi'_A\mid A\in\mathcal{S}'_\alpha\}$ and $\Pi_{\mathcal{S}_\alpha}=\{\pi_A\mid A\in\mathcal{S}_\alpha\}$ satisfying the following property:

$$\pi_{A\to B}\circ\pi'_A(\ell)=\pi'_B(\ell)\qquad \pi_{A\to B}\circ\pi_A(\ell)=\pi_B(\ell)\qquad \forall A,B\in\mathcal{S}'_\alpha,\ell\in[R].$$

Fix a collection $A\in\mathcal{S}'_\alpha$. Let $\sim$ denote that two sets of random variables are identically distributed.

$$\begin{aligned}
\{\pi'_B(\ell_\alpha)\mid B\in\mathcal{S}'_\alpha\}&\sim\{\pi_{A\to B}\circ\pi'_A(\ell_\alpha)\mid B\in\mathcal{S}'_\alpha\}\\
&\sim\{\pi_{A\to B}(a)\mid B\in\mathcal{S}'_\alpha,\ a\text{ is uniformly random in }A\}\\
&\sim\{\pi_{A\to B}\circ\pi_A(\ell_\alpha)\mid B\in\mathcal{S}'_\alpha\}\sim\{\pi_B(\ell_\alpha)\mid B\in\mathcal{S}'_\alpha\}.
\end{aligned}$$

The variables $L = \{\ell_\alpha\}$ are independent of each other. Therefore,

$$\{\pi'_B(\ell_B) \mid B \in \mathcal{S}'\} \sim \{\pi_B(\ell_B) \mid B \in \mathcal{S}'\}.$$

Notice that the choice of $r \in \mathcal{R}$, $H$ and $\kappa$ are independent of the set $\mathcal{S}$. Hence, the final assignments $\{Z^{B,x} \mid B \in \mathcal{S}', x \in \mathbb{F}_q^R\}$ are identically distributed in both cases. □

Returning to the proof of Lemma 8.50, we can write

$$\begin{aligned}\|\mu_{S'} - \operatorname{margin}_{S'} \mu_S\|_1 &= \left\|\mathop{\mathbb{E}}_P \mu_{S'|P} - \mathop{\mathbb{E}}_P \operatorname{margin}_{S'} \mu_{S|P}\right\|_1 \\ &\leqslant \mathop{\mathbb{E}}_P\left[\|\mu_{S'|P} - \operatorname{margin}_{S'} \mu_{S|P}\|_1\right] \qquad \text{(using Jensen's inequality)} \\ &= \mathbb{P}[\operatorname{Incons}(\mathcal{S},P)] \cdot \mathop{\mathbb{E}}_P\left[\|\mu_{S'|P} - \operatorname{margin}_{S'} \mu_{S|P}\|_1 \;\middle|\; \operatorname{Incons}(\mathcal{S},P)\right].\end{aligned}$$

The first step uses that the operator $\mathsf{margin}_{S'}$ is linear. The final step in the above calculation makes use of Claim 8.51. The lemma follows by observing that $\mathbb{P}[\operatorname{Incons}(\mathcal{S},P)] \leqslant |\mathcal{S}|^2 (3/4)^t$ and $\|\mu_{S'|P} - \operatorname{margin}_{S'} \mu_{S|P}\|_1 \leqslant 2$. □

The next corollary follows from the previous lemma (Lemma 8.50) and the triangle inequality.

**Corollary 8.52.** *Let $\mathcal{S}, \mathcal{S}'$ be two subsets of $\mathcal{B}$ and let $S' = \mathcal{S}' \times \mathbb{F}_q^R$ and $S = \mathcal{S} \times \mathbb{F}_q^R$. Then,*

$$\|\operatorname{margin}_{S \cap S'} \mu_S - \operatorname{margin}_{S \cap S'} \mu_{S'}\|_1 \leqslant 4 \max\left(|\mathcal{S}|^2, |\mathcal{S}'|^2\right)(3/4)^t.$$

*Proof.* Suppose $\Phi$ is given by the vertex set $V$, the edge set $E \subseteq V \times V$, and the collection of permutations $\{\pi_e\}_{e \in E}$. Using Lemma 8.16, we obtain a weak $\mathbb{F}_q$-integral SDP solution $\mathcal{B} = \{B_u\}_{u \in V}$ of value $1 - O(\sqrt{\eta \log q})$ for $\Phi$.

We construct a vector solution $\{\boldsymbol{V}_i^{B,x} \mid i \in \mathbb{F}_q, B \in \mathcal{B}, x \in \mathbb{F}_q^R\}$ and local distributions $\{\mu_S \mid S \subseteq \mathcal{B} \times \mathbb{F}_q^R\}$ as defined in Section (Section 8.8).

Note that since each set $B \in \mathcal{B}$ correspond to a vertices in $u \in V$, we can view these vectors and local distributions as an SDP solution for the $\mathsf{E2Lin}_q$ instance $\Phi_{\gamma,q}(\Phi)$. Specifically, we make the identifications $\boldsymbol{V}_i^{u,x} := \boldsymbol{V}_i^{B_u,x}$ and $\mu_S := \mu_{\{(B_u,x) \mid (u,x) \in S\}}$ for all $u \in V$, $x \in \mathbb{F}_q^R$, and sets $S \subseteq V \times \mathbb{F}_q^R$.

Lemma 8.49 and Corollary 8.52 show that this SDP solution is $\varepsilon$-infeasible for $\mathsf{SA}_r$ and $\mathsf{LH}_r$, where $\varepsilon = O(r^2 \cdot (3/4)^{t/2})$. The value of the SDP solution for $\Phi_{\gamma,q}(\Phi)$ (see Fig. 8.1) is given by

$$\mathop{\mathbb{E}}_{v\in V} \mathop{\mathbb{E}}_{\substack{w,w'\in N(v)\\ \pi=\pi_{w,v},\ \pi'=\pi_{w',v}}} \mathop{\mathbb{E}}_{\{x,y\}} \mathop{\mathbb{E}}_{c\in\mathbb{F}_q} \sum_{i=1}^{q} \langle \boldsymbol{V}_i^{w,(x\circ\pi+c\cdot\mathbb{1})}, \boldsymbol{V}_{i-c}^{w',y\circ\pi'} \rangle .$$

Using Eq. (8.11)–(8.12),

$$\langle \boldsymbol{V}_i^{w,(x\circ\pi+c\cdot\mathbb{1})}, \boldsymbol{V}_{i-c}^{w',y\circ\pi'} \rangle$$
$$= \tfrac{1}{q^2} + \tfrac{q-1}{q^2} \mathop{\mathbb{P}}_{P\sim\mathcal{P}}[P(B_w) = P(B_{w'})] \cdot \tfrac{1}{n} \sum_{\ell,\ell'\in[R]} \langle \psi(x_{\pi(\ell)} + c - i), \psi(y_{\pi'(\ell')} - (i-c)) \rangle \langle b_{w,\ell}, b_{w',\ell'} \rangle_\psi^t .$$

Note that $\langle \psi(x_{\pi(\ell)} + c - i), \psi(y_{\pi'(\ell')} - (i-c)) \rangle = \langle \psi(x_{\pi(\ell)}), \psi(y_{\pi'(\ell')}) \rangle$. Using Observation 8.14, we have $\pi_{(w,v)}(\ell) = \pi_{(w',v)}(\ell')$ if and only if $\ell = \pi_{B_w \leftarrow B_{w'}}(\ell')$. Hence, by Lemma 8.20,

$$\begin{aligned} &\tfrac{1}{n} \sum_{\ell,\ell'\in[R]} \langle \psi(x_{\pi(\ell)}), \psi(y_{\pi'(\ell')}) \rangle \langle b_{w,\ell}, b_{w',\ell'} \rangle_\psi^t \\ &\qquad = \tfrac{1}{n} \sum_{\ell} \langle \psi(x_{\pi(\ell)}), \psi(y_{\pi(\ell)}) \rangle \langle b_{w,\pi(\ell)}, b_{w',\pi(\ell)} \rangle_\psi^t \quad \pm 2 \cdot r^2 (3/4)^{t/2} \\ &\qquad = \tfrac{1}{n} \sum_{\ell} \langle \psi(x_\ell), \psi(y_\ell) \rangle \rho(B_w, B_{w'})^t \pm O(\varepsilon). \end{aligned}$$

Note that the distribution of $\{x,y\}$ is independent of the vertices $v, w, w'$, and

$$\mathop{\mathbb{E}}_{\{x,y\}} \tfrac{1}{R} \sum_{\ell\in[R]} \langle \psi(x_\ell), \psi(y_\ell) \rangle = 1 - \gamma .$$

Therefore, if we let $\eta_{w,w'} = \rho(B_w, B_{w'})$, we can lower bound the value of the SDP solution as follows

$$\begin{aligned} &\mathop{\mathbb{E}}_{v\in V} \mathop{\mathbb{E}}_{\substack{w,w'\in N(v)\\ \pi=\pi_{w,v},\ \pi'=\pi_{w',v}}} \mathop{\mathbb{E}}_{\{x,y\}} \mathop{\mathbb{E}}_{c\in\mathbb{F}_q} \sum_{i=1}^{q} \langle \boldsymbol{V}_i^{w,(x\circ\pi+c\cdot\mathbb{1})}, \boldsymbol{V}_{i-c}^{w',y\circ\pi'} \rangle \\ &= \mathop{\mathbb{E}}_{v\in V} \mathop{\mathbb{E}}_{w,w'\in N(v)} \left[ \tfrac{1}{q^2} + \tfrac{q-1}{q^2} \mathop{\mathbb{P}}_{P\sim\mathcal{P}}[P(B_w) = P(B_{w'})] \cdot q \cdot \rho(B_w, B_{w'})^t (1-\gamma) \right] \qquad \pm O(\varepsilon) \\ &\geqslant (1-\gamma) \mathop{\mathbb{E}}_{v\in V} \mathop{\mathbb{E}}_{w,w'\in N(v)} \mathop{\mathbb{P}}_{P\sim\mathcal{P}}[P(B_w) = P(B_{w'})] \rho(B_w, B_{w'})^t \qquad \pm O(\varepsilon) \\ &\geqslant (1-\gamma) \mathop{\mathbb{E}}_{v\in V} \mathop{\mathbb{E}}_{w,w'\in N(v)} (1 - O(t\sqrt{\eta_{w,w'}})) \qquad \pm O(\varepsilon) \qquad \text{(using Lemma 8.42)} \end{aligned}$$

Using Jensen's inequality and the fact that $\mathbb{E}_{v,w,w'} \eta_{v,w,w'} = O(\sqrt{\eta \log q})$ (Lemma 8.16), we see that the the value of our SDP solution is at least $1 - \gamma - O(\varepsilon + t\eta^{1/4})$ (recall that we assume $q$ to be constant).

On smoothing the SDP solution using Theorem 8.10, we lose $O(r^2\varepsilon) = O(r^4(3/4)^t)$ in the SDP value. Thus we can set $t = o(\eta^{-1/4})$ and $r = (3/4)^{t/10}$ in order to get a feasible SDP solution for $\mathsf{LH}_r$ with value $1 - \gamma - o_{\eta,\delta}(1)$.

On smoothing the SDP solution using Theorem 8.11, we lose $O(q^r\varepsilon) = O(q^r(3/4)^t)$ in the SDP value. Thus we can set, $t = o(\eta^{-1/4})$ and $r = t/\log^2 q$, we would get a feasible SDP solution for $\mathsf{SA}_r$ with value $1 - \gamma - o_{\eta,\delta}(1)$. □

**Proof of Theorems 8.1–8.2..** Using Theorem 8.23 with the Khot–Vishnoi integrality gap instance (Lemma 8.15), we have $N = 2^{2^{\log(1/\delta)/\eta}}$ and thus $r = 2^{O((\log\log N)^{1/4})}$. Similarly for $\mathsf{SA}_r$, we get $r = O((\log\log N)^{1/4})$.

## 8.9. Smoothing

Let $\Sigma$ be a finite alphabet of size $q$. Let $\{\chi_1, \ldots, \chi_q\}$ be an orthonormal basis for the vector space $\{f : \Sigma \to \mathbb{R}\}$ such that $\chi_1(a) = 1$ for all $a \in \Sigma$. (Here, orthonormal means $\mathbb{E}_{a\in\Sigma} \chi_i(a)\chi_j(a) = \delta_{ij}$ for all $i, j \in [q]$.) For $R \in \mathbb{N}$, let $\left\{\chi_\sigma \mid \sigma \in [q]^R\right\}$ be the orthonormal basis of the vector space $\left\{f : \Sigma^R \to \mathbb{R}\right\}$ defined by

$$\chi_\sigma(x) \stackrel{\text{def}}{=} \chi_{\sigma_1}(x_1) \cdot \cdots \cdot \chi_{\sigma_R}(x_R), \tag{8.16}$$

where $\sigma = (\sigma_1, \ldots, \sigma_R) \in [q]^R$ and $x = (x_1, \ldots, x_R) \in \Sigma^R$.

For a function $f : \Sigma^R \to \mathbb{R}$, we denote

$$\hat{f}(\sigma) \stackrel{\text{def}}{=} \sum_{x\in\Sigma^R} f(x)\chi_\sigma(x). \tag{8.17}$$

Using the fact $\mathbb{E}_{\sigma\in[q]^R} \chi_\sigma(x)\chi_\sigma(y) = \delta_{xy}$ for all $x, y \in \Sigma^R$, we see that

$$f = \mathop{\mathbb{E}}_{\sigma\in[q]^R} \hat{f}(\sigma)\chi_\sigma.$$

We define the following norm for functions $\hat{f} : [q]^R \to \mathbb{R}$,

$$\|\hat{f}\|_1 \stackrel{\text{def}}{=} \sum_{\sigma\in[q]^R} |\hat{f}(\sigma)|.$$

We say $f : \Sigma^R \to \mathbb{R}$ is a *distribution* if $f(x) \geqslant 0$ for all $x \in \Sigma^R$ and $\sum_{x\in\Sigma^R} f(x) = 1$. We define

$$K \stackrel{\text{def}}{=} \max_{\sigma\in[q]^R, x\in\Sigma^R} |\chi_\sigma(x)|.$$

In the next lemma, we give a proof of the following intuitive fact: If a function $g\colon \Sigma^R \to \mathbb{R}$ satisfies the normalization constraint $\sum_{x\in\Sigma^R} g(x) = 1$ and it is close to a distribution in the sense that there exists a distribution $f$ such that $\|\hat{f} - \hat{g}\|$ is small, then $g$ can be made to a distribution by "smoothing" it. Here, smoothing means to move slightly towards the uniform distribution (where every assignment has probability $q^{-R}$).

**Lemma 8.53.** *Let $f, g\colon \Sigma^R \to \mathbb{R}$ be two functions with $\hat{f}(\mathbb{1}) = \hat{g}(\mathbb{1}) = 1$. Suppose $f$ is a distribution. Then, the following function is also a distribution*

$$(1-\varepsilon)g + \varepsilon q^{-R} \quad \textit{where} \varepsilon = \|\hat{f} - \hat{g}\|_1 \cdot K.$$

*Proof.* It is clear that the function $h = (1-\varepsilon)g + \varepsilon q^{-R}$ satisfies the constraint $\hat{h}(\mathbb{1}) = 1$. For every $x \in \Sigma^R$, we have

$$\begin{aligned}
h(x) &= (1-\varepsilon)g(x) + \varepsilon q^{-R} \\
&\geqslant (1-\varepsilon)\Big(g(x) - f(x)\Big) + \varepsilon q^{-R} \qquad \text{(using } f(x) \geqslant 0\text{)} \\
&= \varepsilon q^{-R} + (1-\varepsilon) \mathop{\mathbb{E}}_{\sigma\in[q]^R} \Big(\hat{g}(\sigma) - \hat{f}(\sigma)\Big)\chi_\sigma(x) \\
&\geqslant \varepsilon q^{-R} - (1-\varepsilon) \mathop{\mathbb{E}}_{\sigma\in[q]^R} \Big|\hat{g}(\sigma) - \hat{f}(\sigma)\Big| \cdot K \\
&= \varepsilon q^{-R} - (1-\varepsilon) K \|\hat{f} - \hat{g}\|_1 \cdot q^{-R} \\
&\geqslant 0. \qquad \text{(by our choice of } \varepsilon\text{)}
\end{aligned}$$

□

Let $V$ be a set. For a function $f : \Sigma^V \to \mathbb{R}$ and a subset $S \subseteq V$, we define the function $\text{margin}_S f : \Sigma^S \to \mathbb{R}$ as

$$\text{margin}_S f(x) \stackrel{\text{def}}{=} \sum_{y\in\Sigma^{V\setminus S}} f(x,y).$$

Note that if $f$ is a distribution over $\Sigma$-assignments to $V$ then $\text{margin}_S f$ is its marginal distribution over $\Sigma$-assignments to $T$.

**Lemma 8.54.** *For every $f : \Sigma^V \to \mathbb{R}$ and $S \subseteq V$,*

$$\mathrm{margin}_S f = \mathop{\mathbb{E}}_{\sigma \in [q]^S} \hat{f}(\sigma, \mathbb{1}) \chi_\sigma .$$

*Here, $\sigma, \mathbb{1}$ denotes the $\Sigma$-assignment to $V$ that agrees with $\sigma$ on $S$ and assigns 1 to all variables in $V \setminus S$.*

*Proof.*

$$\begin{aligned}
\mathrm{margin}_S f(x) &= \sum_{y \in \Sigma^{V \setminus S}} f(x, y) \\
&= \sum_{y \in \Sigma^{V \setminus S}} \mathop{\mathbb{E}}_{\sigma \in [q]^V} \hat{f}(\sigma) \chi_\sigma(x, y) \\
&= \sum_{y \in \Sigma^{V \setminus S}} \mathop{\mathbb{E}}_{\sigma \in [q]^S} \mathop{\mathbb{E}}_{\sigma' \in [q]^{V \setminus S}} \hat{f}(\sigma) \chi_\sigma(x) \chi_{\sigma'}(y) \\
&= \mathop{\mathbb{E}}_{\sigma \in [q]^S} \mathop{\mathbb{E}}_{\sigma' \in [q]^{V \setminus S}} \hat{f}(\sigma, \sigma') \chi_\sigma(x) \cdot \sum_{y \in \Sigma^{V \setminus S}} \chi_{\sigma'}(y) \\
&= \mathop{\mathbb{E}}_{\sigma \in [q]^S} \hat{f}(\sigma, \mathbb{1}) \chi_\sigma(x). \quad (\text{using } \textstyle\sum_{y \in [q]^{V \setminus S}} \chi_{\sigma'}(y) = 0 \text{ for } \sigma' \neq \mathbb{1}.)
\end{aligned}$$

□

The margin operator has the following useful property (which is clear from its definiton).

**Lemma 8.55.** *For every function $f : \Sigma^V \to \mathbb{R}$ and any sets $T \subseteq S \subseteq V$,*

$$\mathrm{margin}_T \, \mathrm{margin}_S f = \mathrm{margin}_T f .$$

**Lemma 8.56.** *Let $V$ be a set and let $\left\{\mu_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\right\}$ be a collection of distributions. Suppose that for all sets $A, B \subseteq V$ with $|A|, |B| \leqslant R$,*

$$\|\mathrm{margin}_{A \cap B} \mu_A - \mathrm{margin}_{A \cap B} \mu_B\|_1 \leqslant \eta .$$

*Then, there exists a collection of distributions $\left\{\mu'_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\right\}$ such that*

- *for all $A, B \subseteq V$ with $|A|, |B| \leqslant R$,*

$$\mathrm{margin}_{A \cap B} \mu'_A = \mathrm{margin}_{A \cap B} \mu'_B .$$

– *for all* $S \subseteq V$ *with* $|S| \leqslant R$,

$$\|\mu'_S - \mu_S\|_1 \leqslant O(\eta q^R K^2),$$

The previous lemma is not enough to establish the robustness of our SDP relaxations. The issue is that we not only require that the distributions are consistent among themselves but they should also also be consistent with the SDP vectors.

The following lemma allows us to deal with this issue.

**Lemma 8.57.** *Let* $V$ *be a set and let* $\left\{\mu_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\right\}$ *be a collection of distributions. Suppose that*

– *for all sets* $A, B \subseteq V$ *with* $|A|, |B| \leqslant R$,

$$\|\text{margin}_{A\cap B}\,\mu_A - \text{margin}_{A\cap B}\,\mu_B\|_1 \leqslant \eta\,.$$

– *for all sets* $A, B \subseteq V$ *with* $|A|, |B| \leqslant 2$,

$$\text{margin}_{A\cap B}\,\mu_A = \text{margin}_{A\cap B}\,\mu_B\,.$$

*Then, for* $\varepsilon \geqslant q^R K^2 \eta$, *there exists a collection of distributions* $\left\{\mu'_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\right\}$ *such that*

– *for all* $A, B \subseteq V$ *with* $|A|, |B| \leqslant R$,

$$\text{margin}_{A\cap B}\,\mu'_A = \text{margin}_{A\cap B}\,\mu'_B\,. \tag{8.18}$$

– *for all* $S \subseteq V$ *with* $|S| \leqslant R$,

$$\|\mu'_S - \mu_S\|_1 \leqslant O(K^2 \eta q^R), \tag{8.19}$$

– *for all* $S \subseteq V$ *with* $|S| \leqslant 2$,

$$\mu'_S = (1-\varepsilon)\mu_S + \varepsilon \cdot q^{-|S|}. \tag{8.20}$$

*Proof.* For $\sigma \in [q]^V$, let $\text{supp}(\sigma)$ denote the set of coordinates of $\sigma$ not equal to $1$, and let $|\sigma|$ denote the number of such coordinates,

$$\text{supp}(\sigma) \stackrel{\text{def}}{=} \{i \in V \mid \sigma_i \neq 1\} \quad \text{and} \quad |\sigma| \stackrel{\text{def}}{=} |\text{supp}(\sigma)|\,.$$

For every $\sigma \in [q]^V$ with $|\sigma| \leqslant R$, we define

$$\hat{f}(\sigma) := \mathop{\mathbb{E}}_{x \sim \mu_S} \chi_\sigma(x) \qquad \text{where} S = \operatorname{supp}(\sigma).$$

For every $\sigma$ with $|\sigma| > R$, we set $\hat{f}(\sigma) := 0$. We define $\mu'_S$ in terms of $f = \mathbb{E}_\sigma \hat{f}(\sigma)\chi_\sigma$,

$$\mu'_S := \operatorname{margin}_S (1-\varepsilon) f + \varepsilon q^{-|V|}.$$

By Lemma 8.55, this choice of $\mu'_S$ satisfies condition (8.18).

First, let us argue that the functions $\mu'_S$ are distributions. Let $S \subseteq V$ with $|S| \leqslant R$. For $\sigma \in [q]^S$ with $T := \operatorname{supp}(\sigma) \subseteq S$, we have

$$\begin{aligned} |\hat{f}(\sigma, \mathbb{1}) - \mathop{\mathbb{E}}_{x \sim \mu_S} \chi_\sigma(x)| &= |\mathop{\mathbb{E}}_{x \sim \mu_T} \chi_\sigma(x) - \mathop{\mathbb{E}}_{x \sim \mu_S} \chi_\sigma(x)| \\ &\leqslant \|\mu_T - \operatorname{margin}_T \mu_S\|_1 \cdot \max|\chi_\sigma| \\ &\leqslant \eta \cdot K. \end{aligned} \tag{8.21}$$

Let $f_S$ denote the function $\operatorname{margin}_S f$. By Lemma 8.54, $\hat{f}_S(\sigma) = \hat{f}(\sigma, \mathbb{1})$ for all $\sigma \in [q]^S$. Hence, $\|\hat{f} - \hat{\mu}_S\|_1 \leqslant q^R \cdot K\eta$. It follows that for $\varepsilon \geqslant q^R K^2 \eta$, the function $\mu'_S = (1-\varepsilon) f_S + \varepsilon q^{-|S|}$ is a distribution (using Lemma 8.53).

Next, let us verify that (8.19) holds. We have

$$\begin{aligned} \|\mu'_S - \mu_S\|_1 &\leqslant O(\varepsilon) + \|\operatorname{margin}_S f - \mu_S\|_1 \\ &\overset{\text{La. 8.54}}{=} O(\varepsilon) + \left\| \mathop{\mathbb{E}}_{\sigma \in [k]^S} \left( \hat{f}(\sigma, \mathbb{1}) - \mathop{\mathbb{E}}_{x \sim \mu_S} \chi_\sigma(x) \right) \chi_\sigma \right\|_1 \\ &\overset{(8.21)}{\leqslant} O\left(\eta K^2 \cdot k^R\right) \quad (\text{using } |\hat{f}(\sigma, \mathbb{1}) - \hat{\mu}_S(\sigma)| \leqslant \eta K \text{ and } |\chi_\sigma(x)| \leqslant K). \end{aligned}$$

Finally, we show that the new distributions satisfy (8.20). Let $S \subseteq V$ be a set of size at most 2. It follows from the consistency assumption that for all $\sigma \in [k]^S$, we have $\hat{f}(\sigma, \mathbb{1}) = \hat{\mu}_S(\sigma)$. Hence, $f_S = \mu_S$, which implies (8.20).

□

**Lemma 8.58.** *Let $V$ be a set and let $\left\{\mu_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\right\}$ be a collection of distributions. Suppose that*

- *for all sets $A, B \subseteq V$ with $|A|, |B| \leqslant R$,*

$$\|\operatorname{margin}_{A \cap B} \mu_A - \operatorname{margin}_{A \cap B} \mu_B\|_1 \leqslant \eta.$$

– *for all sets* $A, B \subseteq V$ *with* $|A|, |B| \leqslant 2$,

$$\operatorname{margin}_{A\cap B} \mu_A = \operatorname{margin}_{A\cap B} \mu_B .$$

*Then, for* $\varepsilon \geqslant kR^2K^2\eta$, *there exists a collection of distributions* $\left\{\mu'_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\right\}$ *such that*

– *for all* $A, B \subseteq V$ *with* $|A|, |B| \leqslant R$ *with* $|A \cap B| \leqslant 2$,

$$\operatorname{margin}_{A\cap B} \mu'_A = \operatorname{margin}_{A\cap B} \mu'_B . \tag{8.22}$$

– *for all* $S \subseteq V$ *with* $|S| \leqslant R$,

$$\|\mu'_S - \mu_S\|_1 \leqslant O(K^2\eta kR^2), \tag{8.23}$$

– *for all* $S \subseteq V$ *with* $|S| \leqslant 2$,

$$\mu'_S = (1-\varepsilon)\mu_S + \varepsilon \cdot k^{-|S|}. \tag{8.24}$$

*Proof.* The proof is along the lines of the proof of the previous lemma.

Define $\hat{f} : [k]^R \to \mathbb{R}$ as before. We define new functions $\{\mu^*_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\}$ such that

$$\hat{\mu}^*_S(\sigma) = \begin{cases} \hat{\mu}_S(\sigma) & \text{if } \operatorname{supp}(\sigma) > 2, \\ \hat{f}(\sigma, \mathbb{1}) & \text{if } 1 \leqslant \operatorname{supp}(\sigma) \leqslant 2, \\ 1 & \text{otherwise.} \end{cases}$$

Since $|\hat{f}(\sigma, \mathbb{1}) - \hat{\mu}_S(\sigma)| \leqslant K\eta$ (see proof of previous lemma), we can upper bound $\|\hat{\mu}^*_S - \hat{\mu}_S\|_1 \leqslant kR^2 \cdot K\eta$ (there are not more than $kR^2$ different $\sigma \in [k]^S$ with $\hat{f}(\sigma, \mathbb{1}) \neq \hat{\mu}_S(\sigma)$. By Lemma 8.53, for $\varepsilon \geqslant kR^2K^2\eta$, the functions $\{\mu'_S : \Sigma^S \to \mathbb{R} \mid S \subseteq V, |S| \leqslant R\}$ defined by $\mu'_S := (1-\varepsilon)\mu^*_S + \varepsilon k^{-|S|}$ are the desired distributions. We can check that the assertions of the lemma are satisfied in the same way as for the proof of the previous lemma. □

*Proofs of Theorem 8.10 and Theorem 8.11 (Sketch).* We apply Lemma 8.58 or Lemma 8.57 to the local distributions $\{\mu_S\}$ of the $\varepsilon$-infeasible $\mathsf{LH}_R$ or $\mathsf{SA}_R$ solution, respectively. We get a new set of local distributions $\{\mu'_S\}$ that have the desired consistency properties. It remains to change the vectors so that their inner product match the corresponding probabilities in the local distributions. Suppose $\{\boldsymbol{v}_{i,a}\}$ is the original vector assignment. Let $\{\boldsymbol{u}_{i,a}\}$ be the vector assignment that corresponds to the uniform distribution over all possible assignments to the variables (this vector assignment is the geometric center of the set of all vector assignments). Then, we define the new vector assignment $\{\boldsymbol{v}'_{i,a}\}$ as

$$\boldsymbol{v}'_{i,a} = \sqrt{1-\delta} \cdot \boldsymbol{v}_{i,a} \oplus \sqrt{\delta}\boldsymbol{u}_{i,a},$$

where $\delta$ is the smoothing parameter in Lemma 8.58 or Lemma 8.57. It is easy to verify that $\{\boldsymbol{v}'_{i,a}\}$ together with $\{\mu'_S\}$ form a feasible $\mathsf{LH}_R$ or $\mathsf{SA}_R$ solution. □

## 8.10. Notes

The material presented in this chapter is based on the paper "Integrality Gaps for Strong SDP Relaxations of Unique Games" [RS09a], joint with Prasad Raghavendra. A preliminary version appeared at FOCS 2009.

### Related Work

In a breakthrough result, Arora et al. [ARV04] used a strong semidefinite program with triangle inequalities to obtain $O(\sqrt{\log n})$ approximation for the Sparsest Cut problem. Inspired by this work, stronger semidefinite programs have been utilized to obtain better approximation algorithms for certain graph coloring problems [Chl07, ACC06, CS08]. We wish to point out that the work of Chlamtac and Singh [CS08] uses the $\mathsf{SA}_r$ hierarchy to obtain approximation algorithms for the hypergraph coloring problem.

In this light, hierarchies of stronger SDP relaxations such as Lovász–Schriver [LS91], Lasserre [Las01], and Sherali–Adams hierarchies [SA90] (See [Lau03] for a comparison) have emerged as possible avenues to obtain better approximation ratios.

Considerable progress has been made in understanding the limits of linear programming hierarchies. Building on a sequence of works [ABL02, ABLT06, Tou05, Tou06], Schoenebeck et al. [STT07a] obtained a $2-\varepsilon$-factor integrality gap for $\Omega(n)$ rounds of Lovász–Schriver LS hierarchy. More recently, Charikar et al. [CMM09] constructed integrality gaps for $\Omega(n^\delta)$ rounds of Sherali–Adams hierarchy for several problems like Max Cut, Vertex Cover, Sparsest Cut and Maximum Acyclic Subgraph. Furthermore, the same work also exhibits $\Omega(n^\delta)$-round Sherali–Adams integrality gap for Unique Games, in turn obtaining a corresponding gap for every problem to which Unique Games is reduced to.

Lower bound results of this nature are fewer in the case of semidefinite programs. A $\Omega(n)$ $LS_+$ round lower bound for proving unsatisfiability of random 3-SAT formulae was obtained in [BOGH+06, AAT05]. In turn, this leads to $\Omega(n)$-round $LS_+$ integrality gaps for problems like Set Cover, Hypergraph Vertex Cover where a matching NP-hardness result is known. Similarly, the $\frac{7}{6}$-integrality gap for $\Omega(n)$ rounds of $LS_+$ in [STT07b] falls in a regime where a matching NP-hardness result has been shown to hold. A significant exception is the result of Georgiou et al. [GMPT07] that exhibited a $2-\varepsilon$-integrality gap for $\Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$ rounds of $LS_+$ hierarchy. More recently, building on the beautiful work of [Sch08] on Lasserre integrality gaps for Random 3-SAT, Tulsiani [Tul09] obtained a $\Omega(n)$-round Lasserre integrality gap matching the corresponding UG-hardness for $k$-CSP [ST06].

# 9. Open Problems

**Number of Large Eigenvalues versus Small-Set Expansion.** Consider the following hypothesis.

**Hypothesis 9.1.** *There exists an absolute constant $\varepsilon_0 > 0$ such that for every constant $\delta > 0$, there exists a function $f : \mathbb{N} \to \mathbb{N}$ with $f(n) \leqslant n^{o(1)}$ such that every graph with $n$ vertices and at least $f(n)$ eigenvalues larger than $1 - \varepsilon_0$ contains a vertex set with volume $\delta$ and expansion at most $1 - \varepsilon_0$.*

The results in Section 5.2 (Subspace Enumeration), e.g. Theorem 5.7, demonstrate that this hypothesis implies that the Small-Set Expansion Hypothesis is false (assuming that NP does not $\exp(n^{o(1)})$-time algorithms).

The graphs that come closest to refuting Hypothesis 9.1 are Boolean noise graph (which are important tools for UG-hardness reductions). There exists (Boolean noise) graphs with $n$ vertices and $\operatorname{poly}\log n$ eigenvalues larger than $1 - 1/\log\log n$, but no set with volume $O(1/\log n)$ and expansion less than $1/2$. These graphs demonstrate that even if Hypothesis 9.1 were true, the function $f(n)$ in the hypothesis must be larger than $\operatorname{poly}\log n$.

We conclude that a resolution of Hypothesis 9.1 would either lead to an algorithm refuting the Small-Set Expansion Hypothesis or it would lead to a family of graphs with "better properties" (in the sense above) than Boolean noise graphs. Such a family of graphs could potentially lead to more efficient hardness reductions or improved integrality gap constructions.

**Subexponential Algorithm for Sparsest Cut.** An interesting question raised by the subexponential algorithms for Small-Set Expansion and Unique Games (see Chapter 5) is whether similar subexponential algorithms exists for other problems that are SSE-hard or UG-hard. Concretely, we can ask if for every constant $\varepsilon > 0$, there exists an algorithm for Sparsest Cut with constant approximation ratio $C = C(\varepsilon)$ and running time $\exp(n^{\varepsilon})$. A positive resolution of the following conjecture would imply such an algorithm. (We omit the proof at this point. It follows from extensions of the results presented in Section 5.2, techniques in [ARV09], and an alternative characterization of the threshold rank of a graph.)

**Conjecture 9.2.** *For every $\varepsilon > 0$, there exists positive constants $\eta = \eta(\varepsilon)$ and $\delta = \delta(\varepsilon)$ such that the following holds: For every collection of unit vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ with $\mathbb{E}_{i,j\in[n]} \left|\langle v_i, v_j\rangle\right| \leqslant n^{-\varepsilon}$, there exists two sets $S, T \subseteq \{1, \ldots, n\}$ with $|S|, |T| \geqslant \delta n$ and $\|v_i - v_j\|^2 \geqslant \eta$ for all $i \in S$ and $j \in T$.*

In words, the conjecture says that if the *global correlation* $\mathbb{E}_{i,j}|\langle v_i, v_j\rangle|$ of a collection of unit vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ is polynomially small ($\leqslant n^{-\varepsilon}$), then an $\Omega_\varepsilon(1)$ fraction of the vectors is $\Omega_\varepsilon(1)$-separated from another $\Omega_\varepsilon(1)$ fraction of the vectors. (An alternative formulation is that the graph formed by joining any two vectors with distance $\eta = \eta(\varepsilon) > 0$ is not a good vertex expander.)

The configuration of vectors that comes closest to refuting Conjecture 9.2 is the $\Omega(\sqrt{d})$-fold tensored Boolean hypercube $\{\pm 1/\sqrt{d}\}^d$. Here, one can show that any two subsets containing an $\Omega(1)$ fraction of the vectors are not $\Omega(1)$ separated (using the vertex expansion of the usual graph defined on Boolean hypercube). However, the global correlation of this configuration is at least $2^{-\tilde{O}(\sqrt{\log n})}$, which is not polynomially small.

Resolving Conjecture 9.2 would either lead to an interesting algorithm for Sparsest Cut or to a family of geometric graphs with "nicer properties" (in the sense above) than the Boolean hypercube.

# Bibliography


[AAT05] Michael Alekhnovich, Sanjeev Arora, and Iannis Tourlakis, *Towards strong nonapproximability results in the Lovász–Schrijver hierarchy*, STOC, 2005, pp. 294–303. 235

[ABL02] Sanjeev Arora, Béla Bollobás, and László Lovász, *Proving integrality gaps without knowing the linear program*, FOCS, 2002, pp. 313–322. 11, 178, 235

[ABLT06] Sanjeev Arora, Béla Bollobás, László Lovász, and Iannis Tourlakis, *Proving integrality gaps without knowing the linear program*, Theory of Computing **2** (2006), no. 1, 19–51. 235

[ABS10] Sanjeev Arora, Boaz Barak, and David Steurer, *Subexponential algorithms for unique games and related problems*, Manuscript, 2010. 20, 118

[ACC06] Sanjeev Arora, Eden Chlamtac, and Moses Charikar, *New approximation guarantee for chromatic number*, STOC, 2006, pp. 215–224. 234

[AIMS10] Sanjeev Arora, Russell Impagliazzo, William Matthews, and David Steurer, *Improved algorithms for unique games via divide and conquer*, Tech. Report ECCC TR10-041, ECCC, 2010. 69, 121

[AK09] Noga Alon and Bo'az Klartag, *Economical toric spines via Cheeger's inequality*, J. Topol. Anal. **1** (2009), no. 2, 101–111. MR MR2541756 94

[AKK+08] Sanjeev Arora, Subhash Khot, Alexandra Kolla, David Steurer, Madhur Tulsiani, and Nisheeth K. Vishnoi, *Unique games on expanding constraint graphs are easy*, STOC, 2008, pp. 21–28. 19, 67, 68, 200, 210, 211, 246

# Bibliography


[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy, *Proof verification and the hardness of approximation problems*, J. ACM **45** (1998), no. 3, 501–555. 4

[Alo86] Noga Alon, *Eigenvalues and expanders*, Combinatorica **6** (1986), no. 2, 83–96. 14

[AM85] Noga Alon and V. D. Milman, $\lambda_1$, *isoperimetric inequalities for graphs, and superconcentrators*, J. Comb. Theory, Ser. B **38** (1985), no. 1, 73–88. 14

[AMS07] Christoph Ambühl, Monaldo Mastrolilli, and Ola Svensson, *Inapproximability results for sparsest cut, optimal linear arrangement, and precedence constrained scheduling*, FOCS, 2007, pp. 329–337. 14, 160

[ARV04] Sanjeev Arora, Satish Rao, and Umesh V. Vazirani, *Expander flows, geometric embeddings and graph partitioning*, STOC, 2004, pp. 222–231. 14, 234

[ARV09] ______, *Expander flows, geometric embeddings and graph partitioning*, J. ACM **56** (2009), no. 2. 124, 237

[AS98] Sanjeev Arora and Shmuel Safra, *Probabilistic checking of proofs: A new characterization of NP*, J. ACM **45** (1998), no. 1, 70–122. 4

[Aus07] Per Austrin, *Towards sharp inapproximability for any 2-CSP*, FOCS, 2007, pp. 307–317. 180, 181

[BHH+08] Boaz Barak, Moritz Hardt, Ishay Haviv, Anup Rao, Oded Regev, and David Steurer, *Rounding parallel repetitions of unique games*, FOCS, 2008, pp. 374–383. 20, 93, 94

[BOGH+06] Joshua Buresh-Oppenheim, Nicola Galesi, Shlomo Hoory, Avner Magen, and Toniann Pitassi, *Rank bounds and integrality gaps for cutting planes procedures*, Theory of Computing **2** (2006), no. 1, 65–90. 235

[Bor85] Christer Borell, *Geometric bounds on the Ornstein-Uhlenbeck velocity process*, Z. Wahrsch. Verw. Gebiete **70** (1985), no. 1, 1–13. MR 795785 (87k:60103) 46

[Bro97] Andrei Broder, *On the resemblance and containment of documents*, Compression and Complexity of Sequences, 1997, pp. 21–29. 95

[Che70] Jeff Cheeger, *A lower bound for the smallest eigenvalue of the Laplacian*, Problems in analysis (Papers dedicated to Salomon Bochner, 1969), Princeton Univ. Press, Princeton, N. J., 1970, pp. 195–199. MR MR0402831 (53 #6645) 12, 13, 31

[Chl07] Eden Chlamtac, *Approximation algorithms using hierarchies of semidefinite programming relaxations*, FOCS, 2007, pp. 691–701. 234

[CKK+06] Shuchi Chawla, Robert Krauthgamer, Ravi Kumar, Yuval Rabani, and D. Sivakumar, *On the hardness of approximating multicut and sparsest-cut*, Computational Complexity **15** (2006), no. 2, 94–114. 181

[CMM06a] Moses Charikar, Konstantin Makarychev, and Yury Makarychev, *Near-optimal algorithms for unique games*, STOC, 2006, pp. 205–214. 20, 72, 208, 209

[CMM06b] Eden Chlamtac, Konstantin Makarychev, and Yury Makarychev, *How to play unique games using embeddings*, FOCS, 2006, pp. 687–696. 47, 95

[CMM09] Moses Charikar, Konstantin Makarychev, and Yury Makarychev, *Integrality gaps for Sherali–Adams relaxations*, STOC, 2009, pp. 283–292. 179, 180, 183, 235

[CS08] Eden Chlamtac and Gyanit Singh, *Improved approximation guarantees through higher levels of sdp hierarchies*, APPROX-RANDOM, 2008, pp. 49–62. 234

[DI98] Tassos Dimitriou and Russell Impagliazzo, *Go with the winners for graph bisection*, SODA, 1998, pp. 510–520. 32

[DKSV06] Nikhil R. Devanur, Subhash Khot, Rishi Saket, and Nisheeth K. Vishnoi, *Integrality gaps for sparsest cut and minimum linear arrangement problems*, STOC, 2006, pp. 537–546. 182

[Dod84] Jozef Dodziuk, *Difference equations, isoperimetric inequality and transience of certain random walks*, Trans. Amer. Math. Soc. **284** (1984), no. 2, 787–794. MR 743744 (85m:58185) 14

[Fei02] Uriel Feige, *Relations between average case complexity and approximation complexity*, IEEE Conference on Computational Complexity, 2002, p. 5. 52

[FKO07] Uriel Feige, Guy Kindler, and Ryan O'Donnell, *Understanding parallel repetition requires understanding foams*, IEEE Conference on Computational Complexity, 2007, pp. 179–192. 9

[FL92] Uriel Feige and László Lovász, *Two-prover one-round proof systems: Their power and their problems (extended abstract)*, STOC, 1992, pp. 733–744. 39, 73

[GMPT07] Konstantinos Georgiou, Avner Magen, Toniann Pitassi, and Iannis Tourlakis, *Integrality gaps of* $2 - o(1)$ *for vertex cover sdps in the Lovász–Schrijver hierarchy*, FOCS, 2007, pp. 702–712. 235

[GMR08] Venkatesan Guruswami, Rajsekar Manokaran, and Prasad Raghavendra, *Beating the random ordering is hard: Inapproximability of maximum acyclic subgraph*, FOCS, 2008, pp. 573–582. 180, 181, 183

[GMT06] Sharad Goel, Ravi Montenegro, and Prasad Tetali, *Mixing time bounds via the spectral profile*, Electron. J. Probab. **11** (2006), no. 1, 1–26 (electronic). MR MR2199053 (2007e:60075) 32

[GW95] Michel X. Goemans and David P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. ACM **42** (1995), no. 6, 1115–1145. 3, 204

[Hås01] Johan Håstad, *Some optimal inapproximability results*, J. ACM **48** (2001), no. 4, 798–859. 4

[Hol09] Thomas Holenstein, *Parallel repetition: Simplification and the no-signaling case*, Theory of Computing **5** (2009), no. 1, 141–172. 93, 95

[IPZ01] R. Impagliazzo, R. Paturi, and F. Zane, *Which problems have strongly exponential complexity?*, Journal of Computer and System Sciences **63** (2001), no. 4, 512–530. 2, 5, 96, 123

[Kar72] Richard M. Karp, *Reducibility among combinatorial problems*, Complexity of Computer Computations, 1972, pp. 85–103. 2

[Kho02] Subhash Khot, *On the power of unique 2-prover 1-round games*, STOC, 2002, pp. 767–775. 6, 41

[Kho10] ———, *On the unique games conjecture (invited survey)*, IEEE Conference on Computational Complexity, 2010, pp. 99–121. 7

[KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell, *Optimal inapproximability results for MAX-CUT and other 2-variable CSPs?*, SIAM J. Comput. **37** (2007), no. 1, 319–357. 41, 47, 180, 181, 184, 185, 186, 201, 202

[Kol10] Alexandra Kolla, *Spectral algorithms for unique games*, IEEE Conference on Computational Complexity, 2010, pp. 122–130. 69, 70, 119, 120

[KORW08] Guy Kindler, Ryan O'Donnell, Anup Rao, and Avi Wigderson, *Rounding schemes and cubical tilings with sphere-like surface area*, FOCS, 2008. 94

[KRT08] Julia Kempe, Oded Regev, and Ben Toner, *Unique games with entangled provers are easy*, FOCS, 2008, pp. 457–466. 39, 95

[KT02] Jon M. Kleinberg and Éva Tardos, *Approximation algorithms for classification problems with pairwise relationships: metric labeling and markov random fields*, J. ACM **49** (2002), no. 5, 616–639. 95

[KT07] Alexandra Kolla and Madhur Tulsiani, *Playing random and expanding unique games*, Unpublished manuscript available from the authors' webpages, to appear in journal version of [AKK+08], 2007. 69, 70, 119, 120

[KV05] Subhash Khot and Nisheeth K. Vishnoi, *The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into* $\ell_1$, FOCS, 2005, pp. 53–62. 68, 69, 120, 179, 180, 181, 182, 183, 184, 185, 186, 189, 197, 200, 201

[Las01] Jean B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM Journal on Optimization **11** (2001), no. 3, 796–817. 52, 234

[Lau03] Monique Laurent, *A comparison of the sherali-adams, lovász-schrijver, and lasserre relaxations for 0-1 programming*, Math. Oper. Res. **28** (2003), no. 3, 470–496. 234

[LR99] Frank Thomson Leighton and Satish Rao, *Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms*, J. ACM **46** (1999), no. 6, 787–832. 14, 124

[LS91] L. Lovász and A. Schrijver, *Cones of matrices and set-functions and 0-1 optimization*, SIAM J. Optim. **1** (1991), no. 2, 166–190. 234

[MM09] Konstantin Makarychev and Yury Makarychev, *How to play unique games on expanders*, CoRR **abs/0903.0367** (2009). 69

[MNRS08] Rajsekar Manokaran, Joseph Naor, Prasad Raghavendra, and Roy Schwartz, *SDP gaps and UGC-hardness for multiway cut, 0-extension, and metric labeling*, STOC, 2008, pp. 11–20. 180, 181, 183

[MOO05] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz, *Noise stability of functions with low influences invariance and optimality*, FOCS, 2005, pp. 21–30. 47

[MR08] Dana Moshkovitz and Ran Raz, *Two query pcp with sub-constant error*, FOCS, 2008, pp. 314–323. 5, 9, 96

[MS07] Rajat Mittal and Mario Szegedy, *Product rules in semidefinite programming*, FCT, 2007, pp. 435–445. 39

[Rag08] Prasad Raghavendra, *Optimal algorithms and inapproximability results for every CSP?*, STOC, 2008, pp. 245–254. 7, 42, 180, 183

[Rao08] Anup Rao, *Parallel repetition in projection games and a concentration bound*, STOC, 2008, pp. 1–10. 124

[Raz98] Ran Raz, *A parallel repetition theorem*, SIAM J. Comput. **27** (1998), no. 3, 763–803. 4, 71

[Raz08] ______, *A counterexample to strong parallel repetition*, FOCS, 2008, pp. 369–373. 39, 93, 94

[RS09a] Prasad Raghavendra and David Steurer, *Integrality gaps for strong SDP relaxations of unique games*, FOCS, 2009, pp. 575–585. 23, 234

[RS09b] ———, *Towards computing the Grothendieck constant*, SODA, 2009, pp. 525–534. 180, 181

[RS10] ———, *Graph expansion and the unique games conjecture*, STOC, 2010, To appear. 21, 69, 159

[RST10a] Prasad Raghavendra, David Steurer, and Prasad Tetali, *Approximations for the isoperimetric and spectral profile of graphs and related parameters*, STOC, 2010, To appear. 32, 125

[RST10b] Prasad Raghavendra, David Steurer, and Madhur Tulsiani, *Reductions between expansion problems*, Manuscript, 2010. 22, 176

[SA90] Hanif D. Sherali and Warren P. Adams, *A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems*, SIAM J. Discrete Math. **3** (1990), no. 3, 411–430. MR MR1061981 (91k:90116) 180, 234

[Sch08] Grant Schoenebeck, *Linear level Lasserre lower bounds for certain k-CSPs*, FOCS, 2008, pp. 593–602. 52, 235

[ST06] Alex Samorodnitsky and Luca Trevisan, *Gowers uniformity, influence of variables, and PCPs*, STOC, 2006, pp. 11–20. 235

[Ste10a] David Steurer, *Fast SDP algorithms for constraint satisfaction problems*, SODA, 2010, To appear. 42

[Ste10b] ———, *Improved rounding for parallel repeated unique games*, APPROX-RANDOM, 2010, pp. 724–737. 20, 93, 94

[STT07a] Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani, *A linear round lower bound for lovasz-schrijver sdp relaxations of vertex cover*, IEEE Conference on Computational Complexity, 2007, pp. 205–216. 183, 235

[STT07b] ———, *Tight integrality gaps for Lovász–Schrijver LP relaxations of vertex cover and max cut*, STOC, 2007, pp. 302–310. 180, 235

[Tou05] Iannis Tourlakis, *Towards optimal integrality gaps for hypergraph vertex cover in the lovász-schrijver hierarchy*, APPROX-RANDOM, 2005, pp. 233–244. 235

[Tou06] ______, *New lower bounds for vertex cover in the Lovász–Schrijver hierarchy*, IEEE Conference on Computational Complexity, 2006, pp. 170–182. 235

[Tre05] Luca Trevisan, *Approximation algorithms for unique games*, FOCS, 2005, pp. 197–205. 68, 121, 210

[TSSW00] Luca Trevisan, Gregory B. Sorkin, Madhu Sudan, and David P. Williamson, *Gadgets, approximation, and linear programming*, SIAM J. Comput. **29** (2000), no. 6, 2074–2097 (electronic). 5

[Tul09] Madhur Tulsiani, *CSP gaps and reductions in the Lasserre hierarchy*, STOC, 2009, pp. 303–312. 180, 183, 235