Title
Hardness of Maximum Constraint Satisfaction

Author
Chan, Siu On

Publication Date
2013

Peer reviewed|Thesis/dissertation

Hardness of Maximum Constraint Satisfaction

by

Siu On Chan

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Elchanan Mossel, Chair
Professor Luca Trevisan
Professor Satish Rao
Professor Michael Christ

Spring 2013

Hardness of Maximum Constraint Satisfaction

by
Siu On Chan

Abstract

Hardness of Maximum Constraint Satisfaction

by

Siu On Chan

Doctor of Philosophy in Computer Science

University of California, Berkeley

Professor Elchanan Mossel, Chair

Maximum constraint satisfaction problem (Max-CSP) is a rich class of combinatorial optimization problems. In this dissertation, we show optimal (up to a constant factor) NP-hardness for maximum constraint satisfaction problem with k variables per constraint (Max-k-CSP), whenever k is larger than the domain size. This follows from our main result concerning CSPs given by a predicate: a CSP is approximation resistant if its predicate contains a subgroup that is balanced pairwise independent. Our main result is related to previous works conditioned on the Unique-Games Conjecture and integrality gaps in sum-of-squares semidefinite programming hierarchies.

Our main ingredient is a new gap-amplification technique inspired by XOR-lemmas. Using this technique, we also improve the NP-hardness of approximating Independent-Set on bounded-degree graphs, Almost-Coloring, Two-Prover-One-Round-Game, and various other problems.

Dedicated to my family

Contents

Acknowledgments

I am indebted to the following people for numerous helpful discussions and continual encouragement: Siavosh Benabbas, Anand Bhaskar, Siu Man Chan, Kelly Sin Man Choi, James Cook, Anindya De, Ilias Diakonikolas, Lap Chi Lau, Mohammad Moharrami, Elchanan Mossel, Lorenzo Orecchia, Prasad Raghavendra, Satish Rao, Grant Schoenebeck, Piyush Srivastava, Luca Trevisan, Madhur Tulsiani, Gregory Valiant, Thomas Vidick, Thomas Watson, and Yi Wu.

I thank Luca Trevisan for pointing out limitations of previous techniques, and Johan Håstad for allowing me to include his predicates in this paper. I am grateful to Venkatesan Guruswami, Subhash Khot, Bundit Laekhanukit, Urmila Mahadev, and Yury Makarychev for suggesting or providing references. I thank Thomas Watson and an anonymous referee for many suggestions that greatly enhance the presentation of the paper. Special thanks to Piyush Srivastava for collaboration on related aspects of this research.

Chapter 1

Introduction

Max-k-CSP is the task of satisfying the maximum fraction of constraints when each constraint involves k variables. Despite much progress on this problem, there remains a huge multiplicative gap between NP-hardness and algorithmic results. When the domain Σ is boolean, the best algorithm by Makarychev and Makarychev [2012] has approximation ratio $\Omega(k/2^k)$, but the best NP-hardness result by Engebretsen and Holmerin [2008] has hardness ratio $2^{O(\sqrt{k})}/2^k$, which is significantly larger by the factor $2^{\Omega(\sqrt{k})}$.

A related question is to identify constraint satisfaction problems (CSPs) that are extremely hard to approximate, so much so that they are NP-hard to approximate better than just outputting a random assignment. Such CSPs are called approximation resistant; famous examples include Max-3-SAT and Max-3-XOR [Håstad 2001]. Previous works focused on CSPs whose constraints involve the same number k of *literals*, and each constraint accepts the same collection $C \subseteq \Sigma^k$ of local assignments. A lot is known about such CSPs of arity at most four. For arity two, a CSP is never approximation resistant [Håstad 2008]. For arity three, a boolean CSP is approximation resistant precisely when its predicate C contains all even-parity or all odd-parity bitstrings [Håstad 2001; Zwick 1998]. For arity four, an extensive study was made by Hast [2005b]. But for higher arity, results were scattered: four families of approximation resistant CSPs were known in [Håstad 2001, Theorem 5.9], [Engebretsen and Holmerin 2008], [Hast 2005b, Theorem 5.2], and [Håstad 2011] (apart from CSPs obtained by padding irrelevant variables).

To make progress, conditional results were obtained assuming the Unique-Games Conjecture of Khot [2002b]. Under this conjecture, Samorodnitsky and Trevisan [2009] showed that Max-k-CSP is NP-hard to approximate beyond $O(k/2^k)$, matching the best algorithm up to a constant factor, and later Raghavendra [2008] obtained optimal inapproximability (and algorithmic) results for every CSP. Under the same conjecture, Austrin and Mossel [2009] showed that a CSP is approximation resistant if its predicate supports a balanced pairwise independent distribution. However, the UG conjecture remains uncertain, and it is desirable to look for new hardness reduction techniques.

In this work, we obtain a general criterion for approximation resistance (unifying [Håstad 2001, Theorem 5.9], [Engebretsen and Holmerin 2008], [Hast 2005b, Theorem 5.2], and

[Håstad 2011, Theorem 4][1]), and settle the NP-hardness of Max-k-CSP (up to a constant factor and modulo P $\neq$ NP). We show hardness for CSPs whose domain is an *abelian group* G, and whose predicate $C \subseteq G^k$ is a *subgroup* satisfying a condition similar to Austrin and Mossel [2009] (see Chapter 4 for definitions).

To state our results, we say it is NP-hard to (c, s)-decide a Max-CSP if given an instance M of the CSP, it is NP-hard to decide whether the best assignment to M satisfies at least c fraction of constraints, or at most s fraction. The parameters c and s are known as completeness and soundness, respectively. The hardness ratio is s/c.

Theorem 1.1 (Main). *Let $k \geqslant 3$ be an integer, G a finite abelian group, and C a balanced pairwise independent subgroup of G^k. For some $\varepsilon = o_{n;k,|G|}(1)$,[2] it is* NP*-hard to $(1 - \varepsilon, |C|/|G|^k + \varepsilon)$-decide an* Max-$C$ *instance of size n.*

A random assignment satisfies $|C|/|G|^k$ fraction of constraints in expectation, so our hardness ratio is tight. Like Austrin and Mossel [2009], we actually show *hereditary* approximation resistance, i.e., any predicate *containing* a pairwise independent subgroup also yields an approximation resistant CSP. Compared with Austrin and Mossel's, our result requires an abelian subgroup structure on the predicate, but avoids their UG Conjecture assumption. Consequently, we throw away the same assumption in an earlier result of Håstad [2009], showing that almost all CSPs given by a predicate are hereditarily approximation resistant, answering his open problem.

Our result is inspired by integrality gaps for sum-of-squares programs. Direct construction of such integrality gaps by Schoenebeck [2008] and Tulsiani [2009] requires both pairwise independence and abelian subgroup structure — abelian subgroup seems indispensable in the Fourier-analytic construction of SDP solution, and in this case balanced pairwise independence is *necessary* [Chan and Molloy 2013] for another ingredient of the construction, namely exponential resolution complexity of random instances. Conversely, these two conditions (pairwise independence and subgroup) are also sufficient for the construction (Appendix D). This observation has motivated our Theorem 1.1, even though the theorem is proved using techniques different from integrality gap construction.

Theorem 1.1 settles the approximability of boolean Max-k-CSP (up to a constant factor), by choosing C to be a Samorodnitsky–Trevisan hypergraph predicate (Appendix C.1).

Corollary 1.2. *For any $k \geqslant 3$, there is $\varepsilon = o_{n;k}(1)$ such that it is* NP*-hard to $(1-\varepsilon, 2k/2^k + \varepsilon)$-decide* Max-$k$-CSP *over boolean domain.*

Below are additional results that follow from our main theorem. Readers not interested in these results may go directly to Chapter 2.

[1] We thank Madhu Sudan for pointing out that Theorem 1.1 also covers [Håstad 2011, Theorem 4].

[2] The notation $\varepsilon = o_{n;k,|G|}(1)$ means that for any fixed k, any fixed $|G|$, the quantity ε goes to zero as n goes to infinity.

Problem	NP-**Hardness**	
Max-k-CSP	$1/2^{\Omega(k)}$	Note (1)
(over $\mathbb{Z}_2$)	$2^{O(\sqrt{k})}/2^k$	Note (2)
	$2k/2^k$	This work
Max-k-CSP	$q^{O(\sqrt{k})}/q^k$	Note (3)
(domain size q)	$O(q^2k/q^k)$	This work
	$O(qk/q^k) \quad (k \geqslant q)$	This work
2-Prover-1-Round-Game	$1/R^{\Omega(1)}$	Note (4)
(alphabet size R)	$4/R^{1/6}$	Note (5)
	$O(\log R)/\sqrt{R}$	This work
Independent-Set	$1/D^{\Omega(1)}$	Note (6)
(degree bound D)	$\exp(O(\sqrt{\log D}))/D$	Note (7)
	$O(\log D)^4/D$	This work
Almost-Coloring	$1/K^2$	Note (8)
(almost K-colorable)	$1/\exp(\Omega(\log K)^2)$	Note (9)
	$1/2^{K/2}$	This work

Table 1.1: Main NP-hardness results

(1) [Håstad 2001; Trevisan 1998; Sudan and Trevisan 1998; Khot et al. 2013]
(2) [Samorodnitsky and Trevisan 2000; Engebretsen and Holmerin 2008]
(3) [Engebretsen 2004]
(4) [Raz 1998; Holenstein 2009; Rao 2011]
(5) [Khot and Safra 2011]
(6) [Alon et al. 1995]
(7) [Trevisan 2001]
(8) [Dinur et al. 2010]
(9) [Khot and Saket 2012]

1.1 Query-efficient PCP

Another way to state Corollary 1.2 is a Probabilistically Checkable Proof (PCP) that is query-efficient, optimally. Put differently, this PCP has the largest gap between completeness and soundness, among all PCPs reading k bits from a proof. Query efficiency is measured by amortized query complexity, defined as $k/\log_2(c/s)$ when a PCP verifier read k bits from a proof, and has completeness c and soundness s [Bellare et al. 1998, Section 2.2.2].

Corollary 1.3. *For every $k \geqslant 3$, for some $\varepsilon = o_{n;k}(1)$, there is a PCP for n-variable* 3SAT *that reads k bits, uses randomness $(1+\varepsilon)k\log n$, has completeness $1-\varepsilon$, and has amortized query complexity $1 + (1 + o_k(1))(\log k)/k + \varepsilon$.*

Our amortized query complexity is tight up to the $o_k(1)$ term unless P = NP [Hast 2005a]. We also reduce the amortized free bit complexity of a PCP. A PCP has free bit complexity f if on every choice of randomness, there are at most 2^f accepting local views (out of the 2^k possibilities for the k bits read). Amortized free bit complexity is then $f/\log_2(c/s)$ [Bellare et al. 1998, Section 2.2.2]. Our PCP has amortized free bit complexity $(1+o_k(1))(\log k)/k$, up to additive $o_{n;k}(1)$. Our result also yields a new simple proof for the inapproximability of Max-Clique within $n^{1-\varepsilon}$, first shown by Håstad [1999] and simplified by Samorodnitsky and Trevisan [2000] and by Håstad and Wigderson [2003] (see also the derandomization by Zuckerman [2007]).

1.2 Independent-Set on bounded-degree graphs

The task of finding an independent set of maximum size in a graph of degree at most D was considered by Papadimitriou and Yannakakis [1991].

Theorem 1.4. *For all sufficiently large D, there is $\nu = o_{n;D}(1)$ such that it* NP*-hard to approximate* Independent-Set *on degree-D graphs beyond $O(\log D)^4/D + \nu$.*

The previous best NP-hardness ratio is $\exp(O(\sqrt{\log D}))/D$ by Trevisan [2001]. Our Theorem 1.4 is not far from factor $\Omega(\log D)/(D \log\log D)$ approximation algorithms of [Halperin 2002; Halldórsson 1998]. The best hardness ratio under the UG Conjecture is $O(\log D)^2/D$ by Austrin et al. [2011]. Theorem 1.4 also slightly improves the hardness of Induced-Matching on d-regular graphs and related problems in game theory [Chalermsook et al. 2013].

1.3 Almost-Coloring

Theorem 1.5. *For any $K \geqslant 3$, there is $\nu = o_{n;K}(1)$ such that given a graph with an induced K-colorable subgraph of fractional size $1-\nu$, it is* NP*-hard to find an independent set of fractional size $1/2^{K/2} + \nu$.*

The previous best NP-hardness result of Khot and Saket [2012] has soundness $\exp(-\Omega(\log K)^2)$. Almost-2-Coloring has arbitrarily small constant soundness under the UG Conjecture [Bansal and Khot 2009]. Given a K-colorable graph, Khot [2001] showed NP-hardness of finding an independent set of fractional size $\exp(-\Omega(\log K)^2)$ for sufficiently large K, and Huang [2013] has subsequently improved Khot's result to $\exp(-\Omega(K^{1/3}))$ using ideas in this paper. See [Khot and Saket 2012] for additional references on approximate coloring problems.

1.4 Non-boolean Max-k-CSP

We can choose C of Theorem 1.1 to be an O'Brien predicate of [Austrin and Mossel 2009, Theorem 1.2].

Corollary 1.6. *For any prime power q, any integer $k \geqslant 3$, there is $\varepsilon = o_{n;k,q}(1)$ such that it is* NP*-hard to* $(1-\varepsilon, q(q-1)k/q^k+\varepsilon)$*-decide* Max-$k$-CSP *over size-$q$ domain.*

The previous best NP-hardness result by Engebretsen [2004] has soundness $q^{O(\sqrt{k})}/q^k$. Like Austrin and Mossel [2009], the soundness in our Corollary 1.6 can be improved to $O(qk/q^k)+\varepsilon$ for infinitely many k. Alternatively, one can plug in Håstad predicates (Appendix C.2) to tighten the hardness ratio for every $k \geqslant q$.

Corollary 1.7. *For any integers $q \geqslant 2$ and $k \geqslant q$, it is* NP*-hard (under randomized reduction) to approximate* Max-k-CSP *over size-q domain beyond $O(qk/q^k)$.*

The randomized reduction can be replaced with a deterministic truth-table reduction, using k-wise δ-dependent distributions [Charikar et al. 2009, Section 3.4] (as pointed out to the author by Yury Makarychev). The best algorithm by Makarychev and Makarychev [2012] has a matching approximation ratio $\Omega(qk/q^k)$ when $k \geqslant \Omega(\log q)$.

1.5 Two-Prover-One-Round-Game

Theorem 1.8. *For any prime power q, there is $\varepsilon = o_{n;q}(1)$ such that it is* NP*-hard to* $(1-\varepsilon, O(\log q/q)+\varepsilon)$*-decide* 2-Prover-1-Round-Game *of alphabet size q^2.*

In terms of alphabet size $R = q^2$, the hardness ratio is $O(\log R/\sqrt{R})$. The previous best inapproximability result by Khot and Safra [2011] has soundness $4/R^{1/6}$ with alphabet size $R = q^6$. 2-Prover-1-Round-Game with perfect completeness have soundness $1/R^{\Omega(1)}$ [Raz 1998; Holenstein 2009; Rao 2011]. Hardness of 2-Prover-1-Round-Game is related to hardness of Quadratic-Programming [Arora et al. 2005], which was the original goal of Khot and Safra. Even though Theorem 1.8 improves soundness of the former problem, it does not imply any quasi-NP-hardness result for Quadratic-Programming, because our 2-Prover-1-Round-Game has a much worse soundness-size tradeoff.

Theorem 1.8 also has applications to many other optimization problems. Recently, Laekhanukit [2012] gave randomized reductions from 2-Prover-1-Round-Game to the following undirected network connectivity problems: Rooted k-Connectivity, Vertex-Connectivity Survivable Network Design, and Vertex-Connectivity k-Route Cut. His hardness results improve a number of previous ones, and our Theorem 1.8 further strengthens his results. See [Laekhanukit 2012] for details.

Problem	**Hardness**	
Rooted k-Connectivity	$k^{\Omega(1)}$	[Cheriyan et al. 2012]
	$\tilde{\Omega}(k^{1/18})$	[Laekhanukit 2012] + [Khot and Safra 2011]
	$\tilde{\Omega}(k^{1/10})$	[Laekhanukit 2012] + This work
Vertex-Connectivity Survivable Network Design	$k^{\Omega(1)}$	[Chakraborty et al. 2008]
	$\tilde{\Omega}(k^{1/16})$	[Laekhanukit 2012] + [Khot and Safra 2011]
	$\tilde{\Omega}(k^{1/8})$	[Laekhanukit 2012] + This work
Vertex-Connectivity k-Route Cut	$k^{\Omega(1)}$	[Chuzhoy et al. 2012]
	$\tilde{\Omega}(k^{1/14})$	[Laekhanukit 2012] + [Khot and Safra 2011]
	$\tilde{\Omega}(k^{1/6})$	[Laekhanukit 2012] + This work

Table 1.2: NP-hardness of connectivity problems

Chapter 2

Techniques

Despite progress on Unique-Games-based conditional results, unconditional NP-hardness of Max-k-CSP has lagged behind. This is due to limitations of existing proof composition techniques [Bellare et al. 1998, Section 3.4], which were known when dictator test was introduced.

To illustrate, consider Håstad's reduction from Label-Cover to Max-3-XOR. For our discussion, think of Label-Cover as a two-party game, where two parties try to convince a verifier that a Max-CSP instance L has a satisfying assignment A. The verifier randomly picks a clause $\boldsymbol{Q}$ from L and randomly a variable $\boldsymbol{u}$ from $\boldsymbol{Q}$. The verifier then asks for the satisfying assignment $A(\boldsymbol{Q})$ to the clause from one party and the assignment $A(\boldsymbol{u})$ to the variable from the other party. The verifier is convinced (and accepts) if $A(\boldsymbol{Q})$ and $A(\boldsymbol{u})$ agree at their assignment to $\boldsymbol{u}$.

When Label-Cover is reduced to Max-3-XOR, the above two-party game is transformed into a three-player game. The verifier now asks for a boolean reply from each player, and will accept or reject based on the XOR of the replies. Therefore the verifier will choose a subset $\boldsymbol{z}^{(1)}$ of assignments to $\boldsymbol{u}$ and ask the first player whether $A(\boldsymbol{u}) \in \boldsymbol{z}^{(1)}$. The verifier also chooses two subsets $\boldsymbol{z}^{(2)}, \boldsymbol{z}^{(3)}$ of satisfying assignments to $\boldsymbol{Q}$ and asks the other two players whether $A(\boldsymbol{Q}) \in \boldsymbol{z}^{(2)}$ and $A(\boldsymbol{Q}) \in \boldsymbol{z}^{(3)}$. The subsets $\boldsymbol{z}^{(1)}, \boldsymbol{z}^{(2)}, \boldsymbol{z}^{(3)}$ will be chosen carefully in a correlated way, and constitute a dictator test.

The above trasformation, known as composition of a dictator test with Label-Cover, naturally generalizes to more than three players. Note that each player belongs to one of the two parties. The above composition scheme is known not to yield optimal hardness for Max-k-CSP ([Bellare et al. 1998, Section 3.4] and [Sudan and Trevisan 1998]), because replies from the same party may conspire and appear correct, even if the Label-Cover instance has no good assignment. To get around the barrier, previous works focused on strengthening Label-Cover and adjusting the composition step (say by creating more parties), as well as improving the dictator test analysis. A sequence of works brought soundness down to $2^{O(\sqrt{k})}/2^k$, which is still far from optimal.

In this work, we leapfrog the barrier with a new approach. We view a Max-k-CSP instance as a k-player game, and reduce soundness by a technique we call *direct sum*, which is inspired by XOR-lemmas. Direct sum is like parallel repetition, aiming to reduce soundness by asking each player multiple questions at once. However, with direct sum each player gives only a single answer, namely the sum of answers to individual questions. Direct sum (or XOR-lemma) is invaluable to average-case complexity [Goldreich et al. 2011] and central to communication complexity [Barak et al. 2010; Sherstov 2012], but (to our knowledge) has never been used for amplifying gap in hardness of approximation. As it turns out, a natural formulation of a multiplayer XOR-lemma is false (see Remark 5.2), which may explain its absence in the inapproximability literature.

Unable to decrease soundness directly, we instead demonstrate *randomness* of replies. Randomness means lack of correlation. The crucial observation is that correlation never increases with direct sum (Lemma 5.3). It remains to show that, in the Soundness case of a single game, we can isolate any player of our choice, so that his/her reply becomes uncorrelated with the other $k-1$ replies after secret shifting (Theorem 5.4). Then the direct sum of k different games will isolate all players one by one, eliminating any correlation in their shifted replies.

We prove Theorem 5.4 using the canonical composition technique. In the soundness analysis of the dictator test, we invoke an invariance-style theorem (Theorem 7.2), based on O'Donnell and Wright [2012] and Wenner [2012]. We show invariance for the *correlation* (Definition 4.2) rather than the objective value.

Our approach also bypasses the composition barrier for other problems, with simple proofs. We improve the hardness of 2-Prover-1-Round-Game as an easy corollary (Chapter 9). Our low free-bit PCP also facilitates further reductions, improving hardness of Almost-Coloring (Chapter 8) and Independent-Set on bounded-degree graphs (Appendix B).

Previous reductions that bypassed the UG Conjecture for other problems [Khot 2002a; Guruswami et al. 2012; Feldman et al. 2009; Khot and Moshkovitz 2011] started from Khot's Smooth-Label-Cover [Khot 2002a]. By contrast, our reduction starts from the usual Label-Cover. In fact, the reduction in Theorem 1.1 maps a 3SAT instance on n variables to an Max-k-CSP instance of size $N = n^{k(1+o_{n;k,|G|}(1))}$. Assuming the Exponential Time Hypothesis [Impagliazzo et al. 2001] (that deciding 3SAT on n variables requires $\exp(\Omega(n))$ time), our Theorem 1.1 implies certain Max-k-CSP remain "approximation resistant" against $\exp(N^{(1-o(1))/k})$ time algorithms — a conclusion unlikely to follow from the UG Conjecture because Unique-Games have subexponential time algorithms [Arora et al. 2010].

Chapter 3

Preliminaries

As usual, let $[q] = \{1, \ldots, q\}$. Denote ℓ^p-norm of a vector $x \in \mathbb{R}^m$ by $\|x\|_{\ell^p} = (\sum_{i\in[m]} |x_i|^p)^{1/p}$.

Let $\triangle_\Sigma = \{x \in \mathbb{R}^\Sigma_{\geqslant 0} \mid \|x\|_{\ell^1} = 1\}$ denote the set of probability distributions over Σ. We also write $\triangle_q$ for $\triangle_{[q]}$.

Random variables are denoted by italic boldface letters, such as $\boldsymbol{x}$.

By the size of a constraint satisfaction problem (including $\mathsf{Label\text{-}Cover}$), we mean the number of constraints/hyperedges (disregarding weights).

We recall basic facts about characters. A character χ of a finite abelian group G is a homomorphism from G to the circle group $\mathbb{T}$ of complex numbers of modulus one (under multiplication). The constant 1 function, denoted $\mathbf{1}$, is always a character, known as the trivial character. Any character χ of a power group G^k has a unique decomposition as a product of characters $\chi_i : G \to \mathbb{T}$ in each coordinate, so that

$$\chi(a_1, \ldots, a_k) = \chi_1(a_1) \ldots \chi_k(a_k) \tag{3.1}$$

for any $(a_1, \ldots, a_k) \in G^k$.

Definition 3.1. Given $j \in [k]$, a character χ of G^k is *j-relevant* if its j-th component χ_j is non-trivial (i.e. not the constant 1 function).

Given two random variables $\boldsymbol{x}$ and $\boldsymbol{y}$ on a set Σ, their statistical distance $d(\boldsymbol{x}, \boldsymbol{y})$ is the statistical distance of their underlying distributions,

$$d(\boldsymbol{x}, \boldsymbol{y}) = \max_{A\subseteq\Sigma} |\mathbb{P}[\boldsymbol{x} \in A] - \mathbb{P}[\boldsymbol{y} \in A]|.$$

The following bound relating statistical distance and character distance is well known, see e.g. [Bogdanov and Viola 2010, Claim 33], who stated the result when G is a finite field, but whose proof can be easily adapted for general abelian groups.

Proposition 3.2. *If $|\mathbb{E}[\chi(\boldsymbol{x})] - \mathbb{E}[\chi(\boldsymbol{y})]| \leqslant \varepsilon$ for all characters χ, then $2d(\boldsymbol{x}, \boldsymbol{y}) \leqslant \sqrt{|G| - 1} \cdot \varepsilon$.*

Chapter 4

Max-CSP given by a predicate

We now define maximum constraint satisfaction problem $\mathsf{Max}\text{-}C$ given by a predicate C. Our definition departs from previous works in that the underlying constraint hypergraph of our instance is k-partite. This is because a $\mathsf{Max}\text{-}C$ instance represents a k-player game, and different players can give different replies on the same question.

Let G be an abelian group and C a subset of G^k. An instance $M = ((V_1, \ldots, V_k), \boldsymbol{Q})$ of $\mathsf{Max}\text{-}C$ is a distribution over constraints of the form $Q = (v, b)$, where $v = (v_1, \ldots, v_k) \in V_1 \times \cdots \times V_k$ is a k-tuple of variables and $b = (b_1, \ldots, b_k) \in G^k$ is a k-tuple of shifts. We think of an instance as a k-player game: a constraint is tuple of questions to the k players, and an assignment $f_i : V_i \to G$ is a strategy of player i. Naturally, upon receiving a variable v_i, player i responds with $f_i(v_i)$. A constraint $Q = (v, b)$ is satisfied if

$$f(v) - b \triangleq (f_1(v_1) - b_1, \ldots, f_k(v_k) - b_k) \in C.$$

The k players aim to satisfy the maximum fraction of constraints. The *value* of the game, denoted by $\mathrm{val}(M)$, is the maximum possible $\mathbb{P}[f(\boldsymbol{v}) - \boldsymbol{b} \in C]$ over k assignments $f_i : V_i \to G$. For boolean domain ($G = \mathbb{Z}_2$), the shifts specify whether the literals are positive or negative. Note that a game without shifts (equivalently, all shifts are the identity element 0_G) is trivial, since players have a perfect strategy by always answering 0_G. The shifts, unknown to the players, make the game challenging.

Definition 4.1. A subset C of G^k is balanced pairwise independent if for every two distinct coordinates $i \neq j \in [k]$ and every two elements $a, b \in G$,

$$\mathbb{P}[\boldsymbol{c}_i = a, \boldsymbol{c}_j = b] = 1/|G|^2,$$

where $\boldsymbol{c} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k)$ is a uniformly random element from C.

We will often choose C to be a subgroup of G^k. Examples of balanced pairwise independent subgroups are dual Hamming codes and Reed–Solomon codes of dimension at least two. Dual Hamming codes have been used to obtain inapproximability results based on the UG

Conjecture [Samorodnitsky and Trevisan 2009] or in the Lasserre hierarchy [Tulsiani 2009]. Reed–Solomon codes have appeared in low-degree tests.

Let $\mathcal{A}$ be the class of predicates over a balanced pairwise independent subgroup $C \subseteq G^k$ for some $k \geqslant 3$. In other words, these are the predicates satisfying the hypothesis of Theorem 1.1. These are also the predicates currently admitting a direct construction of integrality gaps for sum-of-squares programs (Appendix D). The class $\mathcal{A}$ is closely related to the bigger class $\mathcal{B}$ of predicates supporting a balanced pairwise independent distribution (possibly not subgroups), known to give approximation resistant CSPs under the UG Conjecture [Austrin and Mossel 2009] and in weaker SDP hierarchies [Benabbas et al. 2012; Tulsiani and Worah 2013]. Even though $\mathcal{A}$ is a proper subclass of $\mathcal{B}$ (personal communication with Madhur Tulsiani), many interesting predicates in $\mathcal{B}$ also belong to $\mathcal{A}$. In particular, the following predicates implicitly satisfy our abelian subgroup property: [Håstad 2001, Theorem 5.9], Samorodnitsky and Trevisan [2000; 2009], Engebretsen and Holmerin [2008], Guruswami and Raghavendra [2008], O'Brien predicates of [Austrin and Mossel 2009, Theorem 1.2], and Håstad (Appendix C.2). Not all approximation resistant CSPs satisfy our (or Austrin and Mossel's) condition; an notable exception is Guruswami et al. [1998] predicate (see [Hast 2005b, Theorem 7.1]).

When there is no prefect strategy, the shifted replies $f(\boldsymbol{v}) - \boldsymbol{b}$ may not have perfect correlation. We measure correlation of the best strategy by the following quantity.

Definition 4.2. Given Max-C instance M and character $\chi : G^k \to \mathbb{T}$, let

$$\|M\|_\chi \triangleq \max |\mathbb{E}\, \chi(f(\boldsymbol{v}) - \boldsymbol{b})| = \max |\mathbb{E}\, \chi(f_1(\boldsymbol{v}_1) - \boldsymbol{b}_1, \ldots, f_k(\boldsymbol{v}_k) - \boldsymbol{b}_k)|,$$

where the maximum is over k assignments $f_i : V_i \to G$.

Chapter 5

Direct sum

To make the game even more difficult for the players, we can take direct sum of instances. Recall that direct sum is a variant of parallel repetition, where each player receives ℓ questions at once. In direct sum, each player only gives a single answer, namely the sum of answers to the ℓ questions. We first define the direct sum of $\ell = 2$ games.

Definition 5.1. Let $M = ((V_1, \ldots, V_k), \boldsymbol{Q})$ and $M' = ((V_1', \ldots, V_k'), \boldsymbol{Q}')$ be $\textsf{Max-}C$ instances. Their direct sum $M \oplus M'$ is defined as $((V_1 \times V_1', \ldots, V_k \times V_k'), \boldsymbol{Q} \oplus \boldsymbol{Q}')$. Player i in $M \oplus M'$ receives a pair of variables $(v_i, v_i') \in V_i \times V_i'$ from M and M'.

The random question $\boldsymbol{Q} \oplus \boldsymbol{Q}'$ in $M \oplus M'$ is the direct sum of two independent random questions $\boldsymbol{Q}$ and $\boldsymbol{Q}'$, one from M and the other from M'. By the direct sum $Q \oplus Q'$ of two questions $Q = (v, b)$ and $Q' = (v', b')$, we mean sending every player i the variable pair $(v \oplus v')_i \triangleq (v_i, v_i')$ and receiving a reply $g_i(v_i, v_i')$. The shifts for $Q \oplus Q'$ is $b + b'$. To wit, $Q \oplus Q' = (v \oplus v', b + b')$.

We expect players' strategy to be independent across the two coordinates, that is $g_i(v_i, v_i') = (f_i \oplus f_i')(v_i, v_i') \triangleq f_i(v_i) + f_i'(v_i')$, where $f = (f_1, \ldots, f_k)$ is an assignment for M and $f' = (f_1', \ldots, f_k')$ an assignment for M'. However, players need not use such a strategy. Bounding the value of $M \oplus M'$ in terms of the values of M and M' is thus a daunting task.

Remark 5.2. Common sense suggests that by repeatedly taking direct sum, the ℓ-fold repeated game $M^{\oplus \ell} \triangleq M \oplus \ldots \oplus M$ will have no strategy better than a random one, as long as the original game M has no perfect strategy. More precisely, $\mathrm{val}(M^{\oplus \ell})$ should converge to the expected value of a random assignment as $\ell \to \infty$, provided $\|M\|_\chi < 1$ for all non-trivial characters χ (so that shifted replies are never contained in a proper subgroup of G^k). Such a result, if true, may be called a multiplayer XOR-lemma. This result is true for one- and two-player games, but turns out to be *false* for three-player games, as pointed out by Briët et al. [2013]. A counterexample to the three-player XOR-lemma, known as Mermin's game, has a perfect quantum strategy but no perfect classical strategy. Briët et al. observed that certain perfect quantum strategies of the repeated game can be "rounded" to a non-trivial classical strategy, via a multilinear Grothendieck-type inequality. Amazingly, the counterexample was discovered via *quantum* considerations, even though the setting is entirely *classical*.

Fortunately, we can bound the value of $M \oplus M'$ indirectly. As hinted earlier, we instead bound *correlation* of shifted replies. The following lemma shows that correlation can only decrease upon taking direct sum.

Lemma 5.3. *For any* Max-C *instances* M *and* M'*, any character* $\chi : G^k \to \mathbb{T}$*,*

$$\|M \oplus M'\|_\chi \leqslant \min\{\|M\|_\chi, \|M'\|_\chi\}.$$

Proof. Fix arbitrary assignments $f_i : V_i \times V_i' \to G$. The bias is

$$\left| \mathop{\mathbb{E}}_{\boldsymbol{Q}\boldsymbol{Q}'} \chi(f(\boldsymbol{v}, \boldsymbol{v}') - \boldsymbol{b} - \boldsymbol{b}') \right| \leqslant \mathop{\mathbb{E}}_{\boldsymbol{Q}} \left| \mathop{\mathbb{E}}_{\boldsymbol{Q}'} \chi(f(\boldsymbol{v}, \boldsymbol{v}') - \boldsymbol{b} - \boldsymbol{b}') \right|.$$

The RHS is at most $\|M'\|_\chi$, because after fixing a question $\boldsymbol{Q}$ to M, we get assignments $g_i^{\boldsymbol{Q}}(v_i') = f_i(\boldsymbol{v}_i, v_i') - \boldsymbol{b}_i$ to M'. Since f_i's are arbitrary, we have $\|M \oplus M'\|_\chi \leqslant \|M'\|_\chi$. The same argument also yields $\|M \oplus M'\|_\chi \leqslant \|M\|_\chi$. $\square$

Of course, a simple induction shows that $\|M_1 \oplus \ldots \oplus M_\ell\|_\chi \leqslant \min_{i \in [\ell]} \|M_i\|_\chi$.

The following theorem will be proved in Appendix A, based on a dictator test described in Chapter 6. See Definition 3.1 for j-relevant characters.

Theorem 5.4. *Let* C *be a balanced pairwise independent subset of* G^k*. There are* $\eta, \delta = o_{n;k,|G|}(1)$ *such that for any* $j \in [k]$*, it is NP-hard to decide the following cases given a* Max-C *instance* M_j*:*

1. *Completeness:* $\mathrm{val}(M_j) \geqslant 1 - \eta$.
2. *Soundness:* $\|M_j\|_\chi \leqslant \delta$ *for all* j*-relevant characters* $\chi : G^k \to \mathbb{T}$.

We can now prove Theorem 1.1. The reduction constructs k instances $M_1, \ldots, M_k$, one for each $j \in [k]$, as guaranteed by Theorem 5.4. The reduction then outputs the direct sum instance $M = M_1 \oplus \ldots \oplus M_k$. If each M_j has size at most m, then M has size at most m^k, which is polynomial in m for fixed k.

Proof of Theorem 1.1. **Completeness**. For every $j \in [k]$, let $f^{(j)} = (f_1^{(j)}, \ldots, f_k^{(j)})$ be an optimal assignment tuple for M_j. Consider the assignment tuple $g = (g_1, \ldots, g_k)$ for M that is independent across the k component games, that is

$$g_i(v_i^{(1)}, \ldots, v_i^{(k)}) = f_i^{(1)}(v_i^{(1)}) + \cdots + f_i^{(k)}(v_i^{(k)}),$$

Consider a question $\boldsymbol{R} = (\boldsymbol{u}, \boldsymbol{a}) = ((\boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(k)}), \boldsymbol{b}^{(1)} + \cdots + \boldsymbol{b}^{(k)})$ in M. If each of its component question $(\boldsymbol{v}^{(j)}, \boldsymbol{b}^{(j)})$ is satisfied by $f^{(j)}$, then

$$g(\boldsymbol{u}) - \boldsymbol{a} = \textstyle\sum_j f^{(j)}(\boldsymbol{v}^{(j)}) - \boldsymbol{b}^{(j)} \in C,$$

because C is closed under group operations. Hence g also satisfies $\boldsymbol{R}$. Therefore M has value at least $(1 - \eta)^k \geqslant 1 - k\eta$.

Soundness. Fix assignments $f_i : V_i \to G$. Let χ be a non-trivial character of G^k. Then χ is j-relevant for some $j \in [k]$, so

$$|\mathbb{E}\,\chi(f(\boldsymbol{v}) - \boldsymbol{b})| \leqslant \|M\|_\chi \leqslant \|M_j\|_\chi \leqslant \delta,$$

using Definition 4.2, Lemma 5.3, and Theorem 5.4. Let $\boldsymbol{a}$ be a uniformly random element from G^k, so $\mathbb{E}[\chi(\boldsymbol{a})] = 0$ for any non-trivial character χ. By Proposition 3.2, $f(\boldsymbol{v}) - \boldsymbol{b}$ and $\boldsymbol{a}$ have statistical distance

$$d(f(\boldsymbol{v}) - \boldsymbol{b}, \boldsymbol{a}) \leqslant \delta \cdot \sqrt{q^k}/2 =: \varepsilon.$$

Therefore

$$\mathbb{P}[f(\boldsymbol{v}) - \boldsymbol{b} \in C] \leqslant \mathbb{P}[\boldsymbol{a} \in C] + \varepsilon = |C|/|G|^k + \varepsilon. \qquad \square$$

Note that we prove something stronger than the statement of Theorem 1.1: In the Soundness case, the shifted replies are almost uniformly random. This *explains* the approximation resistance of $\mathsf{Max}\text{-}C$, and shows that C is useless in the sense of Austrin and Håstad [2012].

Chapter 6

Dictator test

Theorem 5.4 is based on a natural dictator test T, which we now describe. Throughout this chapter, C is a balanced pairwise independent subset of G^k.

6.1 Properties

We will compose a k-player dictator test with a Label-Cover instance, which is a game involving the clause party and the variable party. Before composition, the clause party replies over alphabet $[dR]$ and the variable party replies over alphabet $[R]$. Both alphabets are partitioned into R blocks, each of which has size 1 for the variable party and size d for the clause party. Define the t-th block

$$B(t) = \{s \in [dR] \mid (t-1)d < s \leqslant td\}$$

as the subset of clause party's alphabet associated with variable party's answer $t \in [R]$. After composition, the players replies over domain G. When assigning players to the parties, we single out player j as the *lonely* player, who is in the variable party, while all other players are in the clause party.

A k-player, j-lonely, d-blocked C-test T is a k-tuple of random variables $(\boldsymbol{z}^{(1)}, \dots, \boldsymbol{z}^{(k)}) \in G^{D_1} \times \cdots \times G^{D_k}$. Here dimension D_i is $D_i = dR$ for $i \neq j$, and $D_j = R$ for the lonely player $j \in [k]$. The test satisfies the completeness property: If players use strategies $f_i : G^{D_i} \to G$ that are "matching dictators" at the same block, the test accepts with high probability, say with probability $c \approx 1$. The test also satisfies the soundness property: If players use strategies far from matching dictators, then player j's reply should be uncorrelated other players' replies.

Formally, the completeness property says that if there are $t \in [R], s \in B(t)$ such that $f_i(z) = z_s$ for $i \neq j$ and $f_j(z) = z_t$, then

$$\mathbb{P}[(f_1(\boldsymbol{z}^{(1)}), \dots, f_k(\boldsymbol{z}^{(k)})) \in C] \geqslant c.$$

To state the soundness property, it is helpful to allow functions f_i to return a random element from G, by considering f_i as having codomain $\triangle_G$ that specifies the distribution of the random element. Functions are far from dictators if they have small influences, a quantity we now define.

Definition 6.1. Let Σ be a set (such as G) and H be a normed linear space (such as $\mathbb{R}^G$). Given $f : \Sigma^D \to H$, define $\|f\|_2^2 = \mathbb{E}_{\boldsymbol{x}\in\Sigma^D}[\|f(\boldsymbol{x})\|_H^2]$ and $\mathrm{Var}[f] = \|f - \mathbb{E}[f]\|_H^2$. The influence of a subset $B \subseteq [D]$ is the expected variance of f after randomly fixing coordinates outside of B, namely

$$\mathrm{Inf}_B[f] \triangleq \mathop{\mathbb{E}}_{\boldsymbol{x}_{\overline{B}}}\left[\mathop{\mathrm{Var}}_{\boldsymbol{x}_B}[f(\boldsymbol{x})]\right],$$

where $\overline{B} = [D] \setminus B$. We also write $\mathrm{Inf}_t[f]$ for $\mathrm{Inf}_{\{t\}}[f]$.

We measure correlation of players' replies f_i by the Fourier coefficients of $f(\boldsymbol{z})$.

Definition 6.2. For a character $\chi : G^k \to \mathbb{T}$, define

$$\mathrm{Bias}_{T,\chi}(f) \triangleq |\mathbb{E}\,\chi(f(\boldsymbol{z}))| = \left|\mathbb{E}\,\chi(f_1(\boldsymbol{z}^{(1)}), \ldots, f_k(\boldsymbol{z}^{(k)}))\right|.$$

Ideally, we want the soundness property that whenever functions $f_i : G^{D_i} \to \triangle_G$ have small common influences,

$$\max_{i\neq j}\left\{\sum_{t\in[R]} \mathrm{Inf}_t[f_j]\,\mathrm{Inf}_{B(t)}[f_i]\right\} \leqslant \tau,$$

then $\mathrm{Bias}_{T,\chi}(f) = o_{\tau;k,|G|}(1)$ for any j-relevant character χ of G^k.

The goal of this chapter is to construct a test T satisfying the completeness and soundness properties for a restricted class of functions.

6.2 Block distribution, noisy functions

The correlated random variables $\boldsymbol{z} = (\boldsymbol{z}^{(1)}, \ldots, \boldsymbol{z}^{(k)})$ in our test T will be independent across the R blocks. Each block is chosen from a block distribution μ over $G^{d_1} \times \cdots \times G^{d_k}$. Here dimension d_i is d for all $i \neq j$, and $d_j = 1$ for the lonely player $j \in [k]$. Therefore $\boldsymbol{z}$ is drawn from the product distribution $T = \mu^{\otimes R}$. We think of $\boldsymbol{z}$ as an $R \times k$ matrix where blocks are rows, and the i-th column is a string in $G^{d_i R}$. Entries in this matrix have different lengths: an entry in column j is an element from the base group G, while entries elsewhere are from the product group G^d.

The distribution μ will be the uniform distribution of choosing length-k tuples $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_d$ independently from C, conditioned on the tuples agreeing at position j. The tuples together represent an element in $G^{d_1} \times \cdots \times G^{d_k}$ because any position other than j gets a sequence of d elements from G, while position j gets the common element of the tuples.

Since C is balanced pairwise independent, the i-column of $\boldsymbol{z}$ is uniformly random over G^{D_i}. In fact, more is true: Looking only at column j and any other column $i \in [k]$ of a single block, the marginal distribution is uniform over $G \times G^d$. We call this property "pairwise independence at column j". This property is weaker than pairwise independence, because columns i and i' need not be independent. To verify this property, let $\boldsymbol{y}$ be a random block over $G^{d_1} \times \cdots \times G^{d_k}$ sampled according to μ. For any $a \in G$ and $b \in G^d$, the event "j-th column of $\boldsymbol{y}$ equals a and i-th column of $\boldsymbol{y}$ equals b" holds with probability

$$\mathbb{P}[\boldsymbol{y}_{(j)} = a, \boldsymbol{y}_{(i)} = b] = \mathbb{P}[\boldsymbol{y}_{(j)} = a] \cdot \mathbb{P}[\boldsymbol{y}_{(i),1} = b_1] \cdot \cdots \cdot \mathbb{P}[\boldsymbol{y}_{(i),d} = b_d],$$

where we have used pairwise independence of C and conditional independence in the definition of μ.

Our test is only sound against η-noisy functions.

Definition 6.3. Given a string $x \in G^m$, an *η-noisy copy* is a random string $\dot{\boldsymbol{x}} \in G^m$, so that independently for each $s \in [m]$, $\dot{\boldsymbol{x}}_s = x_s$ with probability $1-\eta$, and $\dot{\boldsymbol{x}}_s$ is set uniformly at random with probability η. For a function $f : G^m \to \triangle_G$, define the noise operator $\mathrm{T}_{1-\eta} f(x) = \mathbb{E}[f(\dot{\boldsymbol{x}})]$. A function g is η-noisy if $g = \mathrm{T}_{1-\eta} f$ for some function $f : G^m \to \triangle_G$.

When C is the collection of 3-bit strings of even parity, our dictator test becomes the Max-3-XOR test of [Håstad 2001, Section 5].

6.3 Soundness analysis

Inspired by O'Donnell and Wright [2012], we also consider an uncorrelated version of the test in our analysis.

Definition 6.4. The uncorrelated test $T' = (\mu')^{\otimes R}$ has block distribution μ', as defined below. A block from μ' is chosen exactly the same as in μ, and then the j-th entry is re-randomized to be a uniformly random element from G, independent of the other entries.

The following invariance-style theorem will be proved in Chapter 7. The theorem says that functions f_i's with small common influences cannot distinguish between the correlated test T from the uncorrelated version T'.

Theorem 6.5. *Let T be the test from Section 6.2, and T' be its uncorrelated version. Suppose $f_i : G^{D_i} \to \triangle_G$ are η-noisy functions satisfying*

$$\max_{i \neq j} \left\{ \sum_{t \in [R]} \mathrm{Inf}_t[f_j] \, \mathrm{Inf}_{B(t)}[f_i] \right\} \leqslant \tau.$$

Then for any character $\chi : G^k \to \mathbb{T}$,

$$\mathrm{Bias}_{T,\chi}(f) \leqslant \mathrm{Bias}_{T',\chi}(f) + \delta(|G|, k, \eta, \tau).$$

Here $\delta(q, k, \eta, \tau) \leqslant 4^k \,\mathrm{poly}(q/\eta)\sqrt{\tau}$.

We wish to show the term $\mathrm{Bias}_{T',\chi}(f)$ in Theorem 6.5 is negligible. This term is not small in general, if f_i are constant functions. To combat this, we apply the standard trick of folding.

Definition 6.6. Given a function $f : G^m \to G$, its folded version $\tilde{f} : G^m \to \triangle_G$ is the function which, upon receiving $x \in G^m$, picks a random $\boldsymbol{y} \in G$ and returns $f(x+(\boldsymbol{y}, \ldots, \boldsymbol{y}))-\boldsymbol{y}$.

The folding shift $\boldsymbol{y}$ is the same shift appearing in a constraint of $\mathsf{Max}\text{-}C$.

Consider applying the uncorrelated test T' to functions f_i's, where f_j is folded. For any j-relevant character χ,

$$\mathrm{Bias}_{T',\chi}(f) = |\mathbb{E}[\chi_j(f_j(\boldsymbol{z}^{(j)}))]\,\mathbb{E}[\chi_J(f_J(\boldsymbol{z}^{(J)}))]|,$$

where $J = [k] \setminus \{j\}$ denotes all players or columns other than j. The term $\mathbb{E}[\chi_j(f_j(\boldsymbol{z}^{(j)}))]$ is zero, because folding forces $f_j(\boldsymbol{z}^{(j)})$ to be uniformly random over G. Thus

$$\mathrm{Bias}_{T',\chi}(f) = 0.$$

Our preceding discussion implies the following bound on the bias of T for folded functions.

Theorem 6.7. *Let $\chi : G^k \to \mathbb{T}$ be a j-relevant character. Suppose η-noisy functions $f_i : G^{d_i R} \to \triangle_G$ satisfy*

$$\max_{i \neq j} \left\{ \sum_{t \in [R]} \mathrm{Inf}_t[f_j]\,\mathrm{Inf}_{B(t)}[f_i] \right\} \leqslant \tau.$$

Assume further f_j is folded. Then $\mathrm{Bias}_{T,\chi}(f) \leqslant \delta(|G|, k, \eta, \tau) \leqslant 4^k \operatorname{poly}(|G|/\eta)\sqrt{\tau}$.

The test T can be turned into an NP-hardness reduction by standard techniques (Appendix A).

Our dictator test can be generalized to the setting where $C \subseteq G^k$ is a subgroup that supports a balanced pairwise independent distribution. The proof of Theorem 6.5 goes through without change. Therefore our Theorem 1.1 still holds in this more general setting, bringing it closer to Austrin and Mossel [2009]. We choose to state the simpler version in this paper.

Chapter 7

Invariance-style theorem

In this chapter, we prove an invariance-style theorem for functions with small common influences. Our proof is based on [O'Donnell and Wright 2012, Section A] and [Wenner 2012, Theorem 3.21], who used ideas from [Mossel et al. 2010; Mossel 2010; O'Donnell and Wu 2009].

Invariance principle is a generalization of Berry–Esseen Central Limit Theorem. Let us informally recall the principle and the theorem. Berry–Esseen theorem says that a weighted sum of independent Rademacher variables $\{\boldsymbol{x}_i\}$ (weighted by $1/\sqrt{n}$) is close to a standard gaussian. Since a standard gaussian has the same distribution as a weighted sum of independent standard gaussian $\{\boldsymbol{g}_i\}$, we get

$$\frac{1}{\sqrt{n}}\sum_{i\in[n]}\boldsymbol{x}_i \approx \frac{1}{\sqrt{n}}\sum_{i\in[n]}\boldsymbol{g}_i.$$

Invariance principle generalizes this fact, and shows that a polynomial F of independent random variables $\{\boldsymbol{z}_i\}$ is close in distribution to the same polynomial of some other independent random variables $\{\boldsymbol{z}'_i\}$, that is,

$$F(\boldsymbol{z}_1,\ldots,\boldsymbol{z}_n) \approx F(\boldsymbol{z}'_1,\ldots,\boldsymbol{z}'_n),$$

under certain technical conditions. An important condition is that $\boldsymbol{z}_i$ and $\boldsymbol{z}'_i$ have identical first and second moments. Note that a Rademacher variable $\boldsymbol{x}_i$ and a standard gaussian $\boldsymbol{g}_i$ indeed agree in their first and second moments.

When we apply invariance principle in our setting, the random variable $\boldsymbol{z}_i$ is a block rather than a scalar. To demonstrate matching second moments, we need "pairwise independence at column j" (see Chapter 6). This is the intuition behind in an earlier version of our paper on ECCC. Below we give a shorter proof by incorporating ideas of Wenner [2012].[1]

[1] We thank an anonymous referee for suggesting that our proof may be simplified using Wenner's ideas.

7.1 Hoeffding decomposition

We will consider Hoeffding decomposition (or Efron–Stein decomposition) for functions f from Σ^m to a vector space H (such as $\mathbb{R}^q$). We need the following fact from [Mossel 2010, Definition 2.10].

Fact 7.1. *Every function $f : \Sigma^m \to H$ has a unique decomposition $f = \sum_{S \subseteq [m]} f^S$, where the functions $f^S : \Sigma^m \to H$ satisfy*

1. *f^S depends only on $x_S \triangleq \{x_i\}_{i \in S}$.*
2. *For any $T \not\supseteq S$ and any $x_T \in \Sigma^T$, $\mathbb{E}[f^S(\boldsymbol{x}) \mid \boldsymbol{x}_T = x_T] = 0$.*

As a result, we get an orthogonal decomposition whenever H is an inner product space, so that $\mathbb{E}_{\boldsymbol{x} \in \Sigma^m} \langle f^S(\boldsymbol{x}), f^T(\boldsymbol{x}) \rangle_H = 0$ for any $S \neq T$. Therefore $\|f\|_2^2 = \sum_{S \subseteq [m]} \|f^S\|_2^2$. Further, the influence on $B \subseteq [m]$ may be expressed as

$$\mathrm{Inf}_B[f] = \sum_{S : S \cap B \neq \emptyset} \|f^S\|_2^2.$$

A proof is essentially [Blais 2009, Appendix A.1]. As a result, the influence of η-noisy a function equals

$$\mathrm{Inf}_B[\mathrm{T}_{1-\eta} f] = \sum_{S : S \cap B \neq \emptyset} (1-\eta)^{2|S|} \|f^S\|_2^2.$$

This follows from the commutivity relation $(\mathrm{T}_{1-\eta} f)^S = \mathrm{T}_{1-\eta} f^S$ [Mossel 2010, Proposition 2.11], and the fact that $\mathrm{T}_{1-\eta} f^S = (1-\eta)^{|S|} f^S$ by property (2) of the decomposition.

7.2 Complex-valued functions

Theorem 6.5 is based on the following invariance-style theorem. In this version, the functions g_i take values in the closed unit disk $\mathbb{D}$ in the complex plane. In the statement, $\boldsymbol{z}$ has distribution $\mu^{\otimes R}$, where μ is a distribution over $\Sigma_1 \times \cdots \times \Sigma_k$ that is pairwise independent at column j (Section 6.2). Likewise $\boldsymbol{z}'$ has distribution $(\mu')^{\otimes R}$, where μ' is the uncorrelated version of μ (Definition 6.4).

Theorem 7.2. *Suppose $g_i : \Sigma_i^R \to \mathbb{D}$ are functions satisfying $\sum_{t \in [R]} \mathrm{Inf}_t[g_i] \leqslant A$ for all $i \in [k]$, and*

$$\max_{i \neq j} \left\{ \sum_{t \in [R]} \mathrm{Inf}_t[g_j] \, \mathrm{Inf}_t[g_i] \right\} \leqslant \tau.$$

Then

$$|\mathbb{E}[g(\boldsymbol{z})] - \mathbb{E}[g(\boldsymbol{z}')]| \leqslant 4^k \sqrt{A\tau},$$

where $g(z) = \prod_{i \in [k]} g_i(z^{(i)})$.

Proof. Similar to Lindeberg proof of Berry–Esseen theorem, we consider random variables that are hybrids of $\boldsymbol{z}$ and $\boldsymbol{z}'$. For $t = 0, \ldots, R$, the t-th hybrid is $\boldsymbol{z}_{(t)} = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_t, \boldsymbol{z}'_{t+1}, \ldots, \boldsymbol{z}'_R)$, where every $\boldsymbol{z}_s$ is distributed according to μ and every $\boldsymbol{z}'_s$ according to μ', independently. Recall that we think of each $\boldsymbol{z}_s$ or $\boldsymbol{z}'_s$ as a row of the matrix $\boldsymbol{z}_{(t)}$.

Consider the error for switching from $\boldsymbol{z}_{(t-1)}$ to $\boldsymbol{z}_{(t)}$

$$\mathrm{err}_t \triangleq \mathbb{E}[g(\boldsymbol{z}_{(t)})] - \mathbb{E}[g(\boldsymbol{z}_{(t-1)})].$$

Our goal is bounding $\sum_{t\in[R]} |\mathrm{err}_t|$.

Fix $t \in [R]$. Decompose each g_i as $(\mathrm{L}_t^{\|} + \mathrm{L}_t^{\perp}) g_i$ via the operators

$$\mathrm{L}_t^{\|} g_i = \sum_{S \ni t} g_i^S \qquad \text{and} \qquad \mathrm{L}_t^{\perp} g_i = \sum_{S \not\ni t} g_i^S.$$

Note that $\mathrm{L}_t^{\perp} g_i$ is independent of the row t, as guaranteed by Hoeffding decomposition Fact 7.1. We can rewrite err_t as

$$\mathrm{err}_t = \sum_{K \subseteq [k]} \left(\mathbb{E}[\mathrm{L}_t^K g(\boldsymbol{z}_{(t)})] - \mathbb{E}[\mathrm{L}_t^K g(\boldsymbol{z}_{(t-1)})] \right),$$

where

$$\mathrm{L}_t^K g(z) = \prod_{i\in[k]} \mathrm{L}_t^{i,K} g_i(z^{(i)}), \qquad \mathrm{L}^{i,K} = \begin{cases} \mathrm{L}_t^{\|} & \text{if } i \in K \\ \mathrm{L}_t^{\perp} & \text{if } i \notin K \end{cases}.$$

We bound the contribution to err_t for each K. Split K into $K_j = K \cap \{j\}$ and $K_J = K \setminus \{j\}$. We now show that the contribution is zero unless $|K_j| = 1$ and $|K_J| \geqslant 2$. If $|K_j| = 0$, the contribution is zero, because $\mathrm{L}_t^K g$ is independent of the entry (t, j), but $\boldsymbol{z}_{(t)}$ and $\boldsymbol{z}_{(t-1)}$ have identical joint marginal distributions everywhere else. If $|K_J| = 0$, the argument is similar, and now $\mathrm{L}_t^K g$ is independent of the entries (t, i) for all $i \neq j$.

What remains is $|K_j| = |K_J| = 1$. Suppose $K_J = \{h\}$. Then $\mathrm{L}_t^K g$ can only depend on two entries on row t, namely j and h. Since $\boldsymbol{z}_{(t)}$ and $\boldsymbol{z}_{(t-1)}$ have identical joint marginals on all rows except t, and they also have identical joint marginals at (t, j) and (t, h) (by pairwise independence at column j), the contribution is zero.

Let H denote the collection of all $K \subseteq [k]$ such that $|K_j| = 1$ and $|K_J| \geqslant 2$. Therefore we have shown

$$\mathrm{err}_t = \sum_{K \in H} \left(\mathbb{E}[\mathrm{L}_t^K g(\boldsymbol{z}_{(t)})] - \mathbb{E}[\mathrm{L}_t^K g(\boldsymbol{z}_{(t-1)})] \right).$$

Proposition 7.3. *For any hybrid $\boldsymbol{z}$, any $K \in H$, any distinct $h, \ell \in K_J$,*

$$\left| \mathbb{E}[\mathrm{L}_t^K g(\boldsymbol{z})] \right| \leqslant 2^{k-3} \sqrt{\mathrm{Inf}_t[g_j] \, \mathrm{Inf}_t[g_h] \, \mathrm{Inf}_t[g_\ell]}.$$

Assuming the proposition, we can bound

$$\sum_{t\in[R]} |\mathrm{err}_t| \leqslant 2^k \sum_{K\in H} \sum_{t\in[R]} \sqrt{\mathrm{Inf}_t[g_j]\,\mathrm{Inf}_t[g_h]\,\mathrm{Inf}_t[g_\ell]},$$

where $h = h_K, \ell = \ell_K$ are distinct elements in K_J. By Cauchy–Schwarz, the RHS is at most

$$2^k \sum_{K\in H} \sqrt{\sum_{t\in R} \mathrm{Inf}_t[g_j]\,\mathrm{Inf}_t[g_h]} \sqrt{\sum_{t\in[R]} \mathrm{Inf}_t[g_\ell]} \leqslant 2^{2k}\sqrt{A\tau},$$

proving our theorem.

It remains to prove Proposition 7.3. By Hölder's inequality,

$$\mathbb{E}[|\mathrm{L}_t^K g(\boldsymbol{z})|] \leqslant \left\| \mathrm{L}_t^{\|} g_j \mathrm{L}_t^{\|} g_h \right\|_2 \left\| \mathrm{L}_t^{\|} g_\ell \right\|_2 \prod_{i\neq j,h,\ell} \left\| \mathrm{L}_t^{i,K} g_i \right\|_\infty . \tag{7.1}$$

We analyse each factor on the RHS. By pairwise independence at column j,

$$\left\| \mathrm{L}_t^{\|} g_j \mathrm{L}_t^{\|} g_h \right\|_2 = \left\| \mathrm{L}_t^{\|} g_j \right\|_2 \left\| \mathrm{L}_t^{\|} g_h \right\|_2 .$$

Also,

$$\left\| \mathrm{L}_t^{\|} g_j \right\|_2 = \sqrt{\mathrm{Inf}_t[g_j]}.$$

Finally, to bound the sup-norms, note that

$$\mathrm{L}_t^{\perp} g_i(x) = \mathbb{E}\left[g_i(\boldsymbol{x}) \mid \boldsymbol{x}_{[R]\setminus t} = x_{[R]\setminus t} \right] \in \mathbb{D},$$

and likewise $\mathrm{L}_t^{\|} g_i(x) = g_i(x) - \mathrm{L}_t^{\perp} g_i(x) \in 2\mathbb{D}$. Therefore Eq. (7.1) implies Proposition 7.3. □

We remark that by slightly modifying the proof, the bound $4^k\sqrt{A\tau}$ in Theorem 7.2 can be improved to $k^2\sqrt{A\tau}$.

7.3 Simplex-valued functions

Recall the following bound on total influence for η-noisy functions. O'Donnell and Wright [2012] has a different definition of noisy influence, but their noisy influence is always bigger, so their upper bound still holds.

Fact 7.4. *([O'Donnell and Wright 2012, Fact A.3]) Let $A_\eta = \frac{2}{\eta}\ln(\frac{1}{\eta})$. Then for any $d, R \in \mathbb{N}$ and any $h : \Sigma^{dR} \to \mathbb{R}$,*

$$\sum_{t=1}^{R} \mathrm{Inf}_{B(t)}[\mathrm{T}_{1-\eta} h] \leqslant A_\eta \|h\|_2^2.$$

We now prove the invariance theorem for $\triangle_G$-valued functions, restated below.

Theorem 6.5. *Let T be the test from Section 6.2, and T' be its uncorrelated version. Suppose $f_i : G^{D_i} \to \triangle_G$ are η-noisy functions satisfying*

$$\max_{i \neq j} \left\{ \sum_{t \in [R]} \mathrm{Inf}_t[f_j] \, \mathrm{Inf}_{B(t)}[f_i] \right\} \leqslant \tau.$$

Then for any character $\chi : G^k \to \mathbb{T}$,

$$\mathrm{Bias}_{T,\chi}(f) \leqslant \mathrm{Bias}_{T',\chi}(f) + \delta(|G|, k, \eta, \tau).$$

Here $\delta(q, k, \eta, \tau) \leqslant 4^k \operatorname{poly}(q/\eta)\sqrt{\tau}$.

Proof. Apply Theorem 7.2 to the functions $g_i \triangleq \mathrm{X}_i(f_i) : G^{D_i} \to \mathbb{D}$, where $\mathrm{X}_i : \mathbb{R}^G \to \mathbb{C}$ is the linear map naturally derived from χ_i and satisfy

$$\mathrm{X}_i(e_a) = \chi_i(a) \qquad \forall a \in G.$$

To bound $\mathrm{Inf}_t[g_i]$, we will use

$$\mathrm{Inf}_B[\mathrm{X}_i(f_i)] \leqslant \|\mathrm{X}_i\|_{\mathrm{op}} \cdot \mathrm{Inf}_B[f_i], \tag{7.2}$$

where

$$\|\mathrm{X}_i\|_{\mathrm{op}} \triangleq \sup_{y \neq 0} \frac{|\mathrm{X}_i(y)|}{\|y\|_{\ell^2}} \leqslant |G|.$$

To prove Eq. (7.2), fix $x_{\overline{B}} \in G^{\overline{B}}$, and let $h(x_B) = f(x_B, x_{\overline{B}})$. We have

$$|\mathrm{X}_i(h_i) - \mathbb{E}[\mathrm{X}_i(h_i)]| = |\mathrm{X}_i(h_i - \mathbb{E}[h_i])| \leqslant \|\mathrm{X}_i\|_{\mathrm{op}} \|h_i - \mathbb{E}[h_i]\|_{\ell^2}.$$

Taking expectation over $x_{\overline{B}}$ and using Definition 6.1, the last inequality implies Eq. (7.2).

Eq. (7.2) implies

$$\mathrm{Inf}_t[g_i] \leqslant |G| \cdot \mathrm{Inf}_{B(t)}[f_i]$$

where we now interpret g_i as having domain Σ_i^R with $\Sigma_i = G^{d_i}$. Theorem 7.2 now implies Theorem 6.5, because the hypothesis of the former is justified by the hypothesis of the latter, together with the last inequality and Fact 7.4. $\square$

Chapter 8

Almost-Coloring

In this chapter, we prove Theorem 1.5. In our opinion, our proof is simpler than [Dinur et al. 2010; Khot and Saket 2012].

We construct a PCP with small covering parameter apart from small fraction of randomness. Our notion of covering parameter is a variant of Feige and Kilian's [1998]. We then turn the PCP into an FGLSS graph [Feige et al. 1996].

Let M be a $\mathsf{Max}\text{-}C$ instance. We say that M has *covering parameter* K if there are K assignments $f^{(1)}, \ldots, f^{(K)}$ covering every question $(\boldsymbol{v}, \boldsymbol{b})$ of M, that is for every $c \in C$, some $f^{(t)}$ satisfies $f^{(t)}(\boldsymbol{v}) - \boldsymbol{b} = c$.

Proposition 8.1. *Let C be a balanced pairwise independent subset of G^k. There is a $\mathsf{Max}\text{-}C$ instance M_C with covering parameter $|C|$.*

Proof. Let $K = |C|$. Enumerate tuples $c^{(1)}, \ldots, c^{(K)}$ in C. There is only one question $Q = (v, 0_{G^k})$ in M_C, where $v \in V_1 \times \cdots \times V_k$. Consider the assignments $f^{(t)} = (f_1^{(t)}, \ldots, f_k^{(t)})$, where $f_i^{(t)} : V_i \to G$ is given by $f_i^{(t)}(v_i) = c_i^{(t)}$. Then $f^{(t)}(v) = c^{(t)}$, and the K assignments $f^{(1)}, \ldots, f^{(K)}$ cover Q. $\square$

We recall the definition of an FGLSS graph, specialized for $\mathsf{Max}\text{-}C$.

Definition 8.2. Given an $\mathsf{Max}\text{-}C$ instance M, its FGLSS graph H has a vertex (Q, c) for every question $Q = (v, b)$ of M and every $c \in C$. A vertex (Q, c) represents an accepting configuration for M. The vertex has weight $w(Q, c) = \mathbb{P}[\boldsymbol{Q} = Q]/|C|$. Two vertices $((v, b), c)$ and $((v', b'), c')$ are connected if their corresponding configurations are conflicting, that is $v_i = v'_i$ and $b_i + c_i \neq b'_i + c'_i$ for some $i \in [k]$.

Denote by $\mathrm{val}(H)$ the maximum fractional size $w(S) \triangleq \sum_{u \in S} w(u)$ of an independent set S in H (a vertex subset S is an independent set if no edge in H has both endpoints in S).

The value of M determines the fractional size of a maximum independent set in H.

Proposition 8.3 ([Feige et al. 1996, Lemma 3.5])**.** $\mathrm{val}(M) = \mathrm{val}(H)/|C|$.

From now on, C will be a subgroup (not just a subset). Let M be the instance from Theorem 1.1, which either has value at least $1-\eta$ or at most $|C|/|G|^k+\varepsilon$. We construct a PCP M' which is the direct sum $M_C \oplus M$. The output instance is the FGLSS graph H for M'.

Proof of Theorem 1.5. **Completeness**. There are K assignments $g^{(1)},\ldots,g^{(K)}$ covering $1-\eta$ fraction of questions $(\boldsymbol{v},\boldsymbol{b})$ of M'. Indeed, we can take $g^{(t)} = f^{(t)} \oplus f$, where $f^{(t)}$ is a dictator assignment assignment from Proposition 8.1 and f is an assignment satisfying $1-\eta$ questions of M. Then for any question $\boldsymbol{Q} = (\boldsymbol{v},\boldsymbol{b})$ of M satisfied by f and any question $\boldsymbol{Q}_C$ of M_C, the question $\boldsymbol{Q}_C \oplus \boldsymbol{Q}$ is covered by the $g^{(t)}$'s, since the map $c \mapsto c+z$ is a permutation of C whenever $z = f(\boldsymbol{v}) - \boldsymbol{b} \in C$.

In the FGLSS graph H, the K assignments $g^{(t)}$'s correspond to K independent sets containing $1-\eta$ fraction of vertices in total.

Soundness. By the proof of Theorem 1.1, M' inherits the soundness property from M. By Proposition 8.3, no independent set in H has fractional size more than

$$\frac{1}{|C|}\left(\frac{|C|}{|G|^k}+\varepsilon\right) = \frac{1}{|G|^k} + \frac{\varepsilon}{|C|}.$$

To get the result, fix C to be a Samorodnitsky–Trevisan hypergraph predicate (Appendix C.1). Then $K \leqslant 2k$, so soundness is $1/2^k \leqslant 1/2^{K/2}$, up to additive $\varepsilon/|C|$. $\square$

Chapter 9

Two-Prover-One-Round-Game

We prove Theorem 1.8 in this chapter.

Let $M = ((V_1, \dots, V_k), \boldsymbol{Q})$ be an instance of $\mathsf{Max}\text{-}C$. We convert M into a two-prover-one-round game $L_M = ((U, W), \boldsymbol{P})$ between the clause player and the variable player. The variable player receives a variable $u \in U \triangleq V_1 \cup \dots \cup V_k$, and the clause player receives a clause $Q \in W \triangleq \operatorname{supp}(\boldsymbol{Q}) \subseteq (V_1 \times \dots \times V_k) \times G^k$. In the new game L_M, a clause $\boldsymbol{Q} = (\boldsymbol{v}, \boldsymbol{b})$ is chosen from M, and a variable $\boldsymbol{u}$ is chosen uniformly at random from $\boldsymbol{v} = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_k)$, so that $\boldsymbol{u} = \boldsymbol{v}_{\boldsymbol{j}}$ for a random index $\boldsymbol{j} \in [k]$. The clause player responds with a satisfying assignment $g(\boldsymbol{Q}) \in C$ to $\boldsymbol{Q}$; the variable player responds with an assignment $f(\boldsymbol{u}) \in G$ to $\boldsymbol{u}$. The players win if their replies agree,

$$g(\boldsymbol{Q})_{\boldsymbol{j}} = f(\boldsymbol{u}) - \boldsymbol{b}_{\boldsymbol{j}}.$$

Then L_M is a two-prover-one-round game of alphabet size $|C|$. This game (as well as other two-prover-one-round games mentioned in the Introduction) is a *projection* games, i.e., the reply of the first player determines the *only* correct reply of the second player.

Consider the instance L_M when M is the output instance of Theorem 1.1. It is straightforward to show that $\operatorname{val}(L_M) \geqslant 1 - \varepsilon$ if $\operatorname{val}(M) \geqslant 1 - \varepsilon$. For the Soundness case, we again consider randomness in variable player's reply. Define $h(v) = (f(v_1), \dots, f(v_k)) \in G^k$ for $v \in V_1 \times \dots \times V_k$.

Recall the multiplicative Chernoff bound (e.g. [Schmidt et al. 1995, Theorem 2(I)]).

Proposition 9.1. *Suppose $\boldsymbol{Y}$ is a sum of independent $\{0,1\}$-valued random variables. Let $\mu = \mathbb{E}[\boldsymbol{Y}]$. Then for any $\lambda \geqslant 1$,*

$$\mathbb{P}[\boldsymbol{Y} \geqslant (1+\lambda)\mu] \leqslant \exp(-\lambda\mu/3).$$

Proof of Theorem 1.8. **Soundness**. Let $q = |G|$. For a fixed question $Q = (v, b)$, the winning probability (over the random index $\boldsymbol{j}$) is precisely

$$\operatorname{agr}(g(Q), h(v) - b) \triangleq \mathbb{P}[g(Q)_{\boldsymbol{j}} = (h(v) - b)_{\boldsymbol{j}}].$$

We can approximate the random variable $h(\boldsymbol{v})-\boldsymbol{b}$ with a random variable $\boldsymbol{a}$ that is uniform over G^k. Then for any potential answer $c \in C \subseteq G^k$ of the clause player, the fractional agreement $\operatorname{agr}(c, \boldsymbol{a})$ is a random variable $\boldsymbol{Y}/k$, where $\boldsymbol{Y}$ is Binomial with parameters k and $1/q$. Write $t = O(\log(q|C|)) \cdot k/q$, and assume $k \geqslant q$. By multiplicative Chernoff bound (Proposition 9.1),

$$\mathbb{P}[\operatorname{agr}(c, \boldsymbol{a}) \geqslant t/k] = \mathbb{P}[\boldsymbol{Y} \geqslant t] \leqslant 1/(q|C|).$$

It follows by union bound that

$$\mathbb{P}[\exists c \in C,\ \operatorname{agr}(c, \boldsymbol{a}) \geqslant t/k] \leqslant 1/q.$$

Therefore $\operatorname{val}(L_M)$ is bounded by

$$\begin{aligned}\mathbb{E}[\operatorname{agr}(g(\boldsymbol{Q}), h(\boldsymbol{v}) - \boldsymbol{b})] &\leqslant t/k + \mathbb{P}[\exists c \in C,\ \operatorname{agr}(c, h(\boldsymbol{v}) - \boldsymbol{b}) \geqslant t/k] \\ &\leqslant O(\log(q|C|)/q) + 1/q + d(h(\boldsymbol{v}) - \boldsymbol{b}, \boldsymbol{a}).\end{aligned}$$

As in the proof of Theorem 1.1, the statistical distance $d(h(\boldsymbol{v}) - \boldsymbol{b}, \boldsymbol{a}) = o_{n;k,|G|}(1)$ and is negligible.

To bound the first term, we can choose $k = q$ and C to be Reed–Solomon code over $\mathbb{F}_q$ of dimension two, so that $|C| = q^2$. $\square$

Chapter 10

Open problems

Our PCP in Corollary 1.3 has optimal query complexity, but lacks perfect completeness. Getting optimal query complexity and perfect completeness is an interesting open problem. Our PCP has large blow-up in size due to the use of long code, while a previous query-efficient PCP has a smaller variant using the Hadamard code [Khot 2001]. Getting a small PCP with optimal query-efficiency is another natural problem (it requires something different from Hadamard code [Samorodnitsky and Trevisan 2009; Lovett 2008]).

Bibliography

Noga Alon, Uriel Feige, Avi Wigderson, and David Zuckerman. 1995. Derandomized Graph Products. *Computational Complexity* 5, 1 (1995), 60–75.

Sanjeev Arora, Boaz Barak, and David Steurer. 2010. Subexponential Algorithms for Unique Games and Related Problems. In *Symposium on Foundations of Computer Science (FOCS)*. IEEE, Washington, DC, USA, 563–572.

Sanjeev Arora, Eli Berger, Elad Hazan, Guy Kindler, and Muli Safra. 2005. On Non-Approximability for Quadratic Programs. In *Symposium on Foundations of Computer Science (FOCS)*. IEEE, Washington, DC, USA, 206–215.

Per Austrin and Johan Håstad. 2012. On the Usefulness of Predicates. In *Conference on Computational Complexity (CCC)*. IEEE, Washington, DC, USA, 53–63.

Per Austrin, Subhash Khot, and Muli Safra. 2011. Inapproximability of Vertex Cover and Independent Set in Bounded Degree Graphs. *Theory of Computing* 7, 1 (2011), 27–43.

Per Austrin and Elchanan Mossel. 2009. Approximation Resistant Predicates from Pairwise Independence. *Computational Complexity* 18, 2 (2009), 249–271.

Nikhil Bansal and Subhash Khot. 2009. Optimal Long Code Test with One Free Bit. In *Symposium on Foundations of Computer Science (FOCS)*. IEEE, Washington, DC, USA, 453–462.

Boaz Barak, Fernando Guadalupe dos Santos Lins Brandão, Aram Wettroth Harrow, Jonathan Kelner, David Steurer, and Yuan Zhou. 2012. Hypercontractivity, Sum-of-Squares Proofs, and their Applications. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 307–326.

Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. 2010. How to Compress Interactive Communication. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 67–76.

Mihir Bellare, Oded Goldreich, and Madhu Sudan. 1998. Free Bits, PCPs, and Nonapproximability—Towards Tight Results. *SIAM J. Comput.* 27, 3 (June 1998), 804–915.

Siavosh Benabbas, Konstantinos Georgiou, Avner Magen, and Madhur Tulsiani. 2012. SDP Gaps from Pairwise Independence. *Theory of Computing* 8, 1 (2012), 269–289.

Eric Blais. 2009. Testing Juntas Nearly Optimally. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 151–158.

Andrej Bogdanov and Emanuele Viola. 2010. Pseudorandom Bits for Polynomials. *SIAM J. Comput.* 39, 6 (Jan. 2010), 2464–2486.

Jop Briët, Harry Buhrman, Troy Lee, and Thomas Vidick. 2013. Multipartite entanglement in XOR games. *Quantum Information and Computation* 13, 3 & 4 (2013), 334–360.

Tanmoy Chakraborty, Julia Chuzhoy, and Sanjeev Khanna. 2008. Network Design for Vertex Connectivity. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 167–176.

Parinya Chalermsook, Bundit Laekhanukit, and Danupon Nanongkai. 2013. Graph Products Revisited: Tight Approximation Hardness of Induced Matching, Poset Dimension and More. In *Symposium on Discrete Algorithms (SODA)*. SIAM, Philadelphia, PA, USA, 1557–1576.

Siu On Chan and Michael Molloy. 2013. A Dichotomy Theorem for the Resolution Complexity of Random Constraint Satisfaction Problems. *SIAM J. Comput.* 42, 1 (2013), 27–60.

Moses Charikar, Konstantin Makarychev, and Yury Makarychev. 2009. Near-Optimal Algorithms for Maximum Constraint Satisfaction Problems. *ACM Transactions on Algorithms* 5, 3, Article 32 (July 2009), 14 pages.

Joseph Cheriyan, Bundit Laekhanukit, Guyslain Naves, and Adrian Vetta. 2012. Approximating Rooted Steiner Networks. In *Symposium on Discrete Algorithms (SODA)*. SIAM, Philadelphia, PA, USA, 1499–1511.

Julia Chuzhoy, Yury Makarychev, Aravindan Vijayaraghavan, and Yuan Zhou. 2012. Approximation Algorithms and Hardness of the k-Route Cut Problem. In *Symposium on Discrete Algorithms (SODA)*. SIAM, Philadelphia, PA, USA, 780–799.

Irit Dinur and Prahladh Harsha. 2010. Composition of Low-Error 2-Query PCPs Using Decodable PCPs. *Property Testing* (2010), 280–288.

Irit Dinur, Subhash Khot, Will Perkins, and Muli Safra. 2010. Hardness of Finding Independent Sets in Almost 3-Colorable Graphs. In *Symposium on Foundations of Computer Science (FOCS)*. IEEE, Washington, DC, USA, 212–221.

Irit Dinur and Shmuel Safra. 2005. On the Hardness of Approximating Minimum Vertex Cover. *Annals of Mathematics* 162, 1 (2005), 439–485.

Lars Engebretsen. 2004. The Nonapproximability of Non-Boolean Predicates. *SIAM Journal on Discrete Mathematics* 18, 1 (Jan. 2004), 114–129.

Lars Engebretsen and Jonas Holmerin. 2008. More Efficient Queries in PCPs for NP and Improved Approximation Hardness of Maximum CSP. *Random Structures and Algorithms* 33, 4 (Dec. 2008), 497–514.

Uriel Feige, Shafi Goldwasser, Laszlo Lovász, Shmuel Safra, and Mario Szegedy. 1996. Interactive proofs and the hardness of approximating cliques. *J. ACM* 43, 2 (March 1996), 268–292.

Uriel Feige and Joe Kilian. 1998. Zero Knowledge and the Chromatic Number. *J. Comput. System Sci.* 57, 2 (Oct. 1998), 187–199.

Uriel Feige and Eran Ofek. 2006. Random 3CNF formulas elude the Lovász theta function. *arXiv CoRR* abs/cs/0603084 (2006).

Vitaly Feldman, Venkatesan Guruswami, Prasad Raghavendra, and Yi Wu. 2009. Agnostic Learning of Monomials by Halfspaces Is Hard. In *Symposium on Foundations of Computer Science (FOCS)*. IEEE, Washington, DC, USA, 385–394.

Oded Goldreich, Noam Nisan, and Avi Wigderson. 2011. On Yao's XOR-Lemma. In *Studies in Complexity and Cryptography*. Lecture Notes in Computer Science, Vol. 6650. Springer, 273–301.

Dima Grigoriev. 2001. Linear Lower Bound on Degrees of Positivstellensatz Calculus Proofs for the Parity. *Theoretical Computer Science* 259, 1-2 (May 2001), 613–622.

Venkatesan Guruswami, Daniel Lewin, Madhu Sudan, and Luca Trevisan. 1998. A Tight Characterization of NP with 3 Query PCPs. In *Symposium on Foundations of Computer Science (FOCS)*. IEEE, Washington, DC, USA, 8–17.

Venkatesan Guruswami and Prasad Raghavendra. 2008. Constraint Satisfaction over a Non-Boolean Domain: Approximation Algorithms and Unique-Games Hardness. In *International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*. Springer, Berlin, Germany, 77–90.

Venkatesan Guruswami, Prasad Raghavendra, Rishi Saket, and Yi Wu. 2012. Bypassing UGC from some Optimal Geometric Inapproximability Results. In *Symposium on Discrete Algorithms (SODA)*. SIAM, Philadelphia, PA, USA, 699–717.

Magnús Már Halldórsson. 1998. Approximations of Independent Sets in Graphs. In *International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*. Springer, Berlin, Germany, 1–13.

Eran Halperin. 2002. Improved Approximation Algorithms for the Vertex Cover Problem in Graphs and Hypergraphs. *SIAM J. Comput.* 31, 5 (May 2002), 1608–1623.

Gustav Hast. 2005a. Approximating Max kCSP — Outperforming a Random Assignment with Almost a Linear Factor. In *International Colloquium Conference on Automata, Languages and Programming (ICALP)*. 956–968.

Gustav Hast. 2005b. *Beating a Random Assignment*. Ph.D. Dissertation. KTH, Stockholm.

Johan Håstad. 1999. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica* 182, 1 (March 1999), 105–142.

Johan Håstad. 2001. Some Optimal Inapproximability Results. *J. ACM* 48, 4 (July 2001), 798–859.

Johan Håstad. 2008. Every 2-CSP Allows Nontrivial Approximation. *Computational Complexity* 17, 4 (2008), 549–566.

Johan Håstad. 2009. On the Approximation Resistance of a Random Predicate. *Computational Complexity* 18, 3 (Oct. 2009), 413–434.

Johan Håstad. 2011. Satisfying Degree-d Equations over $\mathrm{GF}[2]^n$. In *International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*. Springer-Verlag, Berlin, Heidelberg, 242–253.

Johan Håstad and Avi Wigderson. 2003. Simple Analysis of Graph Tests for Linearity and PCP. *Random Structures and Algorithms* 22, 2 (March 2003), 139–160.

Edwin Hewitt and Kenneth Allen Ross. 1994. *Abstract Harmonic Analysis: Volume 1: Structure of Topological Groups. Integration Theory. Group Representations* (2nd ed.). Springer.

Thomas Holenstein. 2009. Parallel Repetition: Simplification and the No-Signaling Case. *Theory of Computing* 5, 1 (2009), 141–172.

Sangxia Huang. 2013. Improved Hardness of Approximating Chromatic Number. *arXiv CoRR* abs/1301.5216 (2013).

Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. 2001. Which Problems Have Strongly Exponential Complexity? *J. Comput. System Sci.* 63, 4 (2001), 512–530.

Subhash Khot. 2001. Improved Inapproximability Results for MaxClique, Chromatic Number and Approximate Graph Coloring. In *Symposium on Foundations of Computer Science (FOCS)*. IEEE, Washington, DC, USA, 600–609.

Subhash Khot. 2002a. Hardness Results for Coloring 3-Colorable 3-Uniform Hypergraphs. In *Symposium on Foundations of Computer Science (FOCS)*. IEEE, Washington, DC, USA, 23–32.

Subhash Khot. 2002b. On the Power of Unique 2-Prover 1-Round Games. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 767–775.

Subhash Khot and Dana Moshkovitz. 2011. NP-hardness of Approximately Solving Linear Equations Over Reals. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 413–420.

Subhash Khot and Muli Safra. 2011. A Two Prover One Round Game with Strong Soundness. In *Symposium on Foundations of Computer Science (FOCS)*. IEEE, 648–657.

Subhash Khot, Muli Safra, and Madhur Tulsiani. 2013. Towards an Optimal Query Efficient PCP?. In *Innovations in Theoretical Computer Science (ITCS)*. ACM, New York, NY, USA, 173–186.

Subhash Khot and Rishi Saket. 2012. Hardness of Finding Independent Sets in Almost q-Colorable Graphs. In *Symposium on Foundations of Computer Science (FOCS)*. IEEE, Washington, DC, USA, 380–389.

Bundit Laekhanukit. 2012. Parameters of Two-Prover-One-Round Game and the Hardness of Connectivity Problems. (2012). arXiv:1212.0752.

Jean Bernard Lasserre. 2001. Global Optimization with Polynomials and the Problem of Moments. *SIAM Journal on Optimization* 11, 3 (2001), 796–817.

Shachar Lovett. 2008. Lower bounds for adaptive linearity tests. In *Symposium on Theoretical Aspects of Computer Science (STACS)*. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, 515–526.

Konstantin Makarychev and Yury Makarychev. 2012. Approximation Algorithm for Non-boolean MAX k-CSP. In *International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*. Springer, Berlin, Germany, 254–265.

Dana Moshkovitz and Ran Raz. 2010. Two-Query PCP with Subconstant Error. *J. ACM* 57, 5, Article 29 (June 2010), 29 pages.

Elchanan Mossel. 2010. Gaussian Bounds for Noise Correlation of Functions. *Geometric and Functional Analysis* 19 (2010), 1713–1756.

Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. 2010. Noise Stability of Functions with Low Influences: Invariance and Optimality. *Annals of Mathematics* 171, 1 (2010).

Ryan O'Donnell and John Wright. 2012. A New Point of NP-Hardness for Unique Games. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 289–306.

Ryan O'Donnell and Yi Wu. 2009. Conditional Hardness for Satisfiable 3-CSPs. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 493–502.

Ryan O'Donnell and Yuan Zhou. 2013. Approximability and Proof Complexity. In *Symposium on Discrete Algorithms (SODA)*. SIAM, Philadelphia, PA, USA, 1537–1556.

Christos Harilaos Papadimitriou and Mihalis Yannakakis. 1991. Optimization, Approximation, and Complexity Classes. *J. Comput. System Sci.* 43, 3 (1991), 425–440. Extended abstract in STOC 1988.

Pablo Parrilo. 2000. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*. Ph.D. Dissertation. California Institute of Technology.

Prasad Raghavendra. 2008. Optimal Algorithms and Inapproximability Results for Every CSP?. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 245–254.

Anup Rao. 2011. Parallel Repetition in Projection Games and a Concentration Bound. *SIAM J. Comput.* 40, 6 (Dec. 2011), 1871–1891.

Ran Raz. 1998. A Parallel Repetition Theorem. *SIAM J. Comput.* 27, 3 (June 1998), 763–803.

Alex Samorodnitsky and Luca Trevisan. 2000. A PCP Characterization of NP with Optimal Amortized Query Complexity. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 191–199.

Alex Samorodnitsky and Luca Trevisan. 2009. Gowers Uniformity, Influence of Variables, and PCPs. *SIAM J. Comput.* 39, 1 (2009), 323–360.

Jeanette P. Schmidt, Alan Siegel, and Aravind Srinivasan. 1995. Chernoff–Hoeffding Bounds for Applications with Limited Independence. *SIAM Journal on Discrete Mathematics* 8, 2 (1995), 223–250.

Grant Schoenebeck. 2008. Linear Level Lasserre Lower Bounds for Certain k-CSPs. In *Symposium on Foundations of Computer Science (FOCS)*. IEEE, Washington, DC, USA, 593–602. Newer version available at the author's homepage.

Alexander Alexandrovich Sherstov. 2012. The Multiparty Communication Complexity of Set Disjointness. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 525–548.

Madhu Sudan and Luca Trevisan. 1998. Probabilistically Checkable Proofs with Low Amortized Query Complexity. In *Symposium on Foundations of Computer Science (FOCS)*. IEEE, Washington, DC, USA, 18–27.

Luca Trevisan. 1998. Recycling Queries in PCPs and in Linearity Tests. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 299–308.

Luca Trevisan. 2001. Non-Approximability Results for Optimization Problems on Bounded Degree Instances. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 453–461.

Madhur Tulsiani. 2009. CSP Gaps and Reductions in the Lasserre Hierarchy. In *Symposium on Theory of Computing (STOC)*. ACM, New York, NY, USA, 303–312.

Madhur Tulsiani and Pratik Worah. 2013. LS+ Lower Bounds from Pairwise Independence. In *Conference on Computational Complexity (CCC)*. IEEE, Washington, DC, USA. To appear.

Cenny Wenner. 2012. Circumventing d-to-1 for Approximation Resistance of Satisfiable Predicates Strictly Containing Parity of Width Four. In *International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*. Springer Berlin Heidelberg, 325–337.

David Zuckerman. 2007. Linear Degree Extractors and the Inapproximability of Max Clique and Chromatic Number. *Theory of Computing* 3, 1 (2007), 103–128.

Uri Zwick. 1998. Approximation Algorithms for Constraint Satisfaction Problems Involving at Most Three Variables per Constraint. In *Symposium on Discrete Algorithms (SODA)*. SIAM, Philadelphia, PA, USA, 201–210.

Appendix A

Composition

In this chapter, we prove Theorem 5.4. Our reduction closely follows those in previous works [Håstad 2001; O'Donnell and Wright 2012], with one notable difference to Håstad's reduction: we allow different strategies from different players, so our output instance is k-partite. We will need this feature for the direct sum operation.

As usual, we will reduce from Label-Cover $\mathsf{LC}_{R,dR}$. An instance of $\mathsf{LC}_{R,dR}$ is a weighted bipartite graph $((U,V),\boldsymbol{e})$. Vertices from U are variables with domain $[R]$, and vertices from V are variables with domain $[dR]$. Every edge $\boldsymbol{e} = (\boldsymbol{u},\boldsymbol{v}) \in U \times V$ has an associated d-to-1 map $\pi_{\boldsymbol{e}} : [dR] \to [R]$. Given an assignment $A : U \to [R], V \to [dR]$, the constraint on $\boldsymbol{e}$ is satisfied if $\pi_{\boldsymbol{e}}(A(\boldsymbol{v})) = A(\boldsymbol{u})$.

The following theorem of Moshkovitz and Raz [2010] asserts hardness of Label-Cover (see also Dinur and Harsha [2010]).

Theorem A.1. *For some $0 < c < 1$ and some $g(n) = \Omega(\log n)^c$, for any $\sigma = \sigma(n) \geqslant \exp(-g(n))$, there are $d, R \leqslant \exp(\mathrm{poly}(1/\sigma))$ such that the problem of deciding a 3-SAT instance with n variables can be Karp-reduced in $\mathrm{poly}(n)$ time to the problem of $(1,\sigma)$-deciding a $\mathsf{LC}_{R,dR}$ instance L of size $n^{1+o(1)}$. Furthermore, L is a bi-regular bipartite graph with left- and right-degrees $\mathrm{poly}(1/\sigma)$.*

Our reduction from Label-Cover to Max-C produces an instance that is a k-uniform, k-partite hypergraph on the vertex set $V_1 \cup \cdots \cup V_k$. The j-th vertex set V_j is $U \times G^R$, obtained by replacing each vertex in U with a G-ary hypercube. Any other vertex set V_i is a copy of $V \times G^{dR}$. All vertices are variables with domain G (that has q elements). We think of an assignment to variables in $u \in V_j$ as a function $f_{j,u} : G^R \to G$, and likewise an assignment to variables in $v \in V_i$ as a function $f_{i,v} : G^{dR} \to G$.

For every constraint $\boldsymbol{e} = (\boldsymbol{u},\boldsymbol{v})$, the reduction introduces C-constraints on the (folded versions of) η-noisy assignments $f_{j,\boldsymbol{u}}$ and $f_{i,\boldsymbol{v}}$, as specified by a dictator test T under blocking map $\pi_{\boldsymbol{e}}$.

The following theorem, together with Theorem A.1, implies Theorem 5.4.

Theorem A.2. *Let T be the test from Chapter 6. Suppose $\sigma \leqslant \delta\eta\tau^2/(k-1)$, where $\tau = \tau(q,k,\eta,\delta) = \mathrm{poly}(\eta\delta/q)/16^k$ is chosen to satisfy $\delta \leqslant 4^k \mathrm{poly}(q/\eta)\sqrt{\tau}$ in Theorem 6.7.*

The problem of $(1,\sigma)$-deciding a $\mathsf{LC}_{R,dR}$ instance L can be Karp-reduced to the problem of deciding the following cases given a Max-C instance M_j:

1. *Completeness:* $\mathrm{val}(M_j) \geqslant 1-\eta$.
2. *Soundness:* $\|M_j\|_\chi \leqslant 2\delta$ *for all j-relevant characters χ.*

Further, if L has size m, then M_j has size $m \cdot q^{O(kdR)}$.

Proof. **Completeness**. Let A be an assignment to the $\mathsf{Label\text{-}Cover}$ instance with value 1. Consider the assignment $f_{j,u}(z) = z_{A(u)}$ and $f_{i,v}(z) = z_{A(v)}$. These are matching dictators since A satisfies the constraint on $\boldsymbol{e}$. Therefore for every $\boldsymbol{e}$, at least $1-k\eta$ fraction of the associated C-constraints from T are satisfied by $f_{j,\boldsymbol{u}}$ and $f_{i,\boldsymbol{v}}$'s.

Soundness. We prove the contrapositive. Let $\chi : G^k \to \mathbb{T}$ be a j-relevant character. Suppose there are folded assignments $f_{i,v} : G^{d_i R} \to \triangle_q$ for M_j causing the bias to exceed 2δ. Then

$$\|M\|_\chi = \left| \mathop{\mathbb{E}}_{\boldsymbol{e}} \mathop{\mathbb{E}}_{\boldsymbol{z}} \chi(f_{\boldsymbol{e}}(\boldsymbol{z})) \right| \leqslant \mathop{\mathbb{E}}_{\boldsymbol{e}} \left| \mathop{\mathbb{E}}_{\boldsymbol{z}} \chi(f_{\boldsymbol{e}}(\boldsymbol{z})) \right|,$$

where $f_{\boldsymbol{e}} = (f_{1,\boldsymbol{w}_1}, \ldots, f_{k,\boldsymbol{w}_k})$ with $\boldsymbol{w}_i = \boldsymbol{v}$ for $i \neq j$ and $\boldsymbol{w}_j = \boldsymbol{u}$. The RHS is at most

$$\mathop{\mathbb{E}}_{\boldsymbol{e}} \mathrm{Bias}_{T,\chi}(f_{\boldsymbol{e}}).$$

Therefore at least δ fraction of the edges $\boldsymbol{e}$ satisfy $\mathrm{Bias}_{T,\chi}(f_{\boldsymbol{e}}) > \delta$. We call such edges good.

For any good edge e, some $i_e \neq j$ satisfies

$$\sum_{t\in[R]} \mathrm{Inf}_t[f_{j,u}]\, \mathrm{Inf}_{\pi_e^{-1}(t)}[f_{i_e,v}] \geqslant \tau \tag{A.1}$$

by Theorem 6.7.

We use the following randomized decoding procedure to generate an assignment $\boldsymbol{A}$ for the LC instance. Since $f_{i,u}$ is η-noisy, $f_{i,u} = \mathrm{T}_{1-\eta} h_{i,u}$ for some $h_{i,u}$. For every $u \in U$, choose $S \subseteq [R]$ with probability $\|h^S_{j,u}\|_2^2$. (These numbers sum to at most 1 by the discussion following Fact 7.1. For the remaining probability, pick S arbitrarily.) Then pick $\boldsymbol{A}(u)$ as a uniformly random element in S (or assign arbitrarily if $S = \emptyset$). To get a label $\boldsymbol{A}(v)$, we first pick a random position $\boldsymbol{i} \in [k]$ different from j, then go on as before using $\|h^S_{\boldsymbol{i},v}\|_2^2$ as the probability distribution.

Then for any $B \subseteq [R]$ and any $u \in U$,

$$\begin{aligned}
\mathbb{P}[\boldsymbol{A}(u) \in B] &\geqslant \sum_{S : S \cap B \neq \emptyset} \|h_{j,u}^S\|_2^2 \cdot |S \cap B|/|S| \\
&\geqslant \sum_{S : S \cap B \neq \emptyset} \|h_{j,u}^S\|_2^2 \cdot \eta(1-\eta)^{|S|/|S \cap B|} \\
&\qquad \text{(since } \alpha \geqslant \eta(1-\eta)^{1/\alpha} \text{ for } \alpha > 0 \text{ and } 0 \leqslant \eta \leqslant 1\text{)} \\
&\geqslant \eta \cdot \mathrm{Inf}_B[f_{j,u}].
\end{aligned}$$

And similarly

$$\mathbb{P}[\boldsymbol{A}(v) \in B] \geqslant \eta \cdot \mathop{\mathbb{E}}_{\boldsymbol{i} \neq j} \mathrm{Inf}_B[f_{\boldsymbol{i},v}].$$

For a good edge e,

$$\begin{aligned}
\mathbb{P}[\boldsymbol{A}(u) = \pi_e(\boldsymbol{A}(v))] &= \sum_{t \in [R]} \mathbb{P}[\boldsymbol{A}(u) = t \text{ and } \boldsymbol{A}(v) \in \pi_e^{-1}(t)] \\
&= \sum_{t \in [R]} \mathbb{P}[\boldsymbol{A}(u) = t]\, \mathbb{P}[\boldsymbol{A}(v) \in \pi_e^{-1}(t)] \\
&\geqslant \frac{\eta^2}{k-1} \sum_{t \in [R]} \mathrm{Inf}_t[f_{j,u}]\, \mathrm{Inf}_{\pi_e^{-1}(t)}[f_{i_e,v}] \geqslant \frac{\eta^2 \tau}{k-1}.
\end{aligned}$$

Therefore the expected fraction of constraints in L satisfied by $\boldsymbol{A}$ exceeds $\delta\eta^2\tau/(k-1) \geqslant \sigma$. $\square$

Appendix B

Independent-Set

We prove Theorem 1.4 in this chapter. In the Independent-Set problem, a graph H is given, and the goal is to find the largest independent set in H. The application of low free-bit PCP to Independent-Set is well known [Samorodnitsky and Trevisan 2009], but the actual hardness ratio is not explicitly computed before, so we include a proof for completeness.

Our proof closely follows Trevisan's [2001, Section 6]. We will construct an FGLSS graph H (Definition 8.2) for our PCP, and reduce degree by replacing bipartite complete subgraphs in H with "bipartite δ-dispersers" (close relatives of bipartite expanders). The degree bound $O(\delta^{-1}\log(\delta^{-1}))$ for dispersers determines the hardness ratio. Unlike Trevisan [2001], we do not use efficient deterministic constructions of dispersers, since none of the known constructions matches the degree bound offered by probabilistic ones. Luckily, bipartite complete subgraphs in H have size bounded by a function of $1/\varepsilon$ and $1/\eta$, so we can find good dispersers by exhaustive search.

Proof of Theorem 1.4. By Corollary 1.2, there is a PCP Π with completeness $c = 1-\eta$, soundness $s = 2k/2^k + \varepsilon$, and free bit complexity at most $\log_2(2k)$. Construct the FGLSS graph H for Π.

Following [Dinur and Safra 2005, Proposition 8.1], we now turn H into an unweighted graph H' (equivalently, vertices in H' have equal weight), by duplicating vertices. Suppose H is a weighted independent set instance of size m with minimum weight λ and maximum weight κ, and $0 < \sigma \leqslant \lambda$ be a granularity parameter. Construct an unweighted instance H' of size $O(m\kappa^2/\sigma^2)$ as follows: Replicate each vertex u in H of weight $w(u)$ by $\lfloor w(u)/\sigma \rfloor$ copies in H'; if u and v are connected in H, connect all copies of u to all copies of v in H'. Then weights are roughly preserved: any vertex u of weight $w(u)$ in H will have copies of total weight $w(u)(1 \pm O(\lambda/\sigma))$ in H'. Therefore, it is not hard to see that objective value is roughly preserved, $\mathrm{val}(H') = \mathrm{val}(H)(1 \pm O(\lambda/\sigma))$. Further, any vertex u in H has at most κ/σ copies in H'.

As observed by Trevisan [2001], the graph H is a union of bipartite complete subgraphs. More precisely, for every index i in the proof for Π, there is a bipartite complete subgraph between the sets Z_i and O_i of configurations, where configurations in Z_i query index i and expect an answer of zero, and configurations in O_i query index i and expect an answer of one. Further, the set of edges in H is the union of all such bipartite complete subgraphs over index i. This bipartite complete subgraph structure is preserved by the vertex duplication process.

Also, the sets Z_i and O_i in H have the same total weight, and in fact there is a weight-preserving bijection between Z_i and O_i. This bijection is inherited from the corresponding bijection of the subgroup C, thanks to its balanced property. As a result, in the instance H' after duplication of vertices, the vertex sets Z_i and O_i have the same size ℓ_i.

We now replace the bipartite complete subgraph between O_i and Z_i with a bipartite disperser on $([\ell_i], [\ell_i])$, for all index i. The graph after replacement is H''.

Proposition B.1. *For every $\delta > 0$ and any $\ell \geqslant 1$, there is a bipartite graph on $(([\ell], [\ell]), E)$ of degree at most $d = O(\delta^{-1} \log(\delta^{-1}))$ such that for any $A, B \subseteq [\ell]$, $|A| \geqslant \lfloor \delta\ell \rfloor$ and $|B| \geqslant \lfloor \delta\ell \rfloor$, some edge in E goes between A and B, so $(A \times B) \cap E \neq \emptyset$.*

A random bipartite graph is well-known to be a δ-disperser (for completeness, we include a proof below). We can therefore find (and verify) a disperser deterministically by exhaustive search in time $\exp(\text{poly}(\ell_i))$.

To bound ℓ_i, we first bound the maximum size W of Z_i in H (measured by the number of vertices, disregarding weights). Then W times the maximum number of copies of a vertex will upperbound ℓ_i. It is not hard to see that $W = O_{\varepsilon,k}(1)$, where $O_{\varepsilon,k}(1)$ denotes a quantity bounded by a function of ε and k. Indeed, W is at most $2^f \Delta(M)$, where $\Delta(M)$ is the maximum number of constraints incident on a variable in the instance M of Theorem 1.1 (disregarding weight on constraints). To bound $\Delta(M)$, observe that $\Delta(L) = O_{\varepsilon,k}(1)$ for the Label-Cover instance L of Theorem A.1. Also, $\Delta(M_j) = O_{\varepsilon,k}(1)$, where M_j is the instance from Theorem 5.4. Further, direct sum preserves boundedness of Δ, since $\Delta(M \oplus M') = \Delta(M)\Delta(M')$. This shows that $W = O_{\varepsilon,k}(1)$.

We bound the number of copies of a vertex in the replication step by κ/σ. To bound κ/σ, we first bound the ratio $\rho(M) = \kappa(M)/\sigma(M)$ of the maximum weight constraint to minimum weight constraint in a CSP instance M. Then $\rho(L) = 1$ for the Label-Cover instance L in Theorem A.1, because L is a bi-regular bipartite graph. After composing with the dictator test, $\rho(M_j)$ is at most $O_{\varepsilon,\eta,k}(1)$. Finally, $\rho(M \oplus M') = \rho(M)\rho(M')$. Hence the ratio κ/λ for the FGLSS graph H is $O_{\varepsilon,\eta,k}(1)$. If we pick $\sigma = \varepsilon\lambda$, then $\ell_i = O_{\varepsilon,\eta,k}(1)$.

The disperser replacement step increases the objective value by at most $k\delta$ [Trevisan 2001]. We will therefore choose $\delta = s2^{-f}/k$, and the degree bound for H'' becomes $D = O(k/\delta \cdot \log(1/\delta)) = O(k^3 2^k)$. The hardness ratio is $O(c/s) = O(k/2^k) = O(\log D)^4/D$. $\square$

Proof of Proposition B.1. We may assume $\ell \geqslant \delta^{-1}\log(\delta^{-1})$ (otherwise, just take the bipartite complete graph). Assume for now that $\delta\ell$ is an integer.

Denote by U, V the two vertex subsets of size ℓ. We pick a random degree-d bipartite (multi)-graph on (U, V), generated as the union of d independent random perfect matchings.

Consider $A \subseteq U$ of size $\delta\ell$ and $B \subseteq V$ of size $\delta\ell$. The probability that in a perfect matching, all edges from A miss B is $\binom{(1-\delta)\ell}{\delta\ell} / \binom{\ell}{\delta\ell} \leqslant (1-\delta)^{\delta\ell}$. Hence A shares no edges with

B with probability at most $(1-\delta)^{d\delta\ell}$. Taking union bound over all choices of A and B, the random graph is a δ-disperser except with probability at most

$$\binom{\ell}{\delta\ell}\binom{\ell}{\delta\ell}(1-\delta)^{d\delta\ell} \leqslant \left(\frac{e^2}{\delta^2}(1-\delta)^d\right)^{\delta\ell},$$

where we have used $\binom{n}{r} \leqslant (en/r)^r$. The quantity in bracket on the RHS is less than 1 when $d = O(\delta^{-1}\log(\delta^{-1}))$.

When $\delta\ell$ is not an integer, it is easy to get the same conclusion using $\ell \geqslant \delta^{-1}\log(\delta^{-1})$ and appropriate approximations. $\square$

Appendix C

Some predicates

C.1 Samorodnitsky–Trevisan hypergraph predicates

Let $k = 2^r - 1$. The Samorodnitsky and Trevisan [2009] hypergraph predicate of arity k is the dual Hamming code (or truncated Hadamard code) C of block length k and dimension r over $\mathbb{F}_2$. If we index the positions of a codeword $c = (c_S)_{\emptyset \neq S \subseteq [r]}$ by nonempty subsets S of $[r]$, the codewords are given by

$$C = \left\{ c = \left(\sum_{i \in S} y_i \right)_{\emptyset \neq S \subseteq [r]} \middle| \; y_1, \ldots, y_r \in \mathbb{Z}_2 \right\}.$$

C.2 Håstad predicates

We describe a predicate due to Johan Håstad and announced in [Makarychev and Makarychev 2012]. This predicate is used in Corollary 1.7.

Let $k \leqslant 2^t$, $q = 2^s$, and suppose $t \geqslant s$. A Håstad predicate is over $G = \mathbb{Z}_2^s$. We pick a random tuple $\boldsymbol{c} \in G^k$ as follows. Pick random $\boldsymbol{a} \in \mathbb{F}_{2^t}$ and $\boldsymbol{b} \in \mathbb{Z}_2^s$, and set

$$\boldsymbol{c}_i = \pi(\boldsymbol{a} \cdot \bar{i}) + \boldsymbol{b},$$

where $\bar{i}$ denotes the i-th element from $\mathbb{F}_{2^t}$, and $\pi : \mathbb{F}_{2^t} \to \mathbb{Z}_2^s$ is any surjective group homomorphism (e.g. π takes the first s bits in some vector space representation of $\mathbb{F}_{2^t}$ over $\mathbb{F}_2$).

Let C be the collection of random tuples $\boldsymbol{c}$ generated as above. Then C has size at most qk. Further, C is balanced pairwise independent, because for every $i \neq j \in [k]$, the difference

$$\boldsymbol{c}_i - \boldsymbol{c}_j = \pi(\boldsymbol{a} \cdot \bar{i}) - \pi(\boldsymbol{a} \cdot \bar{j}) = \pi(\boldsymbol{a} \cdot (\bar{i} - \bar{j}))$$

is uniformly random over $\mathbb{Z}_2^s$, for any fixed $\boldsymbol{b}$.

Håstad predicates require q to be a prime power. To obtain Corollary 1.7 where q is arbitrary, pick the smallest power of two $q' \geqslant q$, and apply Makarychev's randomized reduction [Austrin and Mossel 2009, Proposition B.1] from domain size q' to domain size q.

Appendix D

Sum-of-squares integrality gaps

Sum-of-squares programs are powerful hierarchies of semidefinite programs proposed independently by Parrilo [2000], Lasserre [2001], and others (see [Barak et al. 2012; O'Donnell and Zhou 2013] for the history). In this chapter, we observe that Schoenebeck's [2008] sum-of-squares gap construction for Max-k-XOR also works for the CSPs in Theorem 1.1, drawing a pleasing parallel between sum-of-squares gap construction and NP-hardness results. Even without the result in this chapter, Theorem 1.1 implies a such a gap via reduction, but the rank of the sum-of-squares solution will not be linear, due to the blow-up in size from direct sum.

Previously, Tulsiani [2009] extended Schoenebeck's construction to any predicate that is a linear code of dual distance at least 3 over a prime field. Later Schoenebeck [2008] simplified his own proof of Max-k-XOR using Fourier analysis. Not surprisingly, his new proof can be further generalized to arbitrary abelian group using Pontryagin duality, as shown below. For intuition about the construction, see [Schoenebeck 2008]. We remark that Schoenebeck's proof was based on Feige and Ofek [2006], and some of Schoenebeck's ideas were applied independently by Grigoriev [2001] to related problems.

D.1 Preliminaries

Given an abelian group G, its dual group $\hat{G}$ is the abelian group of characters on G, under pointwise multiplication. The inverse of $\chi \in \hat{G}$ is therefore $\overline{\chi}$. Pontryagin duality says that G is naturally isomorphic to the dual of $\hat{G}$ (i.e. double dual of G), via the "evaluation map"

$$g \in G \quad \mapsto \quad \{\chi \in \hat{G} \mapsto \chi(g)\}.$$

Given a subgroup H of G, denote by $H^\perp = \{\chi \in \hat{G} \mid \chi(h) = 1 \ \forall h \in H\}$ the annihilator of H. We remark that annihilator is only defined with respect to an ambient group G, which will always be clear from the context. The following fact is well known.

Proposition D.1 ([Hewitt and Ross 1994, Theorems 23.25 and 24.10])**.** *Let Λ be a subgroup of a finite abelian group Γ. Then (a) $\widehat{\Gamma/\Lambda} \cong \Lambda^\perp$ and (b) $(\Lambda^\perp)^\perp = \Lambda$.*

A (linear) equation is a pair $(\chi, z) \in \widehat{G^V} \times \mathbb{T}$, encoding the constraint $\chi(f) = z$ for an assignment $f : V \to G$. Since $\widehat{G^V}$ is isomorphic to $\hat{G}^V$, we write $\hat{G}^V$ in place of $\widehat{G^V}$ for better typography. The support of $\chi \in \hat{G}^V$ is $\operatorname{supp}(\chi) \triangleq \{v \in V \mid \chi \text{ is } v\text{-relevant}\}$, and the degree of χ is the size of its support. Denote by Ω_t the collection of χ of degree at most t.

Definition D.2. Given a collection R of equations, its width-t resolution $\Pi_t(R) \subseteq \hat{G}^V \times \mathbb{T}$ contains all equations in R and those derived via the resolution step

$$(\chi, z), (\psi, y) \in \Pi_t(R) \text{ and } \chi\overline{\psi} \in \Omega_t \quad \Longrightarrow \quad (\chi\overline{\psi}, z\overline{y}) \in \Pi_t(R).$$

The resolution has no contradiction if $(\mathbf{1}, z) \in \Pi_t(R)$ implies $z = 1$.

In this chapter, a $\mathsf{Max}\text{-}C$ instance $M = (V, \boldsymbol{Q})$ will not be k-partite, so all variables $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ of the k-tuple $\boldsymbol{v}$ in a question $\boldsymbol{Q} = (\boldsymbol{v}, \boldsymbol{b})$ come from the same variable set V. Let R_M be the set of equations from constraints in M, defined as

$$R_M \triangleq \{(\chi, \chi(\boldsymbol{b})) \mid (\boldsymbol{v}, \boldsymbol{b}) \in M, \chi \in C^\perp \subseteq \hat{G}^{\boldsymbol{v}}\}.$$

We say that M has resolution width at least t if $\Pi_t(R_M)$ has no contradiction.

We state the results below in terms of Lasserre integrality gaps, but our lower bound also rules out sum-of-squares refutations in Parrilo's hierarchy (Remark D.4). Our definition of Lasserre solution is a rephrasing of the one in [Tulsiani 2009].

Definition D.3. A rank-t Lasserre solution U for a CSP instance $M = (V, \boldsymbol{Q})$ over domain Σ is a collection $\{U_f \mid f \in \Sigma^S, S \subseteq V \text{ s.t. } |S| \leqslant t\}$ of vectors, one for each partial assignment $f : S \to \Sigma$ on a subset S of size at most t.

The Lasserre solution induces a collection of distributions $\{\mu_W \in \triangle_{\Sigma^W} \mid W \subseteq V \text{ s.t. } |W| \leqslant 2t\}$ over partial assignments, subject to the following condition: For any two partial assignments $f \in \Sigma^S$ and $g \in \Sigma^T$ with $|S|, |T| \leqslant t$, we have

$$\langle U_f, U_g \rangle = \mathop{\mathbb{P}}_{\boldsymbol{h} \sim \mu_{S \cup T}}[\boldsymbol{h} \restriction_S = f \text{ and } \boldsymbol{h} \restriction_T = g]. \tag{D.1}$$

The value of the Lasserre solution is $\operatorname{val}(M, U) = \mathbb{E}_{\boldsymbol{Q}} \mathbb{P}[\boldsymbol{Q} \text{ is satisfied under } \mu_{\langle \boldsymbol{Q} \rangle}]$, where $\langle \boldsymbol{Q} \rangle \subseteq V$ denotes the set of variables that $\boldsymbol{Q}$ depends on.

Remark D.4. Sum-of-squares refutations in Parrilo's hierarchy are slightly stronger than sum-of-squares proofs in Lasserre's hirerachy [O'Donnell and Zhou 2013], but the difference is inconsequential in our setting. A degree-t sum-of-squares refutation for a $\mathsf{Max}\text{-}C$ instance $M = (V, \boldsymbol{Q})$ involves multivariate polynomials over indeterminates $\{x_{v,a}\}_{v \in V, a \in G}$. The refutation is associated with equality relations $\{p = 0\}_{p \in A}$ for a collection A of polynomials p; these relations state that (1) $\mathbb{P}[\boldsymbol{Q} \text{ is satisfied under } \mu_{\langle \boldsymbol{Q} \rangle}] = 1$ for all $\boldsymbol{Q}$; (2) total probability mass of local assignments on S is one for any $S \subseteq V, |S| \leqslant t$; and (3) $x_{v,a}$'s are $\{0, 1\}$-indicator variables. The refutation takes the form

$$-1 = s + \sum_{p \in A} q_p p,$$

where s is a sum of squares and q_p's are arbitrary polynomials such that $\deg(s), \deg(q_p p) \leqslant 2t$. Because our rank-$t$ Lasserre solution will satisfy conditions (1), (2) and (3), it also satisfies all the equality relations from A, ruling out any degree-t refutation.

D.2 From resolution complexity to SDP solution

A key step will be the following generalization of [Tulsiani 2009, Theorem B.1].

Theorem D.5. *Let G be an abelian group, and C a subgroup of G^k. If a* Max-C *instance M has resolution width at least $2t$, then there is a rank-t Lasserre solution to M of value* 1*.*

Given the resolution proof $\Pi = \Pi_{2t}(R_M)$, denote by $\Lambda = \{\chi \mid (\chi, z) \in \Pi\}$ the collection of χ's appearing in an equation. If Π has no contradiction, then for every $\chi \in \Lambda$, there is a unique $z(\chi) \in \mathbb{T}$ such that $(\chi, z(\chi)) \in \Pi$. Otherwise the existence of distinct $(\chi, z), (\chi, y)$ in Π implies $(\mathbf{1}, 1) \neq (\mathbf{1}, z\overline{y}) \in \Pi$, a contradiction (pun intended). By definition of the resolution step, if $\chi, \psi, \chi\psi \in \Lambda$, then

$$z(\chi\psi) = z(\chi)z(\psi), \tag{D.2}$$

so $z : \Lambda \to \mathbb{T}$ is a homomorphism wherever it is defined.

The key observation is that if $\chi \notin \Lambda$, then χ does not enforce any constraint on partial assignments. We make this precise in Eq. (D.3) below. For $W \subseteq V$, let $\Lambda_W = \{\chi \in \Lambda \mid \operatorname{supp}(\chi) \subseteq W\}$, which will be considered as a subgroup of $\hat{G}^W$. Let H_W be the set of partial assignments on W that satisfy all the constraints contained in W,

$$H_W = \{h \in G^W \mid \forall \chi \in \Lambda_W,\ \chi(h) = z(\chi)\}.$$

We now show that for every W of size at most $2t$ and every $\chi \in \hat{G}^W \setminus \Lambda_W$,

$$\mathop{\mathbb{E}}_{\boldsymbol{h} \in H_W} \chi(\boldsymbol{h}) = 0. \tag{D.3}$$

Indeed, H_W is a coset of $\Lambda_W^{\perp}$, so Eq. (D.3) follows from Proposition D.6 with $\Lambda := \Lambda_W, \Gamma := \hat{G}_W, H := H_W$.

Proposition D.6. *Let Λ be a subgroup of an abelian group Γ, and $H \subseteq \hat{\Gamma}$ be a coset of $\Lambda^{\perp}$. Then for any $\chi \in \Gamma$,*

$$\chi \in \Lambda \iff \mathop{\mathbb{E}}_{\boldsymbol{h} \in H} \boldsymbol{h}(\chi) \neq 0.$$

Proof. Let $H = h\Lambda^{\perp}$. We have

$$\mathop{\mathbb{E}}_{\boldsymbol{h} \in H} \boldsymbol{h}(\chi) = h(\chi) \cdot \mathop{\mathbb{E}}_{\boldsymbol{h} \in \Lambda^{\perp}} \boldsymbol{h}(\chi) = h(\chi) \cdot \mathop{\mathbb{E}}_{\boldsymbol{z} \in \chi(\Lambda^{\perp})} \boldsymbol{z},$$

where second equality uses the fact that χ is a homomorphism from $\hat{\Gamma}$ to $\mathbb{T}$, by Pontryagin duality. Now the RHS is non-zero if and only if $\chi(\Lambda^{\perp})$ contains only one element, that is $\chi(\Lambda^{\perp})$ is the trivial subgroup $\{1\}$ of $\mathbb{T}$. The latter condition is equivalent to $\chi \in (\Lambda^{\perp})^{\perp}$, and the result follows by Proposition D.1(b). $\square$

Partition Ω_t into equivalence classes $[\chi]$'s so that $[\chi] = [\psi]$ if $\chi\overline{\psi} \in \Lambda$. It is easily checked that the latter condition is indeed an equivalence relation. Also fix an arbitrary representative χ' for each equivalence class $[\chi]$. In the Lasserre vector construction, there will be an orthonormal set of vectors $e_{[\chi]}$'s, one for each equivalent class.

Our goal is Lasserre vectors U_f for partial assignments $f : S \to G$, and to this end we first construct Lasserre vectors U_A for any t-junta A, which is a function $A : G^V \to \mathbb{C}$ depending on at most t variables. Formally, let $\mathrm{supp}(A)$ be the smallset subset $S \subseteq V$ on which there is $B : S \to G$ satisfying $A(h) = B(h \restriction_S)$ for all $h \in G^V$. Then A is a t-junta if $\mathrm{supp}(A)$ has size at most t. Since any t-junta A is a linear combination of characters of degree at most t, it suffices to define the Lasserre vector

$$U_\chi = z(\chi\overline{\chi'})e_{[\chi]}$$

for $\chi \in \Omega_t$ and extend the definition to an arbitrary t-junta A by linearity, i.e.,

$$A = \sum_{\chi\in\hat{G}^S} \hat{A}(\chi)\chi \quad \Longrightarrow \quad U_A = \sum_{\chi\in\hat{G}^S} \hat{A}(\chi)U_\chi,$$

where $S = \mathrm{supp}(A)$.

The following proposition highlights the main property.

Proposition D.7. *For any t-juntas $A, B : G^V \to \mathbb{C}$, let $W = \mathrm{supp}(A) \cup \mathrm{supp}(B)$. Then*

$$\langle U_A, U_B\rangle = \mathop{\mathbb{E}}_{\boldsymbol{h}\in H_W}[A(\boldsymbol{h})\overline{B}(\boldsymbol{h})].$$

Proof. By linearity, it suffices to show that for any $\chi, \psi \in \Omega_t$, if $W = \mathrm{supp}(\chi) \cup \mathrm{supp}(\psi)$ (which has size at most $2t$), then

$$\langle U_\chi, U_\psi\rangle = \mathop{\mathbb{E}}_{\boldsymbol{h}\in H_W}[\chi(\boldsymbol{h})\overline{\psi}(\boldsymbol{h})] = \mathop{\mathbb{E}}_{\boldsymbol{h}\in H_W}[\chi\overline{\psi}(\boldsymbol{h})].$$

When $[\chi] \neq [\psi]$, the LHS is zero because $e_{[\chi]}$ and $e_{[\psi]}$ are orthogonal, and the RHS is also zero by Eq. (D.3).

When $[\chi] = [\psi]$, the LHS is $z(\chi\overline{\chi'})\overline{z}(\psi\overline{\chi'}) = z(\chi\overline{\psi})$ by Eq. (D.2), and the RHS is also $z(\chi\overline{\psi})$ by definition of H_W and the fact that $\chi\overline{\psi} \in \Lambda$. □

Proof of Theorem D.5. We will consider the indicator function $A : G^V \to \mathbb{R}$ for a partial assignment $f : S \to G$, defined as

$$A(h) = \mathbb{I}(h \restriction_S = f).$$

Then A is a t-junta. We then define U_f as U_A.

For any partial assignments $f \in G^S, g \in G^T$,

$$\langle U_f, U_g\rangle = \mathop{\mathbb{E}}_{\boldsymbol{h}\in H_W}[\mathbb{I}(\boldsymbol{h} \restriction_S = f)\mathbb{I}(\boldsymbol{h} \restriction_T = g)]$$

by Proposition D.7. Taking μ_W as the uniform distribution over H_W, the vectors U_f's satisfy the Lasserre constraints Eq. (D.1).

The Lasserre solution has value 1, because every constraint $\boldsymbol{Q} \in M$ is satisfied by every $f \in H_{\langle \boldsymbol{Q} \rangle}$. Indeed, since $\boldsymbol{Q} = (\boldsymbol{v}, \boldsymbol{b})$ induces linear equations $\{(\chi, \chi(\boldsymbol{b})) \mid \chi \in C^{\perp} \subseteq \hat{G}^{\boldsymbol{v}}\}$ in Π, we have

$$f \in H_W \quad \Longrightarrow \quad \chi(f - \boldsymbol{b}) = 1 \quad \forall \chi \in C^{\perp} \quad \Longleftrightarrow \quad f - \boldsymbol{b} \in (C^{\perp})^{\perp} = C,$$

where the equivalence is Pontryagin duality and the last equality is Proposition D.1(b).

The vectors U_f may have complex entries, but equivalent *real* vectors exist. Indeed, the Gram matrix $[\langle U_f, U_g \rangle]_{f,g}$ has only real entries and is positive semidefinite over $\mathbb{C}$, and hence over $\mathbb{R}$. □

D.3 Resolution complexity of random instances

As usual, a random Max-C instance M will be a Lasserre gap instance. To be precise, the m constraints of M are chosen independently (with replacement), where each constraint $\boldsymbol{Q} = (\boldsymbol{v}, \boldsymbol{b})$ is uniformly random in $\binom{V}{k} \times G^k$.

Theorem D.8. *Let G be a finite abelian group, and C a balanced pairwise independent subgroup of G^k for some $k \geqslant 3$. Let M be a random instance of Max-C with $m = \Delta n$ constraints and n variables. Then M has resolution width $n/\Delta^{O(1)}$ with probability $1 - o_{n;\Delta}(1)$.*

Proof sketch. This follows by Tulsiani's proof [Tulsiani 2009, Theorem 4.3]. As in his proof, we need M to be expanding (i.e. every set of $s \leqslant \Omega(1/\Delta)^{25} n$ constraints contains at least $(k - 6/5)s$ variables); the expansion property is guaranteed by [Tulsiani 2009, Lemma A.1(2)]. In our setting, the number of variables involved in an equation (χ, z) is simply the degree of χ.

Also, a subgroup $C \subseteq G^k$ has dual distance at least 3 (i.e. non-trivial characters in $C^{\perp}$ have degree at least 3) if and only if C is balanced pairwise independent. To see this, for any $i \neq j \in [k]$, let $C^{ij} \triangleq \{(c_i, c_j) \mid c \in C\} \subseteq G^{\{i\}} \times G^{\{j\}} \cong G^2$ be the projection of C to i and j coordinates. Balanced pairwise independence of C means for all $i \neq j \in [k]$, we have $C^{ij} \cong G^2$, which is equivalent to $(C^{ij})^{\perp} = \{\mathbf{1}\} \subseteq \mathbb{T}$, by Proposition D.1(a) and the isomorphism $\hat{\Gamma} \cong \Gamma$ for any finite abelian group Γ. Now the condition $(C^{ij})^{\perp} = \{\mathbf{1}\}\ \forall i \neq j \in [k]$ is the same as non-trivial characters in $C^{\perp}$ having degree at least 3.

One can check that Tulsiani's proof goes through. We omit details. □

It is also well known that a random Max-C instance has value close to $|C|/|G|^k$ [Tulsiani 2009, Lemma A.1(1)]. We summarize the result of this chapter in the next theorem, which follows by combining Theorem D.5, Theorem D.8 and [Tulsiani 2009, Lemma A.1(1)], and choosing $\Delta = O(|G|^k/\varepsilon^2)$.

Theorem D.9. *Let G be a finite abelian group, and C be a balanced pairwise independent subgroup of G^k for some $k \geqslant 3$. For any $\varepsilon > 0$, some* Max-C *instance M on n variables has a rank-*$(\mathrm{poly}(\varepsilon/|G|^k) \cdot n)$ *Lasserre solution of value* 1 *and satisfies* $\mathrm{val}(M) \leqslant |C|/|G|^k + \varepsilon$.

Our Theorem D.9 is a generalization of Tulsiani's [2009, Theorem 4.6] and a sum-of-squares gap analogue of Theorem 1.1. Examples of predicates satisfying our theorem but not Tulsiani's are Håstad predicates in Appendix C.2.