Hardness of Approximate Coloring

A Thesis

Submitted to the
Tata Institute of Fundamental Research, Mumbai
for the degree of
Doctor of Philosophy
in
Computer Science

by
Girish Varma

School of Technology and Computer Science
Tata Institute of Fundamental Research
Mumbai
September 2015
Final Version Submitted in January 2016

Thesis Advisor: Prof. Prahladh Harsha Girish Varma

Hardness of Approximate Coloring

Abstract

The graph coloring problem is a notoriously hard problem, for which we do not have efficient algorithms. A *coloring* of a graph is an assignment of colors to its vertices such that the end points of every edge have different colors. A k*-coloring* is a coloring that uses at most k distinct colors. The *graph coloring* problem is to find a coloring that uses the minimum number of colors. Given a 3-colorable graph, the best known efficient algorithms output an $n^{0.199\cdots}$-coloring. It is known that efficient algorithms cannot find a 4-coloring, assuming P$\neq$NP (such results are commonly known as *hardness* results). Hence there is a large gap ($n^{0.199\cdots}$ vs 4) between what current algorithms can achieve and the hardness results known.

In this thesis, we narrow the aforesaid gap for some generalizations of graph coloring, by giving improved hardness results (for exponentially better parameters in some cases). Some of our main results are as follows:

1. For the case of almost 3-colorable graphs, we show hardness of finding a $2^{\text{poly}(\log\log n)}$-coloring, assuming a variant of the Unique Games Conjecture (UGC).
2. For the case of 3-colorable 3-uniform hypergraphs, we show quasi-NP-hardness of finding a $2^{O(\log\log n/\log\log\log n)}$-coloring.
3. For the case of 4-colorable 4-uniform hypergraphs, we show quasi-NP-hardness of finding a $2^{(\log n)^{1/21}}$-coloring.
4. For the problem of the approximating the covering number of CSPs with non-odd predicates, we show hardness of approximation to any constant factor, assuming a variant of UGC.

Contents

അമ്മയ്ക്കും അച്ഛനും ചേട്ടനും

പ്രിയപ്പെട്ട കുടുംബാംഗങ്ങൾക്കും

Acknowledgments

I THANK MY ADVISOR AND TEACHERS AT TIFR. During the days of my course work at TIFR, the lectures of Prahladh Harsha and Jaikumar Radhakrishnan have been great. I learnt the maths required for research from them. Their advise about research as well as personal matters were invaluable. Prahladh has been an excellent advisor. In the later years at TIFR, I had the privilege to work with Srikanth Srinivasan (IITB), whose sheer intensity of doing research, inspired me to make a final push towards this thesis. I also thank Manoj Gopalkrishnan, for introducing me to a wider world of theory in other sciences.

I THANK Kurt Mehlhorn (MPII, Saarbrücken), Irit Dinur (Weizmann Institute) & Subhash Khot (NYU) for inviting me to their respective institutes for short visits. These visits and the discussions with them were enlightening and highly motivating. I thank Ryan O'Donnell (CMU) and Rishi Saket (IBM IRL) for reviewing my thesis. Their comments and suggestions have improved the quality of the write up.

I THANK MY FRIENDS. I could not have held on at TIFR till the completion of this thesis, without the innumerable friendships. I thank my batch mates Rakesh, Shishir, Tapan, Chandra , Mohit & my juniors Swagato, Sagnik, Sarat, Karthyek, Bodhayan, Gugan, Nithin, Deepesh, Kshitij, Nikhil, Suhail, Tulasi & my seniors Ajesh, Kishore, Chinmay, Saswata & post docs Nutan and Sreejith. I thank them for their friendships, and for the discussions about technical matters. I also thank Ajesh, Nutan, Ramprasad, Rakesh and Amey for successful collaboration on papers. I thank Priti, John, Carina,

Parul, Venkatesh, Mathi, Chaitanya, Shampa, Sugato, Yadu, Aiswarya, Atul, Deepti, Chien-Chung & Agnes for their companionship.

I thank my teachers from NIT Calicut, where I got my undergraduate degree. I learnt about the science behind computers from Murali Krishnan, Saleena Nazeer, Vineeth Paleri and Vinod Pathari. They are primarily responsible for my interest in theoretical computer science.

Finally, I thank Google India, for giving me a generous fellowship. I thank Indo-German Max Planck Center for Computer Science (IMPECS) funding my internship at MPII, Saarbrücken. I thank the Deans and the Subject Board Members who have been supportive and given me considerable time to finish my thesis. Thanks to John Barretto and the other members of the administration for being very efficient in all my institutional formalities at TIFR.

Simple can be harder than complex: You have to work hard to get your thinking clean to make it simple. But it's worth it in the end because once you get there, you can move mountains.

Steve Jobs

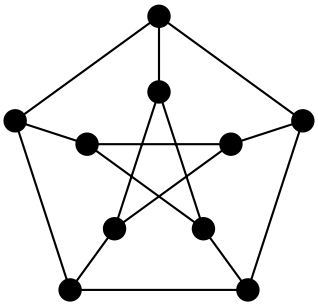
Part I

Introduction

1

Puzzles, Algorithms & Hardness

1.1 Coloring Puzzle



Consider the following puzzle. Given a figure as above, the goal is to give colors to the circles such that for every line, its end points have different colors. Furthermore, the number of colors used needs to be minimized. Without this condition the problem is trivial since using a different color for every circle would be a solution irrespective of the figure.

Puzzles like above are very commonly encountered in a variety of real life instances. However known algorithms takes far too much time to complete even on instances with 20 vertices. This thesis is about an explanation for the lack of efficient algorithms, for coloring like problems.

Figures like above are commonly called *graphs* (strictly speaking they are undirected graphs, but in this thesis will only be concerned with undirected graphs) . A graph $G =$

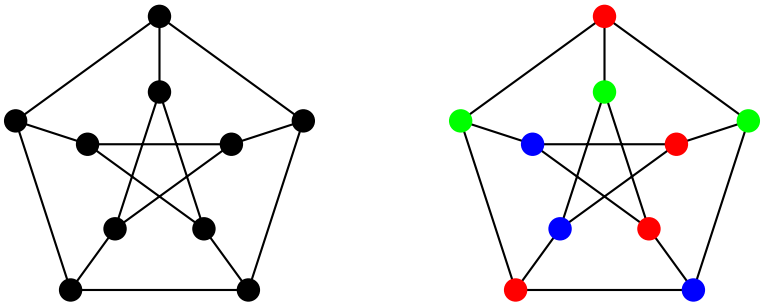


Figure 1.1: The figure on the right is a 3-coloring of the graph on the left. However it does not have a 2-coloring.

(V, E) consists of a set of *vertices* V (the circles) and a set E containing some pairs of vertices, called *edges* (the lines).

Given a graph $G = (V, E)$ and a number C, a C-coloring of G is an assignment of *colors* denoted by $\{1, \cdots, C\}$ to the vertices such that the end points of every edge have different colors. For a given graph, a C-coloring might not exist for all values of C. But for any graph, the coloring which gives different colors to all vertices in V, is an n-coloring, where n is the *size* (the number of vertices in V) of G.

Definition 1.1. *The* Graph-Coloring *problem is, given a graph, find a coloring, using the minimum number of colors. This minimum value is commonly known as the* chromatic number *of the graph.*

1.2 Efficient Algorithms

If the maximum degree of G is Δ then the following simple algorithm computes a $(\Delta+1)$-coloring in time bounded by the number of edges.

for $v \in V$ do
  give v a color in $\{1, \cdots, \Delta + 1\}$ different from the colors of its already colored neighbours.
end

Algorithm 1: For graphs with maximum degree $\leq \Delta$.

Algorithm 1, does not solve the GRAPH-COLORING problem since the chromatic number can be much smaller than $\Delta + 1$. Suppose the graph has a c-coloring there is a simple algorithm to find it, which takes $c^n \cdot m$ time, where m is the number of edges.

for *every assignment* $f \in \{1, \cdots, c\}^V$ *of colors to the vertices* do
    for *every edge* do
        check if f assigns different colors to the end points.
    end
end

Algorithm 2: Brute Force Algorithm

However note that the running time of Algorithm 2 grows *exponentially* in n. Even for $n = 20$, the running time ($> c^{20}$) is prohibitively large, while for Algorithm 1 it is still a reasonable number. This motivates the definition of efficient algorithms, as ones that has running time bounded by a polynomial in the input size. For simplicity, we consider only *decision problems* (i.e. problems that have a Boolean answer). For such problems the inputs are partitioned into YES and NO instances. We will often specify a decision problem by the set of YES instances. Though GRAPH-COLORING problem is not a decision problem, we can consider the problem which have the number C also in the input, where the goal is to check if the graph has a C-coloring.

Informal Definition 1.2. P *is the class of decision problems, that has an algorithm with running time bounded by a polynomial in* n *(the input size).*

Note that for GRAPH-COLORING, $(G, C) \in$ YES (i.e. G has a C-coloring), then there is a certificate that certifies that it is YES instance, that can be verified in time proportional to the number of edges. The certificate is simply the C-coloring of the graph.

And if $(G, C) \in$ NO, then any assignment of colors from $\{1, \cdots, C\}$, will leave some edge monochromatic. This property is true for a large class of problems.

Informal Definition 1.3. NP *is the class of decision problems, for which there is an algorithm V which takes an input x and a proof π, runs in time polynomial in the size of x and satisfies the following properties.*

- *Completeness : If $x \in$* YES *then there exists π such that $V(x, \pi) = 1$.*
- *Soundness : If $x \in$* NO *then for any π, $V(x, \pi) = 0$.*

For any NP problem, there is a trivial algorithm similar to Algorithm 2, which runs in time c^n for some constant c. The algorithm just tries all possible certificates with the verification procedure. It is a major open problem if P= NP. A surprising result is that there are certain class of problems called NP-Complete which capture the hardness of solving problems in NP. That is, an instance of any NP problem could be converted efficiently to an instance of such problems. The Graph-Coloring problem is an NP-Complete problem. Hence unless P $=$ NP, we cannot hope to have polynomial time algorithms for Graph-Coloring. One can ask if there are polynomial time algorithms for a relaxed version of the problem.

1.3 Approximation Algorithms

A relaxed version of the Graph-Coloring problem is to compute the chromatic number approximately. An *approximation algorithm* for chromatic number with approximation *factor* $F \geq 1$, always outputs a number between C and $C \cdot F$, where C is the chromatic number of the input graph.

However for a general graph, this relaxation is also a *hard* problem. That is, assuming the famous P $\neq$ NP conjecture, it is known that this problem cannot be solved

in polynomial time. Feige & Kilian [FK00] showed that computing this number approximately within a factor of $n^{1-\varepsilon}$ for any small constant $\varepsilon > 0$ is known to be hard, assuming a different complexity conjecture. Therefore, there is not much hope of having an efficient algorithm, which does much better than the trivial n-coloring.

Since the general approximation problem is hard, the focus shifted on solving it for subclasses of graphs. A natural subclass of graphs to consider, are the ones for which the chromatic number is a small constant c. For $c = 2$, such graphs are commonly called *bipartite*. There is a simple linear time algorithm for finding a 2-coloring in such graphs. It just assigns a vertex one color, the other color to all its neighbours, and continues until all vertices are colored. However for 3-colorable graphs, finding a 3-coloring is NP-Hard (it cannot be solved by efficient algorithms, assuming P$\neq$ NP).

Hence a long series of works, was aimed at solving this problem approximately.

Definition 1.4. Approximate-Graph-Coloring(c, C) *problem is to find a C-coloring, when the input graph is promised to be c-colorable.*

Remark 1.5. *We will often be considering the decision version of the problem, specified by disjoint sets of* YES *(c-colorable graphs) and* NO *(graphs with chromatic number $> C$) instances. The goal is to have an algorithm to accept all* YES *instances and reject all* NO *instances. The algorithm can either accept or reject inputs which are neither* YES *nor* NO.

For 3-colorable graphs, Wigderson [Wig83] gave the first not trivial improvement of finding a $O(\sqrt{n})$-coloring using combinatorial techniques. This was further improved to $O(n^{3/8})$-coloring by Blum [Blu94]. A major breakthrough was made by Karger, Motwani & Sudan [KMS98], using semi-definite programming (SDP). For a 3-colorable graph with maximum degree d, they gave an $O(d^{1/3})$-coloring algorithm. Combining this with Wigderson's algorithm, they obtained a $O(n^{1/4})$-coloring algorithm. Blum & Karger [BK97] combined the combinatorial methods of Blum [Blu94]

with the SDP to get an $O(n^{3/14})$-coloring. The current best known (see results of Kawarabayashi & Thorup [KT14]) efficient algorithms output a $n^{0.19996}$-coloring.

1.4 What this thesis is about?

This thesis is about giving an explanation for the lack of efficient algorithms for some generalizations of the Approximate-Graph-Coloring(c, C) problem, using the theory of NP-Completeness and complexity conjectures similar to P$\neq$ NP. That is, we prove that efficient algorithms for the problem will imply that the corresponding conjecture is false. Such results are called *hardness results* and the area in general, is commonly called in literature as the *hardness of approximation*. The focus of our hardness results will be the case when c is a small constant and C can be any large constant or a function that depends on n. As described in the previous section, there are algorithms which solve these problems for $C = n^{\alpha}$ for some constant $\alpha < 1$. We prove hardness results for larger values (exponentially larger in some cases) of C than was previously known. We will describe these generalizations in the next three sections.

1.4.1 Almost Coloring

Assuming P$\neq$ NP, Khanna et al. [KLS00] showed hardness for the Approximate-Graph-Coloring$(3, 4)$ problem (that is there is no efficient algorithm which can find a 4-coloring, in any 3-colorable graph). Since then, there has been no progress in this problem. Later, results where proved using a complexity assumption called the Unique Games Conjecture (UGC), which is stronger than the P$\neq$ NP assumption. Starting with the work of Khot [Kho02c], it was shown that UGC, explains the lack of efficient approximation algorithms for a variety of problems (eg. Vertex Cover, MAX-CUT). Dinur, Mossel & Regev [DMR09] showed hardness for Approximate-

Graph-Coloring$(3, C)$ for any constant C using a conjecture similar to UGC. Due to a technical problem (that UGC does not have perfect completeness), their results which used UGC exactly, showed hardness for the ε-Almost-Coloring(c, C) problem, for any small $\varepsilon > 0$.

Definition 1.6. *The ε-Almost-Coloring(c, C) problem is of distiguishing graph from the following cases:*

- YES *: There is a subgraph of size $(1 - \varepsilon)n$ that is c-colorable.*
- NO *: Any independent set is the graph has size at most n/C.*

Contributions of this Thesis: The work of Dinur & Shinkar [DS10] implies hardness results for ε-Almost-Coloring$(3, \text{poly}(\log n))$, using a stronger form of UGC (where the dependence between the soundness and alphabet size is inverse polynomial). In Chapter 8, we show hardness for the same problem, using a weaker form of UGC (in which the aforesaid dependence is super-polynomial) in some respects (joint work with Dinur, Harsha & Srinivasan [DHSV15]).

The previous reductions (by Dinur, Mossel & Regev [DMR09] and Dinur & Shinkar [DS10]) followed the template of Håstad [Hås01], which employed a particular error correcting code known as the long code. As the name implies, this code has a large size which made the reductions inefficient. A shorter code called the low degree long code was proposed by Barak et al. [BGH$^+$12] . Dinur and Guruswami [DG14] showed improved approximate covering (which we define in Section 1.4.3) hardness results using this shorter code. In Chapter 8, we adapt this shorter code to the reduction of Dinur, Mossel and Regev [DMR09] for graph coloring, to get improved results.

1.4.2 Hypergraph Coloring

Guruswami, Håstad & Sudan [GHS02] initiated the study of hypergraph coloring problems, to get a better understanding of graph coloring and since it is a natural generalization. They showed hardness results for $c = 2, C = \text{poly}(\log \log n)$, for hypergraph coloring, assuming that NP does not have quasi-polynomial algorithms (such results are commonly known as *quasi*-NP-*hardness* results). A *k-uniform hypergraph* $G = (V, E)$ is similar to a graph, with the edges $E \subseteq \binom{V}{k}$ containing k vertices. A c-coloring of a hypergraph is coloring of vertices using colors $\{1, \cdots, c\}$, such that every edge has 2 vertices with distinct colors. For $k = 2$, a hypergraph is simply a graph.

Definition 1.7. *The* Approximate-k-Hypergraph-Coloring(c, C) *problem is defined similar to* Approximate-Graph-Coloring(c, C)*, as given a c-colorable k-uniform hypergraph, find a C-coloring. The decision version is also defined analogously.*

When $k > 2$, for any constant $c > 1$, known algorithms only guarantee an n^{α}-coloring for some $\alpha < 1$. Starting with the work of Guruswami, Håstad & Sudan [GHS02], there have been many results in hardness of hypergraph coloring. For the case of constant c, k, strongest known results due to Khot [Kho02a], who showed quasi-NP-hardness for $C = \text{poly}(\log n)$.

Contributions of this Thesis: In Chapter 9, we exponentially improve the hardness results. In Section 9.1, we first show hardness results (joint work with Guruswami, Harsha, Håstad & Srinivasan [GHH+14]) for 3-uniform 3-colorable hypergraphs by a more efficient reduction, that makes use of low degree long code (which we describe in Section 7.1). For the case of 4-uniform 4-colorable hypergraphs, our initial work (joint work with Guruswami, Harsha, Håstad & Srinivasan [GHH+14]) showed the first super-polylogarithmic coloring hardness (i.e. $C >> \text{poly}(\log n)$) results, by

using the low degree long code. Subsequent to our initial work Khot & Saket [KS14a] got hardness results of $2^{(\log n)^{1/21}}$, by using the low degree long code with degree 2. Though their result was for 12-uniform hypergraphs. We further observed [Var14] that by combining their methods with ours, the same hardness results can be obtained for 4-uniform hypergraphs. Hence we improved the hardness results from $\text{poly}(\log n)$ to $2^{(\log n)^{1/21}}$ for the case of 4-colorable 4-uniform hypergraphs.

1.4.3 Covering CSPs

The covering problem for constraint satisfaction (CSP) is a generalization of the hypergraph coloring problem, introduced by Guruswami, Håstad and Sudan [GHS02] and later studied in detail by Dinur & Kol [DK13]. An instance of the problem consists of a hypergraph $G = (V, E)$ along with a *predicate* $P \subseteq \{0,1\}^k$ and a *literal* function $L : E \to \{0,1\}^k$. An *assignment* $f : V \to \{0,1\}$ *covers* an edge $e \in E$, if $f|_e \oplus L(e) \in P$ (by $f|_e$, we mean the k bit string, obtained by restricting f to vertices in e, and the $\oplus$ operation is coordinate-wise parity of the two strings). A *cover* for a CSP instance is a set of assignments such that every edge is covered by one of the assignments. The goal of the covering problem is to find the minimum sized cover.

When the predicate is the 3-OR predicate, we can view the instance as a 3-SAT instance, where each edge is a clause and the literal function specifies which variables are to be negated. Then the satisfiability problem is equivalent to finding a cover of size 1. The covering problem can also be thought of as a generalization of the coloring problem. Consider an instance G with the NAE($:= \{0,1\}^k \setminus \{\bar{0}, \bar{1}\}$) predicate and the trivial literal function $L(e) = 0^k$ for every edge e. It is not difficult to see that G has a cover of size t iff G is 2^t-colorable. The *approximate covering* problem is defined as, given a c-coverable instance, find a C-covering.

Definition 1.8 (Covering-P-CSP(c, C)). *For $P \subseteq [q]^k$ and $c, C \in \mathbb{N}$, the* Cover-

ING-P-CSP(c, C) *problem is, given a* c*-coverable instance* $(G = (V, E), L)$ *of* P-CSP, *find an* C*-covering.*

A predicate P is *odd*, if for every $x \in \{0,1\}^k$ either $x \in P$ or $\overline{x} \in P$. For odd predicates, there is a trivial algorithm with factor 2, since any assignment and its complement covers the CSP instance. Dinur & Kol [DK13] asked the question whether, the approximate covering problem is hard for any constant $C > c$, for all non-odd predicates. Assuming a variant of UGC, they proceeded to show that if a non-odd predicate has a pairwise independent distribution in its support then, this is indeed the case.

CONTRIBUTIONS OF THIS THESIS: In Chapter 10, we answer the question of Dinur & Kol in the affirmative (joint work with Bhangale & Harsha [BHV15]). That is, the approximate covering problem for a non-odd predicate is hard for any constant $C > c$ (assuming the same conjecture as Dinur & Kol used). This leads to a complete characterization of predicates for which this result can be true, since there is a trivial 2-covering algorithm for odd predicates. Our results also holds over non binary alphabets. We also show NP-hardness results, for the approximate covering problem with parameters $c = 2, C = \log\log n$, for a class of predicates. Previously such results were known due to Dinur & Kol for 4-LIN with $c = 2, C = \log\log\log n$.

1.5 Organization of Thesis

This thesis has two main parts. In Part II, we will introduce some of the mathematical techniques used in proving the hardness results. Part III contains all the hardness results.

In Part II, we also prove some lemmas which form the basis of the results in the next part. In Chapter 3, Section 3.3 is a contribution of this thesis, were we prove analogues

of results in Boolean function analysis to functions on subspaces. In Chapter 4, we discuss a linear algebraic result about testing low degree polynomials. Section 4.3 is a contribution of this thesis, though the analysis is similar to the results of Dinur & Guruswami [DG14] mentioned in Section 4.2. In Chapter 5, we also give some combinatorial results about derandomized graph products. Section 5.2 and Section 5.3 are contributions of this thesis which uses the results proved in Section 4.3.

Part III contains all the hardness results about almost graph coloring (Chapter 8; joint work with Dinur, Harsha & Srinivasan [DHSV15]), hypergraph coloring (Chapter 9; joint work with Guruswami, Harsha, Håstad & Srinivasan [GHH+14] and the result [Var14]), and covering problem (Chapter 10; joint work with Bhangale & Harsha [BHV15]). These results are the main contributions of this thesis, though much of the hardness reductions and analysis are similar to previous works.

Part II

Mathematical Techniques

2

Analysis of Functions on Product Spaces

In this chapter, we will describe some results concerning functions on product probability spaces.

2.1 Preliminaries

Let (Ω, μ) be a discrete probability space and $(\Omega^L, \mu^{\otimes L})$ be the corresponding product space. For a function $f : \Omega^L \to \mathbb{R}$, the *Efron-Stein decomposition* of f with respect to the product space is given by

$$f(x_1, \cdots, x_L) = \sum_{\beta \subseteq [L]} f_\beta(x),$$

where f_β depends only on x_i for $i \in \beta$ and

$$\forall \beta' \not\supseteq \beta, a \in \Omega^{\beta'}, \mathop{\mathbb{E}}_{x \in \mu^{\otimes L}} [f_\beta(x) \mid x_{\beta'} = a] = 0.$$

The ℓ_p and ℓ_∞ norms of f with respect to the probability space are defined as

$$\|f\|_p := \mathop{\mathbb{E}}_{x \in \mu^{\otimes L}} [f(x)^p]^{1/p}, \qquad \|f\|_\infty := \max_{x \in \Omega^{\otimes L}} |f(x)|.$$

For $i \in [L]$, the *influence* of the ith coordinate on f is defined as follows.

$$\mathrm{Inf}_i[f] := \mathop{\mathbb{E}}_{x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_L} \mathrm{Var}_{x_i}[f(x_1,\cdots,x_L)] = \sum_{\beta:i\in\beta} \|f_\beta\|_2^2.$$

For an integer d, the *degree d influence* is defined as

$$\mathrm{Inf}_i^{\leq d}[f] := \sum_{\beta:i\in\beta,|\beta|\leq d} \|f_\beta\|_2^2.$$

2.2 Invariance Principle

Let (Ω^k,μ) be a probability space. Let $\mathrm{support}(\mu) := \{x \in \Omega^k \mid \mu(x) > 0\}$.

Definition 2.1 (Connected Sets and Distributions). *We say that $S \subseteq \Omega^k$ is* connected *if for every $x, y \in S$, there is a sequence of strings starting with x and ending with y such that every element in the sequence is in S and every two adjacent elements differ in exactly one coordinate. The probability space is connected if* $\mathrm{support}(\mu)$ *is connected.*

Theorem 2.2 (Mossel [Moso8, Proposition 6.4]). *Let (Ω^k,μ) be a connected probability space such the minimum probability of every atom in* $\mathrm{support}(\mu)$ *is at least $\alpha \in \left(0,\frac{1}{2}\right]$. Then there exists continuous functions $\overline{\Gamma} : (0,1) \to (0,1)$ and $\underline{\Gamma} : (0,1) \to (0,1)$ such that the following holds: For every $\varepsilon > 0$, there exists $\tau > 0$ and an integer d such that if a function $f : \Omega^L \to [0,1]$ satisfies*

$$\forall i \in [n], \mathrm{Inf}_i^{\leq d}(f) \leq \tau$$

then

$$\underline{\Gamma}\left(\mathop{\mathbb{E}}_{\mu}[f]\right) - \varepsilon \leq \mathop{\mathbb{E}}_{(x_1,\ldots,x_k)\sim\mu}\left[\prod_{j=1}^{k} f(x_j)\right] \leq \overline{\Gamma}\left(\mathop{\mathbb{E}}_{\mu}[f]\right) + \varepsilon.$$

There exists an absolute constant C such that one can take $\tau = \varepsilon^{C\frac{\log(1/\alpha)\log(1/\varepsilon)}{\varepsilon\alpha^2}}$ and

$d = \log(1/\tau)\log(1/\alpha)$.

Correlation is a measure of dependence in probability spaces where the sample space is a product set.

Definition 2.3 (Correlated Spaces). *Let $(\Omega_1 \times \Omega_2, \mu)$ be a finite probability space, the correlation between Ω_1 and Ω_2 with respect to μ us defined as*

$$\rho(\Omega_1, \Omega_2; \mu) := \max_{\substack{f:\Omega_1\to\mathbb{R},\mathbb{E}[f]=0,\mathbb{E}[f^2]\leq 1\\ g:\Omega_2\to\mathbb{R},\mathbb{E}[g]=0,\mathbb{E}[g^2]\leq 1}} \mathbb{E}_{(x,y)\sim\mu}[|f(x)g(y)|].$$

For a probability space $\left(\prod_{i=1}^k \Omega_i, \mu\right)$, the correlation is given by

$$\rho\left(\prod_{i=1}^k \Omega; \mu\right) := \max_{i\in[k]} \rho\left(\Omega_i, \prod_{j\in[k],j\neq i} \Omega_i; \mu\right).$$

The following result about correlated spaces is an adaptation of similar results (see Wenner [Wen13, Theorem 3.12] and Guruswami & Lee [GL15, Lemma A.1]) to proving our hardness results.

Theorem 2.4. *Let $(\Omega_1^k \times \Omega_2^k, \mu)$ be a correlated probability space with correlation $\rho < 1$ such that the marginal of μ on any pair of coordinates one each from Ω_1 and Ω_2 is a product distribution. Let μ_1, μ_2 be the marginals of μ on Ω_1^k and Ω_2^k respectively. Let X, Y be two random $k \times L$ dimensional matrices chosen as follows: independently for every $i \in [L]$, the pair of columns $(x^i, y^i) \in \Omega_1^k \times \Omega_2^k$ is chosen from μ. Let x_i, y_i denote the i^{th} rows of X and Y respectively. If $F : \Omega_1^L \to [-1, +1]$ and $G : \Omega_2^L \to [-1, +1]$ are functions such that*

$$\tau := \sqrt{\sum_{i\in[L]} \mathrm{Inf}_i[F]\cdot\mathrm{Inf}_i[G]} \text{ and } \Gamma := \max\left\{\sqrt{\sum_{i\in[L]}\mathrm{Inf}_i[F]}, \sqrt{\sum_{i\in[L]}\mathrm{Inf}_i[G]}\right\},$$

then

$$\left| \mathop{\mathbb{E}}_{(X,Y)\in\mu^{\otimes L}} \left[\prod_{i\in[k]} F(x_i)G(y_i) \right] - \mathop{\mathbb{E}}_{X\in\mu_1^{\otimes L}} \left[\prod_{i\in[k]} F(x_i) \right] \mathop{\mathbb{E}}_{Y\in\mu_2^{\otimes L}} \left[\prod_{i\in[k]} G(y_i) \right] \right| \leq 2^{O(k)}\Gamma\tau. \tag{2.2.1}$$

Proof. We will prove the theorem by using the hybrid argument. For $i \in [L+1]$, let $X^{(i)}, Y^{(i)}$ be distributed according to $(\mu_1 \otimes \mu_2)^{\otimes i} \otimes \mu^{\otimes L-i}$. Thus, $(X^{(0)}, Y^{(0)}) = (X, Y)$ is distributed according to $\mu^{\otimes L}$ while $(X^{(L)}, Y^{(L)})$ is distributed according to $(\mu_1 \otimes \mu_2)^{\otimes L}$. For $i \in [L]$, define

$$\mathrm{err}_i := \left| \mathop{\mathbb{E}}_{X^{(i)},Y^{(i)}} \left[\prod_{j=1}^{k} F(x_j^{(i)})G(y_j^{(i)}) \right] - \mathop{\mathbb{E}}_{X^{(i+1)},Y^{(i+1)}} \left[\prod_{j=1}^{k} F(x_j^{(i+1)})G(y_j^{(i+1)}) \right] \right|. \tag{2.2.2}$$

The left hand side of Equation (2.2.1) is upper bounded by $\sum_{i\in[L]} \mathrm{err}_i$. Now for a fixed i, we will bound err_i. We use the Efron-Stein decomposition of F, G to split them into two parts: the part which depends on the ith input and the part independent of the ith input.

$$F = F_0 + F_1 \text{ where } F_0 := \sum_{\alpha: i\notin\alpha} F_\alpha \text{ and } F_1 := \sum_{\alpha: i\in\alpha} F_\alpha.$$

$$G = G_0 + G_1 \text{ where } G_0 := \sum_{\beta: i\notin\beta} G_\beta \text{ and } G_1 := \sum_{\beta: i\in\beta} G_\beta.$$

Note that $\mathrm{Inf}_i[F] = \|F_1\|_2^2$ and $\mathrm{Inf}_i[G] = \|G_1\|_2^2$. Furthermore, the functions F_0 and F_1 are bounded since $F_0(x) = \mathbb{E}_{x'}[F(x')|x'_{[L]\setminus i} = x_{[L]\setminus i}] \in [-1, +1]$ and $F_1(x) = F(x) - F_0(x) \in [-2, +2]$. For $a \in \{0,1\}^k$, let $F_a(X) := \prod_{j=1}^k F_{a_j}(x_j)$. Similarly G_0, G_1 are bounded and G_a defined analogously. Substituting these definitions in Equation (2.2.2) and expanding the products gives

$$\mathrm{err}_i = \left| \sum_{a,b\in\{0,1\}^k} \left(\mathop{\mathbb{E}}_{X^{(i)},Y^{(i)}} \left[F_a(X^{(i)})G_b(Y^{(i)}) \right] - \mathop{\mathbb{E}}_{X^{(i+1)},Y^{(i+1)}} \left[F_a(X^{(i+1)})G_b(Y^{(i+1)}) \right] \right) \right|.$$

Since both the distributions are identical on $(\Omega_1^k)^{\otimes L}$ and $(\Omega_2^k)^{\otimes L}$, all terms with $a = \bar{0}$ or $b = \bar{0}$ are zero. Because μ is uniform on any pair of coordinates on each from the Ω_1 and Ω_2 sides, terms with $|a| = |b| = 1$ also evaluates to zero. Now consider the remaining terms with $|a|, |b| \geq 1, |a| + |b| > 2$. Consider one such term where $a_1, a_2 = 1$ and $b_1 = 1$. In this case, by Cauchy-Schwarz inequality we have that

$$\left| \mathop{\mathbb{E}}_{X^{(i-1)}, Y^{(i-1)}} \left[F_a(X^{(i-1)}) G_b(Y^{(i-1)}) \right] \right| \leq \sqrt{\mathbb{E}\, F_1(x_1)^2 G_1(y_1)^2} \cdot \|F_1\|_2 \cdot \left\| \prod_{j>2} F_{a_j} \right\|_\infty \cdot \left\| \prod_{j>1} G_{b_j} \right\|_\infty .$$

From the facts that the marginal of μ to any pair of coordinates one each from Ω_1 and Ω_2 sides are uniform, $\mathrm{Inf}_i[F] = \|F_1\|_2^2$ and $|F_0(x)|, |F_1(x)|, |G_0(x)|, |G_1(x)|$ are all bounded by 2, the right side of above becomes

$$\sqrt{\mathbb{E}\, F_1(x_1)^2} \sqrt{\mathbb{E}\, G_1(y_1)^2} \cdot \|F_1\|_2 \cdot \left\| \prod_{j>2} F_{a_j} \right\|_\infty \cdot \left\| \prod_{j>1} G_{b_j} \right\|_\infty \leq \sqrt{\mathrm{Inf}_i[F]^2 \mathrm{Inf}_i[G]} \cdot 2^{2k}.$$

All the other terms corresponding to other (a, b) which are at most 2^{2k} in number, are bounded analogously. Hence,

$$\begin{aligned} \sum_{i \in [L]} \mathrm{err}_i &\leq 2^{4k} \sum_{i \in [L]} \left(\sqrt{\mathrm{Inf}_i[F]^2 \mathrm{Inf}_i[G]} + \sqrt{\mathrm{Inf}_i[F] \mathrm{Inf}_i[G]^2} \right) \\ &= 2^{4k} \sum_{i \in [L]} \sqrt{\mathrm{Inf}_i[F] \mathrm{Inf}_i[G]} \left(\sqrt{\mathrm{Inf}_i[F]} + \sqrt{\mathrm{Inf}_i[G]} \right). \end{aligned}$$

By applying the Cauchy-Schwarz inequality, followed by a triangle inequality, we obtain

$$\sum_{i \in [L]} \mathrm{err}_i \leq 2^{4k} \sqrt{\sum_{i \in [L]} \mathrm{Inf}_i[F] \mathrm{Inf}_i[G]} \left(\sqrt{\sum_{i \in [L]} \mathrm{Inf}_i[F]} + \sqrt{\sum_{i \in [L]} \mathrm{Inf}_i[G]} \right).$$

Thus, proved. □

3

Harmonic Analysis of Functions

In this chapter, we describe some tools for analyzing functions over probability spaces, for which the sample space exhibits a field structure, and the distribution is defined in terms of the field operations. In this case, we can obtain decompositions with special properties, that helps in analysing such distributions. We will be concerned with the sample space which is the vector space of functions of the form $f : \mathbb{F}_p^R \to \mathbb{F}_p$ (p is a prime). We extend these decompositions and the associated results to the case when the domain is the subspace of low degree polynomials over $\mathbb{F}_p$. We prove an analogue of hypercontractivity for functions over this subspace (Lemma 3.18), by reducing it to the result on the full space (Lemma 3.22). This enables us to prove an analogue of a result by Alon et al. [ADFS04] (Lemma 3.6), to functions on the subspace (Lemma 3.19). We use these results later in Chapter 5, for proving some derandomized graph product results.

3.1 Harmonic Analysis for Fields

Consider the probability space $(\mathbb{F}_p, \mu)$ where μ is the uniform probability measure over $\mathbb{F}_p$. We will be working with functions of the form $A : \mathbb{F}_p^R \to \mathbb{C}$. Note that all such functions form a vector space over $\mathbb{C}$ with dimension p^R. Characters are a natural orthogonal basis for this vector space.

Definition 3.1 (Character). *A* character *of* $\mathbb{F}_p^R$ *is a function* $\chi : \mathbb{F}_p^R \to \mathbb{C}$ *such that*

$$\chi(0) = 1 \quad \textit{and} \quad \forall f, g \in \mathbb{F}_p^R,\ \chi(f+g) = \chi(f)\chi(g).$$

The following lists the basic properties of characters, which can be verified easily.

Observation 3.2. *Let* $\{1, \omega, \cdots, \omega^{p-1}\}$ *be the* p*th roots of unity and for* $\beta, f \in \mathbb{F}_p^R$*,*

$$\chi_\beta(f) := \omega^{\beta \cdot f} \quad \textit{where} \quad \beta \cdot f := \sum_{i=1}^{R} \beta_i f_i \mod p.$$

- *The characters of* $\mathbb{F}_p^R$ *are* $\{\chi_\beta : \beta \in \mathbb{F}_p^R\}$.
- *Characters forms an orthonormal basis for the vector space of functions from* $\mathbb{F}_p^R$ *to* $\mathbb{C}$*, under the inner product*

$$\langle A, B \rangle := \mathop{\mathbb{E}}_{f \in \mathbb{F}_p^R} \left[A(f)\overline{B(f)} \right].$$

- *Any function* $A : \mathbb{F}_p^R \to \mathbb{C}$ *can be uniquely decomposed as*

$$A(f) = \sum_{\beta \in \mathbb{F}_p^R} \widehat{A}(\beta)\chi_\beta(f) \quad \textit{where} \quad \widehat{A}(\beta) := \mathop{\mathbb{E}}_{g \in \mathbb{F}_p^R} \left[A(g)\overline{\chi_\beta(g)} \right].$$

- *For any function* $A : \mathbb{F}_p^R \to \mathbb{C}$*,*

$$\sum_{\beta \in \mathbb{F}_p^R} |\widehat{A}(\beta)|^2 = \mathop{\mathbb{E}}_{f \in \mathbb{F}_p^R} \left[|A(f)|^2 \right]. \tag{3.1.1}$$

- *For any function* $A : \mathbb{F}_p^R \to \{1, \omega, \cdots, \omega^{p-1}\}$*,*

$$\sum_{\beta \in \mathbb{F}_p^R} |\widehat{A}(\beta)|^2 = 1. \tag{3.1.2}$$

Remark 3.3. *Note that when $p = 2$, the roots of unity are $\{-1, +1\}$. Then the characters are real valued functions. It is easy to see that they form a orthonormal basis for the real vector space of functions of the form $A : \mathbb{F}_2^R \to \mathbb{R}$. Hence all the above properties hold with respect to this vector space as well.*

Definition 3.4 (Generalized Dictator). *A* generalized dictator *function $A : \mathbb{F}_p^R \to \mathbb{C}$ is one that depends only on a single coordinate. That is, $\exists i \in [R]$ such that for any $\beta \in \mathbb{F}_p^R$, if it has a non-zero entry in some coordinate $j \neq i$ then $\widehat{A}(\beta) = 0$.*

Definition 3.5 (Fourier degree). *The* Fourier degree *of a function $A : \mathbb{F}_p^R \to \mathbb{C}$ is the smallest number d such that A can be written as*

$$A = \sum_{\beta : |\beta| \leq d} \widehat{A}(\beta)\chi_\beta$$

where $|\beta|$ is number of coordinates i where $\beta_i \neq 0$.

An interesting fact about functions of the form $A : \mathbb{F}_2^R \to \{0, 1\}$ is that, if A has Fourier degree 1 then A is a generalized dictator function. The following theorem over $\mathbb{F}_p$, says that if the sum of squares of absolute values of Fourier coefficients of A with $|\beta| > 1$ is small, then it is close to a generalized dictator function. It is a generalization of the well known FKN Theorem (see [FKN02]), which gives the result for $p = 2$.

Lemma 3.6 (Alon et al. [ADFS04]). *For every prime p, there is constant K (that depends on p) such that the following holds: If $A : \mathbb{F}_p^R \to \{0, 1\}$ satisfies*

$$\sum_{|\alpha|>1} |\widehat{A}(\alpha)|^2 \leq \varepsilon \text{ and } \widehat{A}(0) = \delta$$

then there exists a generalized dictator $B : \mathbb{F}_p^R \to \{0, 1\}$ such that

$$\|A - B\|_2 \leq \frac{K\varepsilon}{\delta - \delta^2 - \varepsilon}.$$

The above lemma is proved using the following hypercontractive inequality.

Lemma 3.7 (Hypercontractivity). *For every prime p, there is a constant C (that depends on p) such that for any function $A : \mathbb{F}_p^R \to \mathbb{C}$ with $\widehat{A}(\alpha) = 0$ when $|\alpha| > t$,*

$$\|A\|_4 \leq C^t \|A\|_2.$$

In the next section we will prove derandomized versions of the above lemmas.

3.2 Polynomial Subspaces

Let $\mathsf{P}_{r,d}$ be the set of degree d polynomials on r variables over $\mathbb{F}_p$, with individual degrees $< p$ (the prime p will be clear from the context). Let $\mathfrak{F}_r := \mathsf{P}_{r,(p-1)r}$. Note that $\mathfrak{F}_r$ is the set of all functions from $\mathbb{F}_p^r$ to $\mathbb{F}_p$. $\mathfrak{F}_r$ is a $\mathbb{F}_p$-vector space of dimension p^r and $\mathsf{P}_{r,d}$ is its subspace of dimension $r^{O(d)}$. The *Hamming distance* between f and $g \in \mathfrak{F}_r$, denoted by $\Delta(f, g)$, is the number of inputs on which f and g differ. When $S \subseteq \mathfrak{F}_r$, $\Delta(f, S) := \min_{g \in S} \Delta(f, g)$. We say f is Δ-far from S if $\Delta(f, S) \geq \Delta$ and f is Δ-close to S otherwise. For a polynomial $\alpha \in \mathsf{P}_{r,d}$, the support size of the polynomial is $|\alpha| := |\{x : \alpha(x) \neq 0\}|$. Given $f, g, \in \mathfrak{F}_r$, the *dot product* between them is defined as

$$f \cdot g := \sum_{x \in \mathbb{F}_p^r} f(x)g(x) \mod p.$$

For a subspace $S \subseteq \mathfrak{F}_r$, the *dual subspace* is defined as

$$S^\perp := \{g \in \mathfrak{F}_r : \forall f \in S, g \cdot f = 0\}.$$

The following theorem relating dual spaces is well known.

Lemma 3.8. $\mathsf{P}_{r,d}^\perp = \mathsf{P}_{r,(p-1)r-d-1}$

Proof. First note that the dimensions of the two subspaces are equal by a counting argument. Next we show that $\mathsf{P}_{r,d}^{\perp} \supseteq \mathsf{P}_{r,(p-1)r-d-1}$. We just need to show that for any monomial of degree $(p-1)r-d-1$ with individual degrees $< p$, the dot product with any monomial of degree d with individual degrees $< p$ is 0. The product of any such pair of monomials is a monomial with total degree at most $(p-1)r-1$, and hence has a variable with degree $< p-1$. Without loss of generality, let this variable be x_1 with degree $t < p-1$. Notice that $\sum_{x_1 \in \mathbb{F}_p} x_1^t = 0 \mod p$ and hence the dot product is 0. □

We need the following Schwartz-Zippel-like Lemma for degree d polynomials.

Lemma 3.9 (Schwartz-Zippel lemma [HSS13, Lemma 3.2]). *Let $f \in \mathbb{F}_p[x_1, \cdots, x_r]$ be a non-zero polynomial of degree at most d with individual degrees at most $p-1$. Then the support size ($|f| := |\{x : f(x) \neq 0\}|$) satisfies*

$$|f| \geq p^{r-d/(p-1)}.$$

The following lemma is an easy consequence of Lemma 3.9.

Lemma 3.10. *Let g be a uniformly random polynomial from $\mathsf{P}_{r,d}$. Then its truth table as a random string of length p^r over the alphabet $\mathbb{F}_p$, is $p^{\lfloor (d+1)/(p-1) \rfloor} - 1$-wise independent.*

Proof. From Lemma 3.9 and Lemma 3.8, we know that any non-zero polynomial in $\mathsf{P}_{r,d}^{\perp} = \mathsf{P}_{r,(p-1)r-d-1}$ has support size at least $p^{\lfloor (d+1)/(p-1) \rfloor}$. Suppose there is a subset S of size $p^{\lfloor (d+1)/(p-1) \rfloor} - 1$, where g is not uniform. Let $V \subseteq \mathbb{F}_p^{|S|}$ be the set of restrictions of truth tables of polynomials in $\mathsf{P}_{r,d}$ to S. Note that V is a subspace. Since the truth table of g restricted to S is not uniformly distributed, the dimension of V is $< |S|$. Then $V^{\perp} \subseteq \mathbb{F}_p^{|S|}$ is non-empty. Consider a non-zero $v \in V^{\perp}$. Then the function

f which is zero outside S and $f_{|S} = v$ corresponds to a non-zero polynomial which belongs to $\mathsf{P}^{\perp}_{r,d} = \mathsf{P}_{r,(p-1)r-d-1}$ with support $< p^{\lfloor (d+1)/(p-1) \rfloor}$ which is a contradiction.

□

The following lemma is an easy consequence of Lemma 3.10

Lemma 3.11. *Let $d > 1$, X be a set of $p^d - 1$ points in $\mathbb{F}_p^r$ and $f : X \to \mathbb{F}_p$ an arbitrary function. Then there exists a polynomial q of degree at most $(p-1)d$ such that q agrees with f on all points in X.*

Proof. By Lemma 3.10, the truth table of a random polynomial g of degree $(p-1)d$ is $p^d - 1$-wise independent. Hence $g_{|X} = f$ with non-zero probability. □

3.3 Harmonic Analysis for Polynomial Subspaces

We define a orthonormal basis set of characters for the vector space of functions of the form $A : \mathsf{P}_{r,d} \to \mathbb{C}$.

Definition 3.12 (Character). *A* character *of $\mathsf{P}_{r,d}$ is a function $\chi : \mathsf{P}_{r,d} \to \mathbb{C}$ such that*

$$\chi(0) = 1 \text{ and } \forall f, g \in \mathsf{P}_{r,d},\ \chi(f+g) = \chi(f)\chi(g).$$

The following lists the basic properties of characters (similar to Observation 3.2).

Observation 3.13 ([DG14, Section II C]). *Let $\{1, \omega, \cdots, \omega^{p-1}\}$ be the pth roots of unity and for $\beta \in \mathfrak{F}_r$, $f \in \mathsf{P}_{r,d}$,*

$$\chi_\beta(f) := \omega^{\beta \cdot f} \quad \text{where} \quad \beta \cdot f := \sum_{x \in \mathbb{F}_p^r} \beta(x) f(x) \mod p.$$

- *The characters of $\mathsf{P}_{r,d}$ are $\{\chi_\beta : \beta \in \mathfrak{F}_r\}$.*

- *For any* $\beta, \beta' \in \mathfrak{F}_r$, $\chi_\beta = \chi'_\beta$ *if and only if* $\beta - \beta' \in \mathsf{P}^\perp_{r,d}$.

- *For* $\beta \in \mathsf{P}^\perp_{r,d}$, χ_β *is the constant* 1 *function.*

- $\forall \beta, \exists \beta'$ *such that* $\beta - \beta' \in \mathsf{P}^\perp_{r,d}$ *and* $|\beta'| = \Delta(\beta, \mathsf{P}^\perp_{r,d})$ *(i.e., the constant* 0 *function is (one of) the closest function to* β' *in* $\mathsf{P}^\perp_{r,d}$*). We call such a* β' *a minimum support function for the coset* $\beta + \mathsf{P}^\perp_{r,d}$.

- *Characters forms an orthonormal basis for the vector space of functions from* $\mathsf{P}_{r,d}$ *to* $\mathbb{C}$*, under the inner product*

$$\langle A, B \rangle := \mathop{\mathbb{E}}_{f \in \mathsf{P}_{r,d}} \left[A(f) \overline{B(f)} \right].$$

- *Any function* $A : \mathsf{P}_{r,d} \to \mathbb{C}$ *can be uniquely decomposed as*

$$A(f) = \sum_{\beta \in \Lambda_{r,d}} \widehat{A}(\beta) \chi_\beta(f) \quad \textit{where} \quad \widehat{A}(\beta) := \mathop{\mathbb{E}}_{g \in \mathsf{P}_{r,d}} \left[A(g) \overline{\chi_\beta(g)} \right]$$

  and $\Lambda_{r,d}$ *is the set of minimum support functions, one for each of the cosets in* $\mathfrak{F}_r / \mathsf{P}^\perp_{r,d}$*, with ties broken arbitrarily.*

- *For any function* $A : \mathsf{P}_{r,d} \to \mathbb{C}$*,*

$$\sum_{\beta \in \Lambda_{r,d}} \left| \widehat{A}(\beta) \right|^2 = \mathop{\mathbb{E}}_{f \in \mathsf{P}_{r,d}} \left[|A(f)|^2 \right].$$

- *For any function* $A : \mathsf{P}_{r,d} \to \{1, \omega, \cdots, \omega^{p-1}\}$*,*

$$\sum_{\beta \in \Lambda_{r,d}} |\widehat{A}(\beta)|^2 = 1.$$

The following lemma relates characters over different domains related by co-ordinate projections.

Lemma 3.14. *Let $m \leq r$ and $\pi : \mathbb{F}_p^r \to \mathbb{F}_p^m$ be a (co-ordinate) projection i.e., there exist indices $1 \leq i_1 < \cdots < i_m \leq r$ such that $\pi(x_1, \ldots, x_r) = (x_{i_1}, \cdots, x_{i_m})$. Then for $f \in \mathsf{P}_{m,d},\ \beta \in \mathsf{P}_{r,d}$,*

$$\chi_\beta(f \circ \pi) = \chi_{\pi_p(\beta)}(f),$$

where $\pi_p(\beta)(y) := \sum_{x \in \pi^{-1}(y)} \beta(x)$.

Proof. Without loss of generality, let $\{i_1, \cdots, i_m\} = \{1, \cdots, m\}$. Then

$$\begin{aligned}
\chi_\beta(f \circ \pi) &= \omega^{\sum_{x \in \mathbb{F}_p^r} f \circ \pi(x) \cdot \beta(x)} \\
&= \omega^{\sum_{(x_1, \cdots, x_m) \in \mathbb{F}_p^m} f(x_1, \cdots, x_m) \cdot (\sum_{(x_{m+1}, \cdots, x_n)} \beta(x))} \\
&= \chi_{\pi_p(\beta)}(f)
\end{aligned}$$

□

Influence and generalized dictators can be defined for functions on polynomial subspaces similar to the product setting.

Definition 3.15 (Influence). *For a function $A : \mathsf{P}_{r,d} \to \mathbb{C}$ and a number $k < p^{\lfloor (d+1)/(p-1) \rfloor}/2$, the degree k influence of $a \in \mathbb{F}_p^r$ is defined as*

$$\mathrm{Inf}_a^{\leq k}(A) = \sum_{\beta \in \Lambda_{r,d} : \beta(a) \neq 0 \text{ and } |\beta| \leq k} |\widehat{A}(\beta)|^2.$$

Definition 3.16 (Generalized Dictator). *A function $A : \mathsf{P}_{r,d} \to \mathbb{C}$ is a generalized dictator if there exists $x \in \mathbb{F}_p^r$ and $\widehat{A}_0, \widehat{A}_1, \cdots, \widehat{A}_{p-1} \in \mathbb{C}$ such that A can be written as $A = \widehat{A}_0 + \sum_{i=1}^{p-1} \widehat{A}_i \chi_{ie_x}$ where $e_x : \mathbb{F}_p^r \to \mathbb{F}_p$ the indicator function for x.*

Lemma 3.17. *Let $A : \mathsf{P}_{r,d} \to \{0, 1\}$ be such that all non-zero Fourier coefficients have support size ≤ 1. Then A is a generalized dicator.*

Proof. The proof is similar to the proof of [ADFS04, Lemma 2.3]. Consider the function $(A(f))^2$. Since A is $\{0,1\}$ valued $(A(f))^2 = A(f)$. Equation the Fourier coefficients on both sides will give that there is an $x \in \mathbb{F}_p^r$ such that $A(f)$ only depends on $f(x)$. □

We prove an analogue of Theorem 3.7, to functions over polynomial subspaces.

Lemma 3.18. *For every prime p, there is a constant C such that for $4t \leq p^{d-1}$ and any function $A : \mathsf{P}_{r,(p-1)d} \to \mathbb{C}$ with $\widehat{A}_\alpha = 0$ when $|\alpha| > t$,*

$$\|A\|_4 \leq C^t \|A\|_2.$$

Proof. Follows from Lemma 3.22 and Lemma 3.7. □

We prove an analogue of Theorem 3.6, to functions over polynomial subspaces.

Lemma 3.19. *For every prime p, there is a constant K such that the following holds: If $A : \mathsf{P}_{r,(p-1)d} \to \{0,1\}$ satisfies*

$$\sum_{|\alpha|>1} |\widehat{A}(\alpha)|^2 \leq \varepsilon \text{ and } \widehat{A}(0) = \delta$$

then there exists a generalized dictator $B : \mathsf{P}_{r,(p-1)d} \to \{0,1\}$ such that

$$\|A - B\|_2^2 \leq \frac{K\varepsilon}{\delta - \delta^2 - \varepsilon}.$$

Proof. The proof of the lemma is similar to the proof of [ADFS04, Lemma 2.4]. Let $K = 2 + 32C^8$ where C is the constant from Lemma 3.18. First if $\varepsilon \geq \frac{1}{32C^8}$, then the lemma is true. This is because, for any $B : \mathsf{P}_{r,(p-1)d} \to \{0,1\}$, $A - B$ is a $\{-1,0,1\}$ valued function and $\|A - B\|_2^2 \leq 1$.

Now assume $\varepsilon < \frac{1}{32C^8}$. Let

$$A_S = \sum_{|\alpha|\leq 1} \widehat{A}(\alpha)\chi_\alpha \text{ and } A_L = \sum_{|\alpha|>1} \widehat{A}(\alpha)\chi_\alpha.$$

(S, L stands for small and large). If A_S were Boolean, it has to be a dictator by Lemma 3.17. Then the lemma follows by taking $B = A_S$. Consider the following function which measures the farness of A_S from being Boolean (it is identically 0, for Boolean functions)

$$H := A_S^2 - A_S.$$

Since A_S does not have any Fourier coeffcients with support > 1, H will have only Fourier coefficients with $|\alpha$ to be $0, 1$ and 2. Let e_x be the function with $e_x(x) = 1$ and 0 otherwise. Then for $a, b \in F_p, x, y \in \mathbb{F}_p^r$

$$\widehat{H}(ae_x + be_y) = 2\widehat{A}(ae_x)\widehat{A}(be_y).$$

The following claim says that the norm of H is small.

Claim 3.20.

$$\|H\|^2 \leq 32C^8\varepsilon.$$

The claim is proved later. Let $a_x := \sum_{i\in\mathbb{F}_p} |\widehat{A}(ie_x)|$. Note that

$$\sum_{x,y\in\mathbb{F}_p^r, x\neq y} a_x a_y \leq \frac{\|H\|^2}{4} \leq 8C^8. \tag{3.3.1}$$

Also from assumptions in the claim,

$$\sum_{x\in\mathbb{F}_p^r} a_x = \delta - \delta^2 - \varepsilon. \tag{3.3.2}$$

Let y be such that a_y is maximal. Then

$$(\sum_x a_x)^2 \leq \sum_x a_x^2 + 16C^8 \leq a_y \sum_x a_x + 16C^8.$$

This gives that $a_y \geq \delta - \delta^2 - \varepsilon(1 + 16C^8/(\delta - \delta^2 - \varepsilon))$. If $B' := \widehat{A}(0) + \sum_{i \in \mathbb{F}_p} \widehat{A}(ie_y)\chi_{ie_y}$, then we have that $\|A - B'\|_2^2 \leq \varepsilon(1 + 16C^8/(\delta - \delta^2 - \varepsilon))$. Now rounding B to the closest $[0, 1]$ valued function B' pointwise, we get that

$$\|A - B\|_2^2 \leq 2\|A - B'\|_2^2 \leq \frac{K\varepsilon}{\delta - \delta^2 - \varepsilon}.$$

Proof of Claim 3.20 . First notice,

$$H = A_S^2 - A_S = (A - A_L)^2 - (A - A_L) = A_L^2 + A_L(1 - 2A).$$

Let $k = 2C^4$ and $Z = \{f : |A_L(f) \leq k\sqrt{\varepsilon}\}$. Since $\|A_L\|_2^2 \leq \varepsilon$, by a Markov argument, $\Pr_f[Z] \geq 1 - 1/k^2$. Also for every $f \in Z$, $|H(f)| \leq 2|A_L(f)| \leq 2k\sqrt{\varepsilon}$. Since H has only Fourier coefficients with support size $0, 1, 2$, we can use Lemma 3.18 with $t = 2$. The claim follows from the following

$$\begin{aligned}
\|H\|_2^2 = \mathop{\mathbb{E}}_f |H(f)|^2 &= \Pr[Z] \mathop{\mathbb{E}}_{f \in Z} |H(f)|^2 + (1 - \Pr[Z]) \mathop{\mathbb{E}}_{f \notin Z} |H(f)|^2 \\
&\leq 4k^2\varepsilon + \frac{1}{k^2}\sqrt{\mathop{\mathbb{E}}_{f \notin Z} |H(f)|^4} \\
&\leq 4k^2\varepsilon + \frac{1}{k}\sqrt{\mathop{\mathbb{E}}_f |H(f)|^4} \\
&\leq 4k^2\varepsilon + \frac{1}{k}C^4\|H\|_2^2 \leq 32C^8\varepsilon
\end{aligned}$$

□

□

Definition 3.21 (Lift). *For a function $B : \mathsf{P}_{r,d} \to \mathbb{C}$ with the Fourier decomposition $B = \sum_{\alpha \in \Lambda_{r,d}} \widehat{B}(\alpha)\chi_\alpha$, the lift of B denoted by B' is a function $B' : \mathcal{F}_r \to \mathbb{C}$ with the Fourier decomposition $B' = \sum_{\alpha \in \Lambda_{r,d}} \widehat{B}(\alpha)\chi_\alpha$. In the decomposition of B', χ_α's are functions with domain $\mathcal{F}_r$.*

Lemma 3.22. *If $2kt \leq p^{d-1}$ and $B : \mathsf{P}_{r,(p-1)d} \to \mathbb{C}$ be a function such that $\widehat{B}(\alpha) = 0$ when $|\alpha| > t$ then*

$$\|B\|_{2k} = \|B'\|_{2k}.$$

Proof. From the Lemma 3.9 and Lemma 3.8, we have that $\forall \alpha \in \mathsf{P}^{\perp}_{r,(p-1)d} \setminus \{0\},\ |\alpha| > p^{d-1}$. So if $\exists \{\alpha_i, \beta_i\}_{i\in[k]}$ with $|\alpha_i|, |\beta_i| \leq t$, then

$$\sum_{i\in[k]} \alpha_i - \beta_i \in \mathsf{P}^{\perp}_{r,(p-1)d} \Rightarrow \sum_{i\in[k]} \alpha_i - \beta_i = 0. \tag{3.3.3}$$

This is because $\sum_{i\in[t]} \alpha_i - \beta_i$ has support size at most $2kt < p^{d-1}$. We use this fact to prove the theorem as follows:

$$\begin{aligned}
\|B\|_{2k}^{2k} &= \mathop{\mathbb{E}}_{f\in\mathsf{P}_{r,(p-1)d}} |B(f)|^{2k} = \mathop{\mathbb{E}}_{f\in\mathsf{P}_{r,(p-1)d}} \prod_{i\in[k]} B(f)\overline{B(f)} \\
&= \sum_{\alpha_1,\beta_1,\cdots,\alpha_k,\beta_k \in \Lambda_{r,(p-1)d}} \left(\prod_{i\in[k]} \widehat{B}_{\alpha_i} \overline{\widehat{B}_{\beta_i}} \right) \mathop{\mathbb{E}}_{f\in\mathsf{P}_{r,(p-1)d}} \prod_{i\in[k]} \chi_{\alpha_i}(f)\overline{\chi_{\beta_i}(f)} \\
&= \sum_{\substack{\alpha_1,\beta_1,\cdots,\alpha_k,\beta_k \in \Lambda_{r,(p-1)d} \\ \sum_i \alpha_i - \beta_i \in \mathsf{P}^{\perp}_{r,(p-1)d}}} \prod_{i\in[k]} \widehat{B}_{\alpha_i} \overline{\widehat{B}_{\beta_i}} \\
&= \sum_{\substack{\alpha_1,\beta_1,\cdots,\alpha_k,\beta_k \in \Lambda_{r,(p-1)d} \\ \sum_i \alpha_i - \beta_i = 0}} \prod_{i\in[k]} \widehat{B}_{\alpha_i} \overline{\widehat{B}_{\beta_i}} \ \text{(from (3.3.3))} \\
&= \sum_{\alpha_1,\beta_1,\cdots,\alpha_k,\beta_k \in \Lambda_{r,(p-1)d}} \left(\prod_{i\in[k]} \widehat{B}_{\alpha_i} \overline{\widehat{B}_{\beta_i}} \right) \mathop{\mathbb{E}}_{f\in\mathcal{F}_r} \prod_{i\in[k]} \chi_{\alpha_i}(f)\overline{\chi_{\beta_i}(f)} \\
&= \mathop{\mathbb{E}}_{f\in\mathcal{F}_r} \prod_{i\in[k]} B'(f)\overline{B'(f)} = \mathop{\mathbb{E}}_{f\in\mathcal{F}_r} |B'(f)|^{2k} = \|B'\|_{2k}^{2k}
\end{aligned}$$

□

3.3.1 Folding over Subspace

Definition 3.23 (Folded function over a subspace). *For any set S, a function $A : \mathsf{P}_{r,(p-1)d} \to S$ is said to be folded over a subspace $J \subseteq \mathsf{P}_{r,(p-1)d}$ if A is constant over cosets of J in $\mathsf{P}_{r,(p-1)d}$.*

Fact 1. *Given a function $A : \mathsf{P}_{r,(p-1)d}/J \to S$ there is a unique function $A' : \mathsf{P}_{r,(p-1)d} \to S$ that is folded over J such that for $g \in \mathsf{P}_{r,(p-1)d}$, $A'(g) = A(g + J)$.*

Given $q_1, \cdots, q_k \in \mathsf{P}_{r,3(p-1)}$, let

$$J(q_1, \ldots, q_k) := \left\{ \sum_i r_i q_i : r_i \in \mathsf{P}_{r,(p-1)(d-3)} \right\}.$$

The following lemma shows that if a function is folded over $J = J(q_1, \ldots, q_k)$, then it cannot have weight on small support characters that are non-zero on J (this is a generalization of the corresponding lemma by Dinur & Guruswami [DG14] to arbitrary fields).

Lemma 3.24. *Let $\beta \in \mathfrak{F}_r$ is such that $|\operatorname{support}(\beta)| < p^{d-3}$, and there exists $x \in \operatorname{support}(\beta)$ with $q_i(x) \neq 0$ for some i. Then if $A : \mathsf{P}_{r,d} \to \mathbb{C}$ is folded over $J = J(q_1, \ldots, q_k)$, then $\widehat{A}(\beta) = 0$.*

Proof. Construct a polynomial t which is zero at all points in support of β except at x. From Lemma 3.11, its possible to construct such a polynomial of degree at most $(p -$

$1)(d-3)$. Then we have that $tq_i \in J$ and $\langle \beta, tq_i \rangle \neq 0$. Now

$$\begin{aligned}
\mathop{\mathbb{E}}_{h}[A(h)\chi_\beta(h)] &= \frac{1}{p}\mathop{\mathbb{E}}_{h}[A(h)\chi_\beta(h) + A(h+tq_i)\chi_\beta(h+tq_i) + \cdots \\
&\qquad + A(h+(p-1)tq_i)\chi_\beta(h+(p-1)tq_i)] \\
&= \frac{1}{p}\mathop{\mathbb{E}}_{h}[A(h)\chi_\beta(h) + A(h)\chi_\beta(h+tq_i) + \cdots \\
&\qquad + A(h)\chi_\beta(h+(p-1)tq_i)] \\
&= \frac{1}{p}\mathop{\mathbb{E}}_{h}[A(h)\chi_\beta(h)(1+\chi_\beta(tq_i) + \cdots + \chi_\beta((p-1)tq_i))] \\
&= \frac{1}{p}\mathop{\mathbb{E}}_{h}[A(h)\chi_\beta(h)(1+\omega^{t(\beta\cdot q_i)} + \cdots + \omega^{(p-1)t(\beta\cdot q_i)})] = 0 \qquad \square
\end{aligned}$$

The last step is due to the fact that the sum $(1+\omega+\cdots+\omega^{p-1}) = 0$. Since $t(\beta\cdot q_i) \neq 0$, the previous equation contains this sum.

4

Testing of Low Degree Polynomials

In this chapter, we will describe some results about testing of low degree polynomials. These developments form the technical basis of our improvement to the hardness of coloring problems. The result in Section 4.1 is due to Haramaty et al. [HSS13] and Section 4.2 due to Dinur & Guruswami [DG14]. Section 4.3 is a contribution of this thesis and forms the basis for the combinatorial results in Chapter 5. It was first proved for showing the hardness of finding independent sets in 3-uniform 3-colorable hypergraphs (see Section 9.1).

We will be working with the field $\mathbb{F}_p$ with p being 2 or 3. Recall that $\mathsf{P}_{r,d}$ is the set of degree d polynomials on r variables with individual degrees $\leq p-1$. A test for $\mathsf{P}_{r,d}$ is an algorithm which has oracle access to the truth table of an input polynomial $f : \mathbb{F}_p^r \to \mathbb{F}_p$ and satisfies the following properties:

- Completeness : If $f \in \mathsf{P}_{r,d}$ then the algorithm accepts with probability 1.
- Soundness : For $f \notin \mathsf{P}_{r,d}$ with Δ being the distance of f from $\mathsf{P}_{r,d}$,

$$\Pr[\text{Test accepts}] \leq 1 - \Omega\left(\frac{\Delta}{p^r}\right).$$

Remark 4.1. *In common literature, testing also requires that the number of queries to be bounded. However for our purposes of designing hardness of approximation reductions,*

this is not a restriction.

4.1 Affine Subspace Test

In this section, we will be working over the field $\mathbb{F}_3$. Haramaty et al. [HSS13] analyzes the following test for checking whether a polynomial is of degree $\leq 2r - 2d - 1$. We take degree to be $2r - 2d - 1$, as it is convenient later in Chapter 5 and Chapter 8, were we use it.

Affine Subspace Test(f):

- Choose a random affine subspace S of dimension $r - d$.
- Check if the input f is of degree $2r - 2d - 1$ on S.

The test clearly satisfies completeness, since if the input f is of degree $\leq 2r - 2d - 1$, then its degree remains $\leq 2r - 2d - 1$ on any subspace. Note that $f|_S$ is of degree $\leq 2r - 2d - 1$ iff

$$\sum_{x \in \mathbb{F}_3^r} \mathbb{1}[x \in S] f(x) = \sum_{x \in S} f(x) = 0.$$

Since S is of dimension $r - d$, there exists linearly independent $\ell_1, \cdots \ell_d \in \mathsf{P}_{r,1}$ such that $\mathbb{1}[x \in S] = 2^d \times \prod_{i=1}^{d} (\ell_i - 1)(\ell_i - 2)$. Haramaty et al. [HSS13] proved the following lemma.

Lemma 4.2 (Haramaty et al. [HSS13, Theorem 1.3]). *There exists constants C_1, C_2 such that for any $\alpha \in \mathcal{F}_r$,*

$$\Pr[\text{Test accepts}] = \Pr_{\ell_i}\left[\langle \alpha, \prod_{i=1}^{d} (\ell_i - 1)(\ell_i - 2)\rangle = 0\right] \leq \max\left\{1 - \frac{C_1 \Delta(\alpha, \mathsf{P}_{r,2r-2d-1})}{3^d}, C_2\right\}$$

where $\ell_1, \cdots, \ell_i \in \mathsf{P}_{r,1}$ are random linearly independent polynomials.

4.2 Product Test

In this section we will be working over the field $\mathbb{F}_2$. Dinur & Guruswami [DG14] considered the following test for checking whether $f \in \mathsf{P}_{r,r-d-1}$.

PRODUCT TEST(f):

- Choose a random $g \in \mathsf{P}_{r,d/4}, h \in \mathsf{P}_{r,3d/4}$.
- Accept iff $\langle g \times h, f \rangle = \sum_{x \in \mathbb{F}_2^r} g(x) \cdot h(x) \cdot f(x) = 0$.

If $f \in \mathsf{P}_{r,r-d-1}$ then $g \times h \times f \in \mathsf{P}_{r,r-1}$ and hence $\sum_{x \in \mathbb{F}_2^r} (g \times h \times f)(x) = 0$. For analyzing the soundness, we consider two cases. Suppose the distance of f from $\mathsf{P}_{r,r-d-1}$ denoted by Δ is $< 2^{d/2}$. We can assume without loss of generality that $f = f' + \alpha$ where $f' \in \mathsf{P}_{r,r-d-1}$ and α has support size Δ. Note that the acceptance probability for f and α is the same. Since a uniformly random $g \in \mathsf{P}_{r,d/2}$ is $2^{d/2}$-wise independent as a string in 2^r, $\langle g \times h, \alpha \rangle$ is a uniformly random bit. Note that if p_{acc} is the acceptance probability of the test, then $\mathbb{E}\, \chi_\gamma(g \times h) = 2p_{\text{acc}} - 1$. Hence $p_{\text{acc}} = 1/2$. For the case when $\Delta \geq 2^{d/2}$, Dinur & Guruswami proved the following theorem, which implies that $p_{\text{acc}} \leq 1/2 + 2^{-4 \cdot 2^{d/4}}$.

Theorem 4.3 (Dinur & Guruswami, [DG14, Theorem 1]). *If $\gamma \in \mathfrak{F}_r$ has distance from $\mathsf{P}_{r,d}^{\perp} = \mathsf{P}_{r,r-d-1}$ at least $2^{d/2}$, then*

$$\mathop{\mathbb{E}}_{g \in \mathsf{P}_{r,d/4}} \left[\left| \mathop{\mathbb{E}}_{h \in \mathsf{P}_{r,3d/4}} [\chi_\gamma(g \times h)] \right| \right] \leq 2^{-4 \cdot 2^{d/4}}.$$

4.3 Square Test

In this section, we consider the following test for degree $2r - 2d - 1$ polynomials over $\mathbb{F}_3$.

SQUARE TEST(f):

- Choose a random $g \in \mathsf{P}_{r,d}$.
- Accept iff $\langle g^2, f \rangle = \sum_{x \in \mathbb{F}_2^r} (g \times g \times f)(x) = 0$.

As seen in the previous tests, the completeness condition is clearly satisfied. For the soundness, we analyze the quantity

$$\langle \beta, g^2 \rangle,$$

where $g \in \mathsf{P}_{r,d}$ is chosen uniformly at random and $\beta : \mathbb{F}_3^r \to \mathbb{F}_3$ is a fixed function having distance exactly Δ from $\mathsf{P}_{r,2r-2d-1}$. Similar to the previous section, it is easy to see that, if the distribution of above is ε-close to uniform distribution on $\mathbb{F}_3$, then $p_{\text{acc}} \leq 1/3 + \varepsilon$.

For $a \in \{0,1,2\}^r$, let $|a| := \sum_i a_i$ and x^a denote the monomial $\prod_i x_i^{a_i}$. In this notation, $g(x) = \sum_{|a| \leq d} g_a x^a$ where $g_a \in \mathbb{F}_3$ are chosen independently and uniformly at random. For $x \in \mathbb{F}_3^r$, let e_x be the column vector of evaluation of all degree d monomials at x, i.e., $e_x := (x^a)_{|a| \leq d}$. Then $g(x) = g^T e_x$ where g is now thought of as the column vector $(g_a)_{|a| \leq d}$ and hence, $g^2(x) = (g^T e_x)^2 = g^T (e_x e_x^T) g$.

$$\langle \beta, g^2 \rangle = \sum_x \beta(x) \left(g^T e_x e_x^T g \right) = g^T \left(\sum_x \beta(x) e_x e_x^T \right) g.$$

We are thus, interested in the quadratic form represented by the matrix $Q^\beta := \sum_x \beta(x) e_x^T e_x$. Observe that all β belonging to the same coset in $\mathsf{P}_{r,2r}/\mathsf{P}_{r,2r-2d-1}$ have the same value for $\langle \beta, g^2 \rangle$ and the matrix Q^β. Hence, by Lemma 3.2, we might without loss of generality, assume that β satisfies $\text{support}(\beta) = \Delta$. The following lemma (an easy consequence of Theorem 6.21 in book by Rudolf & Niederreiter [LN97]), shows that it suffices to understand the rank of Q^β.

Lemma 4.4. *Let A be a $n \times n$, symmetric matrix with entries from $\mathbb{F}_3$. The statistical distance of the random variable $p^T Ap$ from uniform is* $\exp(-\Omega(\text{rank}(A)))$.

In the next sequence of lemmas, we relate $\text{rank}(Q^\beta)$ to Δ. In particular, we show that $\text{rank}(Q^\beta)$ is equal to Δ if $\Delta \leq 3^{d/2}$ and is exponential in d otherwise. Recall that over $\mathbb{F}_3$, $\mathsf{P}_{r,2r}$ is the set of all function from $\mathbb{F}_3^r$ to $\mathbb{F}_3$ and $(\mathsf{P}_{r,2d})^\perp = \mathsf{P}_{r,2r-2d-1}$.

Lemma 4.5. $\text{rank}(Q^\beta) \leq \Delta$.

Proof. By assumption, β satisfies $\Delta = \text{support}(\beta)$. The lemma follows from that fact that $e_x e_x^T$ are rank one matrices and $Q^\beta = \sum_x \beta(x) e_x e_x^T$. □

Lemma 4.6. *If $\Delta < 3^{d/2}$, then* $\text{rank}(Q^\beta) = \Delta$.

Proof. By assumption, β satisfies $\Delta = \text{support}(\beta)$ and $Q^\beta = \sum_x \beta(x) e_x e_x^T$. Since $\mathsf{P}_{r,d}^\perp = \mathsf{P}r, 2r - d - 1$ and any non-zero polynomial with degree $\leq 2r - d - 1$ has support at least $3^{d/2}$ (Lemma 3.9), any $\lceil 3^{d/2} \rceil - 1$ vectors e_x are linearly independent. In particular, the Δ vectors e_x for x in $\text{support}(\beta)$ are linearly independent. Consider any non-zero v in the kernel of the matrix Q^β. The linear independence of e_x's gives that $e_x^T v = 0$ for all $x \in \text{support}(\beta)$. Hence, the kernel of Q^β resides in a Δ-codimensional space which implies that $\text{rank}(Q^\beta) = \Delta$. □

We conjecture that Lemma 4.6 holds for larger values of Δ, but for our purposes we only need a lower bound on the rank when $\Delta \geq 3^{d/2}$.

Lemma 4.7. *There exists a constant d_0 such that if $d > d_0$ and $\Delta > 3^{d/2}$ then* $\mathrm{rank}(Q^\beta) \geq 3^{d/9}$.

Proof. The proof of this theorem is similar to the proofs by Dinur & Guruswami [DG14, Theorems 15,17] for the $\mathbb{F}_2$ case and we follow it step by step. Define

$$B^r_{d,k}(\beta) := \{q \in \mathsf{P}_{r,k} : q\beta \in \mathsf{P}_{r,2r-2d-1+k}\}.$$

Claim 4.8. $\mathrm{kernel}(Q^\beta) = B^r_{d,d}(\beta)$.

Proof. The matrix Q^β satisfies that $Q^\beta(a,b) = \langle \beta, x^a x^b \rangle$, for all $a, b \in \{0,1,2\}^r$, $|a|, |b| \leq d$. Using this description of Q^β, we obtain the following description of $\ker(Q^\beta)$.

$$\begin{aligned}
(h_a)_{|a|\leq d} \in \mathrm{kernel}(Q^\beta) &\iff \forall a : |a| \leq d, && \sum_{b:|b|\leq d} \langle \beta, x^a x^b \rangle h_b = 0 \\
&\iff \forall a : |a| \leq d, && \left\langle \beta, x^a \sum_{b:|b|\leq d} h_b x^b \right\rangle = 0 \\
&\iff \forall a : |a| \leq d, && \langle \beta x^a, h \rangle = 0 \\
&\iff \forall q \in \mathsf{P}_{r,d}, && \langle \beta q, h \rangle = 0 \\
&\iff \forall q \in \mathsf{P}_{r,d}, && \langle \beta h, q \rangle = 0 \\
&\iff \beta h \in \mathsf{P}_{r,2r-d-1} && \square
\end{aligned}$$

Thus to prove Lemma 4.7, it suffices to show that $\mathrm{rank}(Q^\beta) = \dim(\mathsf{P}^r_d / B^r_{d,d}(\beta)) \geq 3^{d/9}$. Towards this end, we define

$$\Phi_{d,k}(D) := \min_{r>d/2, \beta \in \mathsf{P}_{r,2r} : \Delta\left(\beta, \mathsf{P}_{r,2r-2d-1}\right) > D} \dim(\mathsf{P}_{r,k} / B^r_{d,k}(\beta)). \tag{4.3.1}$$

In terms of $\Phi_{d,k}$, Lemma 4.7 now reduces to showing that $\Phi_{d,d}(3^{d/2}) \geq 3^{d/9}$. We obtain this lower bound by recursively bounding this quantity. The following serves as

the base case of the recursion.

Claim 4.9. *For $k > 2d$, $\forall D$, $\Phi_{d,k}(D) = 0$ and for $k \leq 2d$, $\Phi_{d,k}(1) \geq 1$.*

Proof. Let β be the polynomial which attains the minimum in (4.3.1). The first part of the claim follows from the fact that if $k > 2d$ then $B^r_{d,k}(\beta) = \mathsf{P}_{r,k}$.

Now for the second part. Since $\beta \notin \mathsf{P}_{r,2r-2d-1}$, there is a monomial x^a with $|a| \leq 2d$ such that

$$\langle \beta, x^a \rangle \neq 0 \iff \langle \beta x^a, 1 \rangle \neq 0 \iff \beta x^a \notin \mathsf{P}_{r,2r-1}.$$

If $|a| \leq k$, $x^a \notin B^r_{d,k}(\beta)$ and we are done. Otherwise, consider b such that $b \leq a$ coordinate-wise and $|b| = k$. Suppose $x^b\beta \in \mathsf{P}_{r,2r-2d-1+k}$ then $x^a\beta \in \mathsf{P}_{r,2r-1}$ which is a contradiction. Hence, $x^b\beta \notin \mathsf{P}_{r,2r-2d-1+k}$ and the second part of the claim follows. □

For the induction step, we need the following result from Haramaty, Shpilka & Sudan [HSS13].

Claim 4.10 (Haramaty, Shpilka & Sudan [HSS13, Theorems 4.16, 1.7]). *There exists a constant d_0 such that if $3^5 < \Delta < 3^d$, $d > d_0$ where β is Δ-far from $\mathsf{P}_{r,2r-2d-1}$, then there exists nonzero $\ell \in \mathsf{P}_{r,1}$ such that $\forall c \in \mathbb{F}_3$, $\beta|_{\ell=c}$ are $\Delta/27$ far from the restriction of $\mathsf{P}_{r,2r-2d-1}$ to affine hyperplanes.*

Claim 4.11. *If $3^5 \leq D \leq 3^d$ and $d > d_0$, then*

$$\Phi_{d,k}(D) \geq \Phi_{d-1,k}(D/27) + \Phi_{d-1,k-1}(D/27) + \Phi_{d-1,k-2}(D/27).$$

Proof. From Lemma 4.10, we get that there exists non-zero $\ell \in \mathsf{P}_{r,1}$ such that for all $c \in \mathbb{F}_3$, $\beta|_{\ell=c}$ is $\Delta/27$ far from $\mathsf{P}_{r-1,2r-2d-1}$. By applying a change of basis, we can assume that $\ell = x_r$.

Let $\beta = (x_r^2-1)\gamma + x_r\eta + \theta$ and $q = (x_r^2-1)u + (x_r-1)s + t$ where $\gamma, \eta, \theta, u, s, t$ do not depend on x_r. Note that $\theta - \gamma, \theta + \eta, \theta - \eta$ are $D/27$ far from $\mathsf{P}_{r-1,2r-2d-1}$. Expanding the product βq, we have

$$\beta q = (x_r^2 - 1)\left((\theta - \gamma)u + \gamma t + \eta s - \gamma s\right) + (x_r - 1)\left((\theta - \eta)s + \eta t\right) + (\theta + \eta)t.$$

Comparing terms, we observe that $\beta q \in \mathsf{P}_{r,2r-2d-1+k}$ iff the following three items are true:

1. $(\theta - \gamma)u + \gamma t + \eta s - \gamma s \in \mathsf{P}_{r-1,2r-2d-1+k-2}$,
2. $(\theta - \eta)s + \eta t \in \mathsf{P}_{r-1,2r-2d-1+k-1}$,
3. $(\theta + \eta)t \in \mathsf{P}^{r-1,2r-2d-1+k}$.

Since $u \in \mathsf{P}_{r,k-2}, s \in \mathsf{P}_{r,k-1}, t \in \mathsf{P}_{r,k}$, this is equivalent to the following (written in reverse order):

1. $t \in B^{r-1}_{d-1,k}(\theta + \eta)$,
2. $s \in -\eta t + B^{r-1}_{d-1,k-1}(\theta - \eta)$,
3. $r \in \gamma s - \eta s - \gamma t + B^{r-1}_{d-1,k-2}(\theta - \gamma)$.

Since t, s, u belongs to sets with the same size as $B^{r-1}_{d-1,k}(\theta+\eta)$, $B^{r-1}_{d-1,k-1}(\theta-\eta)$, $B^{r-1}_{d-1,k-2}(\theta-\gamma)$ respectively and each choice gives a distinct element of $B^r_{d,k}(\beta)$, we get the following equality.

$$\dim(B^r_{d,k}(\beta)) = \dim(B^{r-1}_{d-1,k}(\theta+\eta)) + \dim(B^{r-1}_{d-1,k-1}(\theta-\eta)) + \dim(B^{r-1}_{d-1,k-2}(\theta-\gamma))$$

Combining this with $\dim(\mathsf{P}_{r,k}) = \dim(\mathsf{P}_{r-1,k}) + \dim(\mathsf{P}_{r-1,k-1}) + \dim(\mathsf{P}_{r-1,k-2})$, we obtain

$$\begin{aligned}\dim(\mathsf{P}_{r,k}/B^r_{d,k}(\beta)) &= \dim(\mathsf{P}_{r-1,k}/B^{r-1}_{d-1,k}(\theta+\eta)) + \dim(\mathsf{P}_{r-1,k-1}/B^{r-1}_{d-1,k-1}(\theta-\eta)) \\ &\quad + \dim(\mathsf{P}^{r-1,k-2}/B^{r-1}_{d-1,k-2}(\theta-\gamma)) \\ &\geq \Phi_{d-1,k}(D/27) + \Phi_{d-1,k-1}(D/27) + \Phi_{d-1,k-2}(D/27).\end{aligned}$$

The last inequality follows from the fact that $\theta-\gamma, \theta+\eta, \theta-\eta$ are $D/27$ far from $\mathsf{P}_{r-1,2r-2d-1} = \mathsf{P}_{r-1,2(r-1)-2(d-1)-1}$. Thus, proved. □

To prove Lemma 4.7, we start with $\Phi_{d,d}(3^{d/2})$ and apply Claim 4.11 recursively $d/6-2$ times and finally use the base case from Claim 4.9 (this can be done as long as $d/6 - 2 \leq d/2$). This gives $\text{rank}(Q^\beta) \geq \Phi_{d,d}(3^{d/2}) \geq 3^{d/6-2} \geq 3^{d/9}$ as long as d_0 is large enough.

□

Lemma 4.12. *If* $\alpha : \mathbb{F}_3^r \to \mathbb{F}_3$ *such that* $\Delta(\alpha, \mathsf{P}^\perp_{r,2d}) > 3^{d/2}$ *then*

$$\left|\mathop{\mathbb{E}}_{p\in\mathsf{P}_{r,d}} \chi_\alpha(p^2)\right| \leq 3^{-\Omega(3^{d/9})}.$$

Proof. By definition $\left|\mathbb{E}_{p\in\mathsf{P}_{r,d}} \chi_\alpha(p^2)\right| = \left|\mathbb{E}_{p\in\mathsf{P}_{r,d}} \omega^{\langle\alpha,p^2\rangle}\right|$. If $\alpha : \mathbb{F}_3^r \to \mathbb{F}_3$ is such that $\Delta(\alpha, \mathsf{P}^\perp_{r,2d}) > 3^{d/2}$ then for a random $p \in \mathsf{P}_{r,d}$, $\langle\alpha, p^2\rangle$ is $3^{-\Omega(3^{d/9})}$-close to the uniform distribution on $\mathbb{F}_3$ according to Lemma 4.7 and Lemma 4.4. □

5

Independent Sets in Graph Products

In this chapter, we will use the testing results from the previous chapter, to obtain derandomized graph product results. Gadgets constructed from graph products were previously used in graph coloring hardness reductions. By derandomized graph products, we mean that there are subgraphs of these graph products, with much smaller size, that have the properties to make the hardness reductions work. These properties are also of independent interest as a combinatorial question.

5.1 Graph Products

We consider the following graph product result due to Alon et al. [ADFS04]. Consider the undirected weighted graph K_3 on the three vertices $V = \{0, 1, 2\}$ and edges weighted as follows: $W(f, f') = 1/2$ if $f' \neq f \in \{0, 1, 2\}$ and 0 otherwise. Let $K_3^{\otimes R}$ be the graph with vertex set $V^{\otimes R}$ and weights-matrix the R-wise tensor of the matrix W. Consider independent sets in this wieghted graph (which are sets such that for every edge with non-zero weight, both end points does not lie in the set). Clearly, for any $i \in [R]$ and $a \in \{0, 1, 2\}$, the set $V_{i,a} := \{v \in V^{\otimes R} : v_i = a\}$ is an independent set in $K_3^{\otimes R}$ of fractional size $1/3$ since K_3 does not have any self loops. We call such an independent set a *dictator* for obvious reasons. Alon et al. [ADFS04] showed that these are the maximal independent sets in $K_3^{\otimes R}$ and in fact any independent set of size close

to the maximum is close to a dictator.

Theorem 5.1 (Alon et al. [ADFS04]). *Let A be an independent set in $K_3^{\otimes R}$ of size $\delta 3^R$. Then,*

1. $\delta \leq 1/3$.
2. $\delta = 1/3$ *iff A is a dictator.*
3. *If $\delta \geq 1/3 - \varepsilon$, then A is $O(\varepsilon)$-close to a dictator. That is, there is a dictator A' such that $|A \Delta A'| = O(\varepsilon 3^R)$.*

One may ask if something can be said about independent set of constant density. While studying the hardness of approximate graph coloring, Dinur, Mossel & Regev [DMR09] proved the following "majority is stablest" type of result: if there is a pair of subsets of vertices in $K_3^{\otimes R}$ of sufficiently large size such that the average weight of edges between them is small, then their indicator functions must have a common influential coordinate. Subsequently, Dinur & Shinkar [DS10] obtained the following quantitative improvement to the above theorem.

Theorem 5.2 (Dinur & Shinkar [DS10, Theorem 1.3]). *For all $\mu > 0$ there exists $\delta = \mu^{O(1)}$ and $k = O(\log 1/\mu)$ such that the following holds: For any two functions $A, B : \{0,1,2\}^R \to [0,1]$ if*

$$\mathbb{E} A > \mu, \ \mathbb{E} B > \mu, \ \text{and} \ \mathop{\mathbb{E}}_{f,f'} A(f)B(f') \leq \delta^{\star}$$

where f is chosen randomly from $V^{\otimes R}$ and f' is chosen with probability $W^{\otimes R}(f, f')$ then

$$\exists x \in [R] \text{ such that } \mathrm{Inf}_x^{\leq k}(A) \geq \delta \text{ and } \mathrm{Inf}_x^{\leq k}(B) \geq \delta.$$

[⋆] The hypothesis in the theorem statement of Dinur-Shinkar [DS10] requires $\mathbb{E}_{f,f'} A(f)B(f') = 0$, however it is easy to check that their theorem also holds good under the weaker hypothesis $\mathbb{E}_{f,f'} A(f)B(f') \leq \delta$.

5.2 Derandomized Graph Products

The product graph $K_3^{\otimes R}$ has 3^R vertices. In this section , we show that that there exists a considerably smaller subgraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of $K^{\otimes R}$ with only $3^{\text{poly}(\log R)}$ vertices that has the same properties. In order to describe the subgraph, it will be convenient to think of K_3 as having vertex set $\mathbb{F}_3$ and

$$W(f, f') = \Pr_{p \in \mathbb{F}_3, a \in \{1,2\}}[f' = f + a(p^2 + 1)].$$

Let r and d be two parameters and let $R = 3^r$. Note that $V^{\otimes R}$ can be identified with $\mathsf{P}_{r,2r}$, since $\mathsf{P}_{r,2r}$ is the set of all functions from $\mathbb{F}_3^r$ to $\mathbb{F}_3$. The subgraph $\mathcal{G}_d = (\mathcal{V}, \mathcal{E})$ is as follows : $\mathcal{V} := \mathsf{P}_{r,2d}$ and the edges are given by the weights-matrix defined below

$$\mathcal{W}(f, f') = \Pr_{p \in \mathsf{P}_{r,d}, a \in \{1,2\}}[f' = f + a(p^2 + 1)].$$

Note that since $\mathsf{P}_{r,2d}$ is a subspace of dimension $r^{O(d)}$, the size of the vertex set is $3^{r^{O(d)}}$, which is considerably smaller than 3^R for constant d.

Theorem 5.3. *There is a constant d for which the following holds. If A is an independent set of size $\delta|\mathcal{V}|$ in $\mathcal{G}_d$ then*

1. $\delta \leq 1/3$.
2. $\delta = 1/3$ *iff A is a dictator.*
3. *If $\delta \geq 1/3 - \varepsilon$ then A is $O(\varepsilon)$-close to a dictator.*

A crucial element in the proof of Theorem 5.1 is a hypercontractivity theorem for functions which do not have any heavy Fourier coefficients. Theorem 5.3 is proved by observing that a similar hypercontractivity theorem also holds good in the low-degree long code setting (see Lemma 3.18).

Proof of 1. For $f \in V$, consider the set $\{f, f+1, f+2\} \subseteq V$. These sets form a partition of V and are triangles in the graph. Hence $\delta \leq 1/3$. □

Proof of 2. Let $A : \mathsf{P}_{r,2d} \to \{0,1\}$ be the indicator set of the independent set of size $\delta|V|$. By Parseval's equation and the fact that $\widehat{A}_0 = \delta$, we have that

$$\sum_{\alpha \in \Lambda_{r,2d} \setminus \{0\}} |\widehat{A}_\alpha|^2 = \delta - \delta^2. \tag{5.2.1}$$

Since A is an independent set,

$$\mathop{\mathbb{E}}_{p \in \mathsf{P}_{r,d}, a \in \mathbb{F}_3, f \in \mathsf{P}_{r,2d}} A(f)A(f+a(p^2+1)) = \sum_{\alpha \in \Lambda_{r,2d}} |\widehat{A}_\alpha|^2 \mathop{\mathbb{E}}_{p \in \mathsf{P}_{r,d}, a \in \mathbb{F}_3} \chi_\alpha(a(p^2+1)) = 0.$$

Taking the real parts of the equation on both sides and rearranging, we get

$$\sum_{\alpha \in \Lambda_{r,2d} \setminus \{0\}} |\widehat{A}_\alpha|^2 \operatorname{Re}\left(\mathop{\mathbb{E}}_{p \in \mathsf{P}_{r,d}} \chi_\alpha(p^2+1) \right) = -\delta^2. \tag{5.2.2}$$

Let T be a random variable such that $\Pr[T = \alpha] = |\widehat{A}_\alpha|^2/(\delta - \delta^2)$ and X be the random variable $X(T) = \operatorname{Re}\left(\mathbb{E}_{p \in \mathsf{P}_{r,d}, a \in \mathbb{F}_3} \chi_\alpha(a(p^2+1))\right)$. From (5.2.1) and (5.2.2), we have that

$$\mathbb{E}\, X = \frac{-\delta}{1-\delta}.$$

Since p is a random degree d polynomial, it is $3^{d/2}$-wise independent from Lemma 3.10. So if $|T| \leq 3^{d/2}$,

$$\begin{aligned}
\left| \operatorname{Re}\left(\mathop{\mathbb{E}}_{p \in \mathsf{P}_{r,d}, a \in \mathbb{F}_3} \chi_\alpha(a(p^2+1)) \right) \right| &= \left| \frac{1}{2} \operatorname{Re}\left(\left(\frac{\omega^2-1}{3} \right)^{|\alpha|_1} \left(\frac{\omega-1}{3} \right)^{|\alpha|_2} \right.\right. \\
&\qquad \left.\left. + \left(\frac{\omega-1}{3} \right)^{|\alpha|_1} \left(\frac{\omega^2-1}{3} \right)^{|\alpha|_2} \right) \right| \\
&\leq \left(\frac{1}{\sqrt{3}} \right)^{|\alpha|}
\end{aligned}$$

where $|\alpha|_a = \{x \in \mathbb{F}_3^r : \alpha(x) = a\}$.

If $|T| > 3^{d/2}$, we know from Lemma 4.12 that $|X(T)| \leq 3^{-\Omega(3^{d/9})}$.

Note that for T with $|T| = 1$, $X(T) = -1/2$. For T with $|T| = 2$, $X(T) \geq 0$. For T with $|T| \geq 3$, $X(T) \geq \frac{-1}{3\sqrt{3}}$. So if $\mathbb{E}\, X = -1/2$ then $\Pr[|T| = 1] = 1$. So A is a Boolean valued function with non zero Fourier coefficients of support size only 0 and 1. From Lemma 3.17, A is a generalized dictator function. □

Proof of 3. Suppose $\delta = 1/3 - \varepsilon$. First we show that most of Fourier weights are concentrated in the first two levels

Lemma 5.4.

$$\sum_{\alpha \in \Lambda_{r,2d} : |\alpha| > 1} |\widehat{A}_\alpha|^2 \leq 2\varepsilon$$

Proof. Consider the random variables X and T defined in the Proof of 2. Since $\delta = 1/3 - \varepsilon$ and since $\varepsilon < 1/3$, $\mathbb{E}\, X = -1/2 + \varepsilon$. Let Y be the random variable $X + 1/2$. Note that $Y \geq 0$ and when $Y > 0$, $Y \geq 1/6$. Therefore by Markov, $\Pr[Y > 0] \leq 6\varepsilon$ and

$$\sum_{\alpha \in \Lambda_{r,2d} : |\alpha| > 1} |\widehat{A}_\alpha|^2 \leq (\delta - \delta^2) \Pr[Y > 0] \leq 2\varepsilon.$$

□

Then we use Lemma 3.19 to obtain the result.

□

5.3 Derandomized Majority is Stablest

We also show that an analogue of Theorem 5.2 also holds for the subgraph $\mathcal{G}$. For defining influence for real valued functions on $\mathsf{P}_{r,2d}$, we note that the characters of $\mathsf{P}_{r,2d}$ are

restrictions of characters of $\mathbb{F}_3^R \equiv \mathsf{P}_{r,2r}$. So the definition of influence for functions on $\mathbb{F}_3^R$ also extends naturally to functions on $\mathsf{P}_{r,2d}$.

Theorem 5.5. *For all $\mu > 0$ there exists $\delta = \mu^{O(1)}, k = O(\log 1/\mu), d = O(\log 1/\mu)$ such that the following holds: For any two functions $A, B : \mathsf{P}_{r,2d} \to [0,1]$ if*

$$\mathbb{E} A > \mu, \ \mathbb{E} B > \mu, \ \text{and} \ \mathop{\mathbb{E}}_{f,f'} A(f)B(f') \leq \delta$$

where f is chosen randomly from $\mathsf{P}_{r,2d}$, $f' = f + a(p^2+1)$, p are chosen randomly from $\mathsf{P}_{r,d}$ and $a \in \{1,2\}$ then

$$\exists x \in \mathbb{F}_3^r \text{ such that } \mathrm{Inf}_x^{\leq k}(A) \geq \delta \text{ and } \mathrm{Inf}_x^{\leq k}(B) \geq \delta.$$

In this section, we prove Theorem 5.5. The graphs described in Theorem 5.5 and Theorem 5.2 can be viewed as Cayley graphs on a suitable group. For the proof, we will need bounds on the eigenvalues of these Cayley graphs. These bounds are obtained from the testing results from the previous chapter. For a group G, $\mathbb{R}^G$ denotes the vector space of real valued functions on G.

Definition 5.6 (Cayley Operator). *For a group G with operation $+$, an operator $M : \mathbb{R}^G \to \mathbb{R}^G$ is a* Cayley operator *if there is a distribution μ on G such that for any function $A : G \to \mathbb{R}$,*

$$(MA)(f) = \mathop{\mathbb{E}}_{\eta \in \mu} A(f + \eta).$$

It is easy to see that a character $\chi : G \to \mathbb{C}$ is an eigenvector of M with eigenvalue $\mathbb{E}_{\eta \in \mu} \chi(\eta)$.

Definition 5.7. *We define the following Cayley operators:*

1. *For the group $\mathbb{F}_3$, let $T : \mathbb{R}^{\mathbb{F}_3} \to \mathbb{R}^{\mathbb{F}_3}$ be the Cayley operator corresponding to the distribution μ that is uniform on $\mathbb{F}_3 \setminus \{0\}$. Let λ be the second largest eigenvalue*

in absolute value of T.

2. *For the group* $\mathcal{F}_r$*, let* $T_r : \mathbb{R}^{\mathcal{F}_r} \to \mathbb{R}^{\mathcal{F}_r}$ *be the Cayley operator corresponding to the distribution* μ_r *that is uniform on* $\{f \in \mathcal{F}_r : f^{-1}(0) = \emptyset\}$*. Let* $\lambda_r(\alpha)$ *be the eigenvalue of* T_r *corresponding to the eigenvector* χ_α*, for* $\alpha \in \mathcal{F}_r$*.*

3. *For the group* $\mathsf{P}_{r,2d}$*, let* $T_{r,d} : \mathbb{R}^{\mathsf{P}_{r,2d}} \to \mathbb{R}^{\mathsf{P}_{r,2d}}$ *be the Cayley operator corresponding to the distribution* $\mu_{r,2d}$ *of choosing a uniformly random element in* $\{p^2+1, -p^2-1\}$ *where* $p \in \mathsf{P}_{r,d}$ *is chosen uniformly at random. Let* $\lambda_{r,d}(\alpha)$ *be the eigenvalue of* $T_{r,d}$ *corresponding to* χ_α*, for* $\alpha \in \mathcal{F}_r$*.*

4. *For the group* $\mathsf{P}_{r,2d}$*, let* $S_{r,d} : \mathbb{R}^{\mathsf{P}_{r,2d}} \to \mathbb{R}^{\mathsf{P}_{r,2d}}$ *be the Cayley operator corresponding to the distribution of* $a \cdot \prod_{i=1}^{d}(\ell_i - 1)(\ell_i - 2)$ *where* $\ell_1, \cdots, \ell_d$ *are linearly independent degree* 1 *polynomials chosen uniformly at random and* a *is randomly chosen from* $\mathbb{F}_3$*. Let* $\rho_{r,d}(\alpha)$ *be the eigenvalue of* $S_{r,d}$ *corresponding to* χ_α*, for* $\alpha \in \mathcal{F}_r$*.*

Now we will list some known bounds of the eigenvalues of the above operators. It is easy to see that λ is a constant < 1. Since $\mathbb{F}_3^R$ can be identified with $\mathcal{F}_r$, $T^{\otimes R}$ can be identified with T_r. Hence we have the following lemma.

Lemma 5.8.

$$|\lambda_r(\alpha)| \leq |\lambda|^{|\alpha|}.$$

Lemma 5.9. *For* $\alpha \in \Lambda_{r,2d}$*,*

$$|\lambda_{r,d}(\alpha)| \begin{cases} = |\lambda_r(\alpha)| & \text{if } |\alpha| \leq 3^{d/2} \\ \leq 3^{-3^{C_1 d}} & \text{otherwise.} \end{cases} \tag{5.3.1}$$

Proof. The first case follows from the fact that a random element η according to $\mu_{r,2d}$ (the distribution that defines $T_{r,d}$) is $3^{d/2}$-wise independent (see Lemma 3.10) as a string

of length 3^r over alphabet $\mathbb{F}_3$. The latter case follows from Lemma 4.12. □

We will derive bounds on the eigenvalues of $S_{r,d}$ using the results of Haramaty et al. [HSS13].

Lemma 5.10. *There exists constants C_1', C_2' such that, for $\alpha \in \Lambda_{r,2d}$,*

$$1 - \frac{2|\alpha|}{3^d} \leq |\rho_{r,d}(\alpha)| \leq \max\left\{1 - \frac{C_1'\Delta(\alpha, \mathsf{P}_{r,2r-2d-1})}{3^d}, C_2'\right\} \tag{5.3.2}$$

Proof. First we will prove the lower bound. By definition

$$\rho_{r,d}(\alpha) = \mathop{\mathbb{E}}_{\ell_i,a} \omega^{a\cdot\sum_x \alpha(x)\prod_{i=1}^d(\ell_i(x)-1)(\ell_i(x)-2)}.$$

For any x in support of α, the probability that $\prod_{i=1}^d(\ell_i(x)-1)(\ell_i(x)-2) \neq 0$ is $1/3^d$. Hence by union bound, $\prod_{i=1}^d(\ell_i(x)-1)(\ell_i(x)-2) = 0$ for every x in support of α with probability $1 - |\alpha|/3^d$ and when this happens the expectation is 1. Also note that the quantity inside the expectation has absolute value 1.

For proving the upper bound we will use Lemma 4.2. Let p_{acc} be the probability mentioned in Lemma 4.2. Then

$$\rho_{r,d}(\alpha) = \mathop{\mathbb{E}}_{\ell_i,a} \omega^{a\langle\alpha,\prod_{i=1}^d(\ell_i-1)(\ell_i-2)\rangle} = p_{\mathsf{acc}} + \frac{1-p_{\mathsf{acc}}}{2}(\omega+\omega^2) = \frac{3}{2}p_{\mathsf{acc}} - \frac{1}{2}.$$

From the above equation and Lemma 4.2, the constants C_1', C_2' can be obtained. □

Lemma 5.11. *For $A, B : \mathsf{P}_{r,2d} \to [0,1]$, let $A' := S_{r,d}^t A$ and similarly define B'. Then*

$$|\langle A, T_{r,d}B\rangle - \langle A', T_{r,d}B'\rangle| \leq 2dt/3^d$$

Proof.

$$\begin{aligned}
|\langle A, T_{r,d}B\rangle - \langle A', T_{r,d}B'\rangle| &\leq |\langle A, T_{r,d}B\rangle - \langle A, T_{r,d}B'\rangle| \\
&\quad + |\langle A, T_{r,d}B'\rangle - \langle A', T_{r,d}B'\rangle| \\
&= \left|\langle A - \mathbb{E}\, A, T_{r,d}(1 - S^t_{r,d})(B - \mathbb{E}\, B)\rangle\right| \\
&\quad + \left|\langle T_{r,d}(1 - S^t_{r,d})(A - \mathbb{E}\, A), B' - \mathbb{E}\, B'\rangle\right| \\
&\leq \|T_{r,d}(1 - S^t_{r,d})(B - \mathbb{E}\, B)\| + \|T_{r,d}(1 - S^t_{r,d})(A - \mathbb{E}\, A)\| \\
&\leq 2td/3^d
\end{aligned}$$

The last step follows from the fact that the operators $T_{r,d}$, $(1-S^t_{r,d})$ have the same set of eigenvectors and the largest eigenvalue in absolute value of $T_{r,d}(1 - S^t_{r,d})$ is $2td/3^d$ from Lemma 5.9 and Lemma 5.10. □

Theorem 5.5 will follow from the following lemma.

Lemma 5.12. $\forall \varepsilon > 0, \exists k = O(1/\varepsilon^2), d = O(\log(1/\varepsilon))$ *such that the following holds: if* $A, B : \mathsf{P}_{r,2d} \to [0,1]$ *then* $\exists \mathcal{A}, \mathcal{B} : \mathcal{F}_r \to [0,1]$ *such that*

1. $|\mathbb{E}\, A - \mathbb{E}\, \mathcal{A}|, |\mathbb{E}\, B - \mathbb{E}\, \mathcal{B}| \leq \varepsilon$,
2. *For all* $x \in \mathbb{F}_3^r, k' \leq k$,

$$\mathrm{Inf}_x^{\leq k'}(\mathcal{A}) \leq \mathrm{Inf}_x^{\leq k'}(A) + \varepsilon$$

$$\mathrm{Inf}_x^{\leq k'}(\mathcal{B}) \leq \mathrm{Inf}_x^{\leq k'}(B) + \varepsilon$$

3. $|\langle A, T_{r,d}B\rangle - \langle \mathcal{A}, T_r\mathcal{B}\rangle| \leq \varepsilon$.

Proof of Theorem 5.5. We will show that if Theorem 5.5 is false then Theorem 5.2 is also false. First using Lemma 5.12 with parameter $\varepsilon = \mu^{O(1)}$, we obtain functions $\mathcal{A}, \mathcal{B}$:

$\mathcal{F}_r \to [0,1]$ such that

1. $\mathbb{E}\,\mathcal{A}, \mathbb{E}\,\mathcal{B} \geq \mu - \varepsilon$,

2. For all $x \in \mathbb{F}_2^r$, $k' \leq k$,

$$\mathrm{Inf}_x^{\leq k'}(\mathcal{A}) \leq \delta + \varepsilon \text{ and } \mathrm{Inf}_x^{\leq k'}(\mathcal{B}) \leq \delta + \varepsilon$$

3. $|\langle \mathcal{A}, T_r \mathcal{B} \rangle| \leq |\langle A, T_{r,d} B \rangle| + \varepsilon$.

Now applying Theorem 5.2 to the functions $\mathcal{A}, \mathcal{B}$, we obtain that $|\langle \mathcal{A}, T_r \mathcal{B} \rangle| \geq \delta'$, where $\delta' = \mu^{O(1)}$. Hence $|\langle A, T_{r,d} B \rangle| \geq \delta' - \varepsilon$, and we set the parameters $\delta = \delta' - \varepsilon$, $d = O(\log 1/\mu)$ and $k = O(\log 1/\mu)$.

□

5.3.1 Proof of Lemma 5.12

For proving Lemma 5.12, crucially use the following lemma by Kane & Meka [KM13].

Lemma 5.13. *Let $\xi : \mathbb{R} \to \mathbb{R}_+$ be the function $\xi(x) := (\max\{-x, x-1, 0\})^2$ [†]. For any parameters $k \in \mathbb{N}$ and $\varepsilon \in (0,1)$, there is a $d = O(\log(k/\varepsilon))$ such that the following holds: If the polynomial $P : \mathcal{F}_r \to \mathbb{R}$ satisfies $\|P\| \leq 1$ and $\widehat{P}(\alpha) = 0$ for $\alpha \in \Lambda_{r,d}$ such that $|\alpha| > k$, then*

$$\left| \mathop{\mathbb{E}}_{f \in \mathcal{F}_r} \xi(P(f)) - \mathop{\mathbb{E}}_{f \in \mathsf{P}_{r,d}} \xi(P(f)) \right| \leq \varepsilon.$$

Remark 5.14. *For proving Lemma 5.13, we need a generalization of [KM13, Lemma 4.1] in the paper of Kane & Meka. [KM13, Lemma 4.1] considers uniform distribution over $\{-1,+1\}^R$, but we need a similar result for uniform distribution over $\{1, \omega, \omega^2\}^R$.*

[†] $\xi(x)$ represents the distance of x from the interval $[0,1]$

However, we observe that the polynomials we consider are real-valued $P : \mathcal{F}_r \to \mathbb{R}$ and hence satisfy $\widehat{P}(\alpha) = \overline{P}(-\alpha)$.

Using this observation, the proof of Kane & Meka [KM13, Lemma 4.1] generalizes to our setting (the above property is preserved throughout the proof). The result of [KM13] also requires an earlier result of Diakonikolas, Gopalan, Jaiswal, Servedio & Viola [DGJ+10] on fooling Linear Threshold functions (LTFs) with sample spaces of bounded independence. A generalization of [DGJ+10] is also known due to Gopalan et al. ([GOWZ10, Theorem I.5]) for uniform distribution over $\{1, \omega, \omega^2\}^R$.

Proof of Lemma 5.12. Let $t = \frac{3^d \log(10/\varepsilon)}{2k}$, and $A_1 = S^t_{r,d} A, B_1 = S^t_{r,d} B$. Then from Lemma 5.11

$$|\langle A, T_{r,d} B\rangle - \langle A_1, T_{r,d} B_1\rangle| \leq 2dt/3^d \tag{5.3.3}$$

and similarly for B_1. Let k be a number $< 3^{d/2}$ and $A_2 = \mathrm{Re}(A_1^{\leq k})$. Using the fact that A_1 is real valued,

$$\|A_1 - A_2\| \leq \|A_1 - A_1^{\leq k}\| \leq (1 - 2k/3^d)^t \leq e^{-2tk/3^d} = \varepsilon/10 \tag{5.3.4}$$

Let $A_3 : \mathcal{F}_r \to \mathbb{R}$ be defined as $A_3 := \mathrm{Re}((A_1^{\leq k})')$ where $(A_1^{\leq k})'$ is the lift (Definition 3.21) of $A_1^{\leq k}$. Since a random degree d polynomial is $3^{d/2}$-wise independent (see Lemma 3.10),

$$\langle A_2, T_{r,d} B_2\rangle = \langle A_3, T_r B_3\rangle \tag{5.3.5}$$

Note that A_3 may not be a $[0,1]$-valued function. But since A is $[0,1]$-valued, so is A_1. Let $\xi : \mathbb{R} \to \mathbb{R}_+$ be the function $\xi(x) := (\max\{-x, x-1, 0\})^2$. Notice that $\mathbb{E}_f \xi \circ A(f)$ gives the ℓ_2^2 distance of A from $[0,1]$-valued functions. Using Lemma 5.13,

for $d = O(\log(k/\varepsilon))$,

$$\left| \mathop{\mathbb{E}}_{f \in \mathsf{P}_{r,2d}} \xi(A_2(f)) - \mathop{\mathbb{E}}_{f \in \mathcal{F}_r} \xi(A_3(f)) \right| \leq \varepsilon/10 \tag{5.3.6}$$

and similarly for B_2. Hence there exists functions $\mathcal{A}, \mathcal{B} : \mathcal{F}_r \to [0,1]$ such that

1. $|\mathbb{E}\, A - \mathbb{E}\, \mathcal{A}| \leq ||A_1' - \mathcal{A}|| \leq \varepsilon$ (similarly for B),
2. For all $x \in \mathbb{F}_3^r$, $k' \leq k$, $\mathrm{Inf}_x^{\leq k'}(\mathcal{A}) \leq \mathrm{Inf}_x^{\leq k'}(A) + \varepsilon$ (similarly for B),
3. $|\langle A, T_{r,d} B \rangle - \langle \mathcal{A}, T_r \mathcal{B} \rangle| \leq \varepsilon$.

□

Part III

Hardness of Approximate Coloring

6

PCPs & Hardness of Approximation

In this chapter, we will review some of the fundamental results in hardness of approximation. One of the main results is an alternate characterization for NP. Recall the definition of NP (Definition 1.3). For the verifier of an NP problem, the proof length is at most a polynomial in n. The alternate characterization deals with probabilistic verifiers, that query only a few bits in the proof.

Definition 6.1 ($\text{PCP}_{c,s}[t, q, R]$). *A $\text{PCP}_{c,s}[t(n), q(n), R]$-Verifier (PCP stands for Probabilistically Checkable Proofs) for a decision problem L, is an algorithm V which takes an input x, a random string $y \in \{0,1\}^{t(n)}$, has oracle access to a proof π over alphabet $[R]$ of length $2^{t(n)}$ and satisfies the following properties:*

- *V runs in polynomial time in the length of x for all y, π.*
- *V queries π in at most $q(n)$ bits on all inputs.*
- *Completeness: If $x \in L$ then there exists π such that $\Pr_y[V(x, \pi) = 1] \geq c$.*
- *Soundness : If $x \notin L$ then for any π, $\Pr_y[V(x, \pi) = 1] \leq s$.*

$\text{PCP}_{c,s}[t, q, R]$ is the class of decision problems that have a $\text{PCP}_{c,s}[t, q, R]$-Verifier.

It is not difficult to see that, for any q, constant R and $s < c$,

$$\text{PCP}_{c,s}[O(\log n), q, R] \subseteq \text{NP}.$$

A major breakthrough in PCP characterizations, which lead to many hardness of approximation results was the following theorem due to Arora & Safra [AS98] and Arora et al. [ALM+98].

Theorem 6.2 (PCP Theorem). *There exists constant $s < 1, q, R$ such that*

$$\text{NP} \subseteq \text{PCP}_{1,s}[O(\log n), q, R].$$

In the above form, it is not clear how the theorem might be useful in hardness of approximation results. The theorem can be stated equivalently as a hardness of approximation result.

Theorem 6.3 (Hardness of Approximating Max-3-SAT). *There exists a constant $s < 1$, such that it is* NP-Hard *to distinguish satisfiable* Max-3-SAT *instances, from ones for which any assignment can satisfy at most s fraction of the clauses.*

It is easy to see that the above theorem is equivalent to having a $\text{PCP}_{1,s}[O(\log n), 3, 2]$-Verifier for Max-3-SAT, whose checks are Max-3-SAT clauses. The equivalence follows from viewing such verifiers as Max-3-SAT instances and vice versa. Any such verifier could be converted to a Max-3-SAT instance. The instance is obtained by adding a variable for every bit of the proof, and adding a Max-3-SAT clause for every check the verifier makes. Any Max-3-SAT instance, has a trivial PCP verification procedure, whose proof is the assignment and the verifier chooses a random clause, and checks if it is satisfied by the assignment.

6.1 Label Cover

The PCP theorem stated in terms hardness of approximation of Max-3-SAT, does not give optimal inapproximability results (see Håstad [Hås01]). To get strong results, it is

used in conjunction with the parallel repetition theorem of Raz [Raz98]. This strong version of PCP theorem is usually stated in terms of the Label-Cover problem.

Definition 6.4 (Label-Cover). *An instance $G = (U, V, E, L, R, \{\pi_e\}_{e \in E})$ of the* LC *constraint satisfaction problem consists of a bi-regular bipartite graph (U, V, E), two sets of alphabets R and L and a projection map $\pi_e : R \to L$ for every edge $e \in E$. Given a labeling $\ell : U \to R, \ell : V \to L$, an edge $e = (u, v)$ is said to be satisfied by ℓ if $\pi_e(\ell(v)) = \ell(u)$.*

G is said to be at most δ-satisfiable *if every labeling satisfies at most a δ fraction of the edges.*

An instance of Unique-Game *is a label cover instance where $L = R$ and the constraints π are permutations.*

We consider label cover instances obtained from 3-SAT instances in the following natural manner.

Definition 6.5 (r-repeated label cover). *Let φ be a* 3-SAT *instance with X as the set of variables and C the set of clauses. The r-repeated bipartite label cover instance $I(\varphi)$ is specified by:*

- *A graph $G := (U, V, E)$, where $U := C^r, V := X^r$.*
- $\Sigma_U := \{0,1\}^{3r}, \Sigma_V := \{0,1\}^r$.
- *There is an edge $(u, v) \in E$ if the tuple of variables v can be obtained from the tuple of clauses u by replacing each clause by a variable in it.*
- *The constraint $\pi_{uv} : \{0,1\}^{3r} \to \{0,1\}^r$ is simply the projection of the assignments on $3r$ variables in all the clauses in u to the assignments on the r variables in v.*

- *For each u there is a set of r functions $\{f_i^u : \{0,1\}^{3r} \to \{0,1\}\}_{i=1}^r$ such that $f_i^u(a) = 0$ iff the assignment a satisfies the ith clause in u. Note that f_i^u depends only on the 3 variables in the ith clause.*

A labeling $L_U : U \to \Sigma_U, L_V : V \to \Sigma_V$ satisfies an edge (u,v) iff $\pi_{uv}(L_U(u)) = L_V(v)$ and $L_U(u)$ satisfies all the clauses in u. Let $\mathrm{OPT}(I(\varphi))$ *be the maximal fraction of constraints that can be satisfied by any labeling.*

The following theorem is obtained by applying parallel repetition theorem of Raz [Raz98] with r repetitions on hard instances of Max-3-SAT where each variable occurs the same number of times (see Feige's result [Fei98]) and a structural property proved by Håstad [Hås01, Lemma 6.9].

Theorem 6.6. *There is an algorithm which on input a* 3-SAT *instance φ and $r \in \mathbb{N}$ outputs an r-repeated label cover instance $I(\varphi)$ in time $n^{O(r)}$ with the following properties.*

- *Completeness: If $\varphi \in$* 3-SAT*, then* $\mathrm{OPT}(I(\varphi)) = 1$.
- *Soundnes: If $\varphi \notin$* 3-SAT*, then* $\mathrm{OPT}(I(\varphi)) \leq 2^{-\varepsilon_0 r}$ *for some universal constant $\varepsilon_0 \in (0,1)$.*
- *Smooth Projections:*

$$\forall v \in V, \alpha \subset R, \qquad Pr_u\left[|\pi_{uv}(\alpha)| < |\alpha|^{c_0}\right] \leq \frac{1}{|\alpha|^{c_0}}.$$

Moreover, the underlying graph G is both left and right regular.

For our hardness results for 3-uniform 3-colorable hypergraphs, we need a multipartite version of label cover, satisfying a smoothness condition.

Definition 6.7 ([Kho02b]). *Let I be a bipartite label cover instance specified by $((U, V, E), \Sigma_U, \Sigma_V, \Pi)$. Then I is* η-smooth *iff for every $u \in U$ and two distinct labels $a, b \in \Sigma_U$*

$$\Pr_v[\pi_{uv}(a) = \pi_{uv}(b)] \leq \eta,$$

where v is a random neighbour of u.

Definition 6.8 (r-repeated ℓ-layered η-smooth label cover). *Let $T := \lceil \ell/\eta \rceil$ and φ be a* 3-SAT *instance with X as the set of variables and C the set of clauses. The r-repeated ℓ-layered η-smooth label cover instance $I(\varphi)$ is specified by:*

- *An ℓ-partite graph with vertex sets $V_0, \cdots V_{\ell-1}$. Elements of V_i are tuples of the form (C', X') where C' is a set of $(T + \ell - i)r$ clauses and X' is a a set of ir variables.*
- *$\Sigma_{V_i} := \{0,1\}^{m_i}$ where $m_i := 3(T + \ell - i)r + ir$ which corresponds to all Boolean assignments to the clauses and variables corresponding to a vertex in layer V_i.*
- *For $0 \leq i < j < \ell$, $E_{ij} \subseteq V_i \times V_j$ denotes the set of edges between layers V_i and V_j. For $v_i \in V_i, v_j \in V_j$, there is an edge $(v_i, v_j) \in E_{ij}$ iff v_j can be obtained from v_i by replacing some $(j - i)r$ clauses in v_i with variables occurring in the clauses respectively.*
- *The constraint $\pi_{v_i v_j}$ is the projection of assignments for clauses and variables in v_i to that of v_j.*
- *For each $i < \ell$, $v_i \in V_i$, there are $(T+\ell-i)r$ functions $f_j^{v_i} : \{0,1\}^{3(T+\ell-i)r+ir} \to \{0,1\}$, one for each clause j in v_i such that $f_j^{v_i}(a) = 0$ iff a satisfies the clause j. This function only depends on the 3 coordinates in j.*

Given a labeling $L_i : V_i \to \Sigma_{V_i}$ *for all the vertices, an edge* $(v_i, v_j) \in E_{ij}$ *is satisfied iff* $L_i(v_i)$ *satisfies all the clauses in* v_i*,* $L_j(v_j)$ *satisfies all the clauses in* v_j *and* $\pi_{v_i v_j}(L_i(v_i)) = L_j(v_j)$*. Let* $\mathrm{OPT}_{ij}(I(\varphi))$ *be the maximum fraction of edges in* E_{ij} *that can be satisfied by any labeling.*

The following theorem was proved by Dinur et al. [DGKR05] in the context of hypergraph vertex cover inapproximability (also see results of Dinur, Regev & Smyth [DRS05]).

Theorem 6.9. *There is an algorithm which on input a* 3-SAT *instance* φ *and* $\ell, r \in \mathbb{N}, \eta \in [0, 1)$ *outputs a* r*-repeated* ℓ*-layered* η*-smooth label cover instance* $I(\varphi)$ *in time* $n^{O((1+1/\eta)\ell r)}$ *with the following properties.*

1. $\forall\, 0 \le i < j < \ell$*, the bipartite label cover instance on* $I_{ij} = \big((V_i, V_j, E_{ij}), \Sigma_{V_i}, \Sigma_{V_j}, \Pi_{ij}\big)$ *is* η*-smooth.*

2. *For* $1 < m < \ell$*, any* m *layers* $0 \le i_1 < \cdots < i_m \le \ell - 1$*, any* $S_{i_j} \subseteq V_{i_j}$ *such that* $|S_{i_j}| \ge \frac{2}{m}|V_{i_j}|$*, there exists distinct* i_j *and* $i_{j'}$ *such that the fraction of edges between* S_{i_j} *and* $S_{i_{j'}}$ *relative to* $E_{i_j i_{j'}}$ *is at least* $1/m^2$*.*

3. *If* $\varphi \in$ 3-SAT*, then there is a labeling for* $I(\varphi)$ *that satisfies all the constraints.*

4. *If* $\varphi \notin$ 3-SAT*, then*

$$\mathrm{OPT}_{i,j}(I(\varphi)) \le 2^{-\Omega(r)}, \quad \forall 0 \le i < j \le \ell.$$

6.2 Unique Games Conjecture

Khot observed [Kho02c] that if the label sets in the Label-Cover instance are the same and the projections are permutations, then the hardness reductions could be simplified. He made the conjecture that Label-Cover is hard to approximate to any constant factor, restricted to such kind of instances.

Definition 6.10 (Unique Games Conjecture). *For every δ there exists a large enough R such that it is hard to distinguish between* Unique-Game *instances G have label size R from the following cases.*

- YES *case :* $\mathrm{OPT}(G) \geq 1 - \delta$
- NO *case :* $\mathrm{OPT} \leq \delta$

Starting with the work of Khot [Kho02c], it was shown that UGC, explains the lack of efficient approximation algorithms for a variety of problems (eg. Vertex Cover, MAX-CUT).

7

Long Code Bottleneck

The last two decades have seen tremendous progress in understanding the hardness of approximating constraint satisfaction problems. Despite this progress, the status of approximate coloring of constant colorable (hyper)graphs is not resolved and in fact, there is an exponential (if not doubly exponential) gap between the best known approximation algorithms and inapproximability results. The current best known approximation algorithms require at least $n^{\Omega(1)}$ colors to color a constant colorable (hyper)graph on n vertices while the best inapproximability results only rule out at best $(\log n)^{O(1)}$ (and in fact, in most cases, only $o(\log n)$) colors.

Given this disparity between the positive and negative results, it is natural to ask why current inapproximability techniques get stuck at the poly $\log n$ color barrier. The primary bottleneck in going past polylogarithmic colors is the use of the *long code*, a quintessential ingredient in almost all tight inapproximability results, since it was first introduced by Bellare, Goldreich & Sudan [BGS98].

Definition 7.1 (Long Code). *For a label $\ell \in \{0,1\}^r$, the long code encoding $A_\ell : \mathcal{F}_r \to \{0,1\}$ is given by*

$$\forall f \in \mathcal{F}_r, A_\ell(f) := f(\ell).$$

The long code, as the name suggests, is the most redundant encoding, wherein a r-bit Boolean string x is encoded by a 2^{2^r}-bit string which consists of the evaluation of

all Boolean functions on r bits at the point x. It is this doubly exponential blow-up of the long code which prevents the coloring inapproximability to go past the poly $\log n$ barrier.

7.1 Low-Degree Long Code

Recently, Barak et al. [BGH+12], while trying to understanding the tightness of the Arora-Barak-Steurer algorithm for Unique Games, introduced the *short code*, also called the *low-degree long code* [DG14]. The low-degree long code is a puncturing of the long code in the sense, that it contains only the evaluations of low-degree functions (opposed to all functions). Barak et al. [BGH+12] introduced the low-degree long code to prove exponentially stronger integrality gaps for Unique Games, and construct small set expanders whose Laplacians have many small eigenvalues,

Definition 7.2 (Low-Degree Long Code). *For $a \in \mathbb{F}_p^n$, the degree d long code for a is a function $\mathsf{LC}_d(a) : \mathsf{P}_d^n \to \mathbb{F}_p$ defined as*

$$\mathsf{LC}_d(a)(f) := f(a).$$

Note that for $d = (p-1)n$, this matches with the definition of the original long code over the alphabet $\mathbb{F}_p$.

Being a derandomization of the long code, one might hope to use the low-degree long code as a more size-efficient surrogate for the long code in inapproximability results. In fact, Barak et al. [BGH+12] used it obtain a more efficient version of the KKMO alphabet reduction [KKMO07] for Unique Games. However, using the low-degree long code towards improved reductions from Label Cover posed some challenges related to folding, and incorporating noise without giving up perfect completeness (which is crucial for results on coloring). Recently, Dinur & Guruswami [DG14] introduced a

very elegant set of techniques to adapt the long code based inapproximability results to low-degree long codes. Using these techniques, they proved (1) improved inapproximability results for gap-$(1, \frac{15}{16} + \varepsilon)$-4SAT for $\varepsilon = \exp(-2^{\Omega(\sqrt{\log \log N})})$ (long code based reductions show for $\varepsilon = 1/\text{poly} \log N$) and (2) hardness for a variant of approximate hypergraph coloring, with a gap of 2 and $\exp(2^{\Omega(\sqrt{\log \log N})})$ number of colors (where N is the number of vertices). It is to be noted that the latter is the first result to go beyond the logarithmic barrier for a coloring-type problem. However, the Dinur-Guruswami [DG14] results do not extend to standard (hyper)graph coloring hardness due to a multipartite structural bottleneck in the PCP construction, which we elaborate below.

As mentioned earlier, the two main contributions of Dinur-Guruswami [DG14] are (1) folding mechanism over the low-degree long code and (2) noise in the low-degree polynomials. The results of Bhattacharyya et al. [BKS+10] and Barak et al. [BGH+12] suggest that the product of d linearly independent affine functions suffices to work as noise for the low-degree long code setting (with degree = d) in the sense that it attenuates the contribution of large weight Fourier coefficients. However, this works only for PCP tests with imperfect completeness. Since approximate coloring results require perfect completeness, Dinur & Guruswami [DG14] inspired by the above result, develop a noise function which is the product of two random low-degree polynomials such that the sum of the degrees is at most d. This necessitates restricting certain functions in the PCP test to be of smaller degree which in turn requires the PCP tests to query two types of tables – one a low-degree long code of degree d and another a low-degree long code of smaller degree. Though the latter table is a part of the former, a separate table is needed since otherwise the queries will be biased to the small degree portion of the low-degree long code. This multipartite structure is what precludes them from extending their result for standard coloring results. (Clearly, if the query of the PCP tests straddles two tables, then the associated hypergraph is trivially 2-colorable.)

Building on the Dinur-Guruswami framework, in Chapter 9, by a simple test for low-degree long code, we show that it is quasi-NP-hard to color a 4-colorable 4-uniform hypergraph with $2^{2^{\Omega(\sqrt{\log \log n})}}$ colors.

7.2 Quadratic Label Cover

Both the Dinur-Guruswami and our results were obtained by modifying the innermost PCP verifier to work with the low-degree long code. Shortly thereafter, in a remarkable improvement, Khot & Saket [KS14a] showed that it is quasi-NP-hard to color a 2-colorable 12-uniform hypergraph with $2^{(\log n)^{\Omega(1)}}$ colors. They obtained this result by using an 12-query inner PCP verifier based on the quadratic code, ie., a low-degree long code with degree two. Since degree 2 functions can be represented as matrices, the quadratic code has an alternative simpler definition.

Our focus will be the case when the field $\mathbb{F}$ has characteristic 2. Let $\mathbb{F}^{m \times m}$ be the vector space of $m \times m$ matrices over the field $\mathbb{F}$.

Definition 7.3 (Quadratic Code). *The quadratic code of* $x \in \mathbb{F}^m$ *is a function* $A_x : \mathbb{F}^{m \times m} \to \mathbb{F}$ *defined as* $A_x(X) := \langle X, x \otimes x \rangle$.

However, to use a quadratic code based inner verifier, they needed an outer PCP verifier with a significantly stronger soundness guarantee than the standard outer PCP verifier obtained from parallel repetition of the PCP Theorem. In particular, they needed an outer PCP verifier, which in the soundness case, would not be satisfied by a short list of proofs even in *superposition*[*]. The construction of this outer PCP verifier with this stronger soundness guarantee is the main technical ingredient in the result of Khot & Saket [KS14a].

[*]We will not require the exact definition of *satisfying in superposition* for this note. See Theorem 7.4 for the details of the Khot-Saket outer PCP verifier.

Our reductions makes use of the following outer PCP verifier of Khot & Saket [KS14a]. As stated in the introduction, these instances have stronger soundness conditions which make them amenable for composition with a quadratic code based inner verifier.

Theorem 7.4 (Khot & Saket [KS14a, , Theorem 7.2]). *There is a quasi-polynomial time reduction from an instance of* 3-SAT *to a bi-regular instance* (U, V, E, Π) *of Label Cover such that*

- *Vertex sets* U *and* V *are bounded in size by* N.
- *The label sets are* $\mathbb{F}_2^{r\times r}, \mathbb{F}_2^{m\times m}$ *for* U *and* V *respectively.*
- *For* $e \in E$*, the map* $\pi^e : \mathbb{F}_2^{m\times m} \to \mathbb{F}_2^{r\times r}$ *is a linear transformation that maps symmetric matrices to symmetric matrices*[†]*. For an* $r \times r$ *matrix* X*,* $X \circ \pi^e$ *is the unique* $m \times m$ *matrix such that* $\langle X \circ \pi^e, Y\rangle = \langle X, \pi^e Y\rangle$.
- *For each vertex* $v \in V$*, there is a constraint* C_v *that is a a conjunction of homogeneous linear equations on the entries of the* $m \times m$ *matrix label.*
- $\delta \leq 2^{-\log^{1/3} N}$ *and* $k \geq (\log N)^{1/9}$.

The reduction satisfies:

1. *Completeness : If the* 3-SAT *instance is satisfiable then there is a labeling* $x_u \otimes x_u$ *for* $u \in U$ *and* $y_v \otimes y_v$ *for* $v \in V$ *such that*
   - *for each* $v \in V$*,* $y_v \in \mathbb{F}_2^m$ *has the* m^{th} *coordinate* 1 *and* $y_v \otimes y_v$ *satisfies the constraint* C_v*,*
   - *for each* $(u, v) \in E$*,* $\pi_{u,v}(y_v \otimes y_v) = x_u \otimes x_u$.

[†]The property that π maps symmetric matrices to symmetric matrices is easy to see from the proof of [KS14a, Theorem 7.2].

2. *Soundness : If the* 3-SAT *instance is not satisfiable then the following cannot hold: There are symmetric matrices* $M_u \in \mathbb{F}_2^{r\times r}, M_v \in \mathbb{F}_2^{m\times m}$ *for* $u \in U, v \in V$ *of rank* $\leq k$ *such that*

   - *for each* $v \in V$, $M_v \in \mathbb{F}_2^{m\times m}$ *has the* $(m,m)^{th}$ *coordinate* 1 *and* M_v *satisfies the constraint* C_v,
   - *for* δ *fraction of edges* e, $\pi_e(M_v) = M_u$.

3. *Smoothness : For any* $v \in V$ *and any symmetric non-zero matrix* M_v *with rank* $\leq k$, *over a random choice of an edge* e *incident on* v,

$$\Pr[\pi_e(M_v) = 0] \leq \delta/2.$$

8

Almost Coloring of Graphs

In this chapter, we describe our hardness results on almost graph coloring (joint work with Dinur, Harsha & Srinivasan [DHSV15]). Recall the discussion on graph coloring algorithms and hardness results from Section 1.3. Given a 3-colorable graph, best known algorithms give a $n^{0.19996}$-coloring [KT14] and it is known that finding a 4-coloring is NP-Hard [GK04]. Assuming UGC, Dinur, Mossel & Regev [DMR09], showed that, given an almost 3-colorable graph, it is hard to find a C-coloring for any constant C. Their exact result is as follows:

Theorem 8.1 (Dinur, Mossel & Regev [DMR09]). *There is a reduction from* Unique-Game *instances G with n vertices and label set $[R]$ to graphs $\mathcal{G}$ of size $n3^R$ such that*

- YES*: If G is an instance of* Unique-Game *with* $\mathrm{OPT}(G) \geq 1 - \varepsilon$ *then there is a subgraph of $\mathcal{G}$ with fractional size* $1 - \mathrm{poly}(\varepsilon)$ *that is 3-colorable.*
- NO*: If G is an instance of* Unique-Game *with soundness* $\mathrm{OPT}(G) \leq \delta$ *then $\mathcal{G}$ does not have any independent sets of fractional size* $O(1/\log(1/\delta))$.

For any constant C, taking δ to be a small enough constant will ensure that the chromatic number is $\geq C$ in the NO case. For getting hardness results with super-constant C, requires us to have sub-constant δ which depends on n. In UGC, the relation between the the label size R and soundness δ is not specified. But $R = \Omega(1/\delta)$, since a

random labeling to a Unique-Game instance, satisfies $O(1/R)$ fraction of the constraints. Assuming UGC, with $R = \text{poly}(1/\delta)$ and $\delta = 1/\text{poly}(\log n)$, will ensure that (1) in the NO case, the chromatic number is $\Omega(\log \log n)$ and (2) size of $\mathcal{G}$ is $\text{poly}(n)$.

Dinur & Shinkar [DS10] improved the analysis of the reduction, to show that there are no independent sets of fractional size $O(1/\text{poly}(\delta))$ in the NO case. Using the same assumption mentioned earlier, this implies hardness for chromatic number $\Omega(\text{poly}(\log n))$. However for these results to hold, the alphabet size R has to be $O(\log n)$.

In this chapter, we give a more efficient reduction, which ensures that the size of $\mathcal{G}$ remains small even when $R = 2^{2^{O(\sqrt{\log \log n})}}$. The reduction of Dinur, Mossel & Regev [DMR09], used the graph product described in Section 5.1 as a gadget. This is the reason why the size of $\mathcal{G}$ is $n3^R$. We replace this gadget by the derandomized graph product construction from Section 5.2. This ensures that the size of $\mathcal{G}$ is $n3^{\text{poly}_\delta(\log R)}$. Hence the reduction remains polynomial time, even when the alphabet size is much larger than $\log n$.

For getting the hardness result, we need to assume a conjecture similar to the Unique Games Conjecture with specific parameters.

Conjecture 8.2 ($(c(n), s(n), r(n))$-UG Conjecture). *It is NP-Hard to distinguish between unique label cover instances (U, V, E, R, Π) on n vertices and $R = \mathbb{F}_3^{r(n)}$ from the following cases:*

- *YES Case : There is a labeling and a set $S \subseteq V$ of size $(1 - c(n))|V|$ such that all edges between vertices in S are satisfied.*
- *NO Case : For any labeling, at most $s(n)$ fraction of edges are satisfied.*

Khot & Regev [KR08] proved that the Unique Games Conjecture implies that for any constants $c, s \in (0, 1/2)$ there is a constant r such that (c, s, r)-UG Conjecture

is true. We also require that the constraints of the Unique Games instance are full rank linear maps.

Definition 8.3 (Linear constraint). *A constraint $\pi : R \to L$ is a linear constraint of iff $R = L = \mathbb{F}_3^r$, and π is a linear map of rank r.*

Theorem 8.4. *There is a reduction from (c, s, r)-Unique Label Cover instances G with n vertices, label set $\mathbb{F}_3^r$ and linear constraints to graphs $\mathcal{G}$ of size $n3^{r^{O(\log 1/\mu)}}$ where $\mu =$ poly(s) such that*

- *If G belongs to the YES case of (c, s, r)-UG then there is a subgraph of $\mathcal{G}$ with fractional size $1 - c$ that is 3-colorable.*
- *If G belongs to the NO case of (c, s, r)-UG Conjecture then $\mathcal{G}$ does not have any independent sets of fractional size μ.*

Due to the efficiency of reduction, we are able to get hardness results even if the label cover instances had super-polylogarithmic sized label sets of size at most $2^{2^{O(\sqrt{\log \log n})}}$ while the reduction due to Dinur and Shinkar only worked if label set is of size $O(\log^c n)$ for some constant c. However we get this improvement only when soundness of the label cover $s(n) = 1/2^{O(\sqrt{\log \log n})}$. We remark that we can improve the conclusion if Theorem 5.5 can be proved even when $d = O(\log \log 1/\mu)$.

Corollary 8.5. *Let c, s, r be functions such that $s^{-1}(n), r(n) = 2^{O(\sqrt{\log \log n})}$. Assuming (c, s, r)-UG Conjecture on instances with linear constraints, given a graph on N vertices which has an induced subgraph of relative size $1 - c$ that is 3-colorable, no polynomial time algorithm can find an independent set of size* poly$(s(N))$.

8.1 Reduction

In this section we prove Theorem 8.4. Let $G = (U, V, R, E, \Pi)$ be a unique games cover instance with label set $R = \mathbb{F}_3^r$ and the constraints π are full rank linear transformations. We will construct a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = V \times \mathsf{P}_{r,2d}$, where d is a parameter to be fixed later. Let $T_{r,d}$ be the operator in Definition 5.7. There is an edge in $\mathcal{G}$ between (v, f) and (w, g) if there is a $u \in U$ such that $(u, v), (u, w) \in E$ and $T_{r,d}(f \circ \pi_{u,v}^{-1}, g \circ \pi_{u,w}^{-1}) > 0$, where $\pi_{u,v}$ is the full rank linear map that maps a label of v to label of u.

Lemma 8.6 (Completeness). *If G belongs to the YES case of (c, s, r)-UG Conjecture then $\mathcal{G}$ has a induced subgraph of relative size $1 - c$ that is 3-colorable.*

Proof. Suppose the label cover instance has a labeling $\ell : V \to \mathbb{F}_3^r$ and a set $S \subseteq V, |S| = (1 - c)|V|$, such that ℓ satisfies all the edges incident on vertices in S. We will show that $A_v(f) := f(\ell(v))$ for $v \in V$, is a 3-coloring for the induced subgraph of $\mathcal{G}$ on the set $S \times \mathsf{P}_{r,2d}$. For any $u \in U, v, w \in S$ having edges $(u, v), (u, w) \in E$, consider the edge $((v, f), (w, g)) \in \mathcal{E}$. The colors given to the end points are $f(\ell(v))$ and $g(\ell(w))$. Since $T_{r,d}(f \circ \pi_{u,v}^{-1}, g \circ \pi_{u,w}^{-1}) > 0$,

$$g \circ \pi_{u,w}^{-1} = f \circ \pi_{u,v}^{-1} + a(p^2 + 1) \text{ for some } p \in \mathsf{P}_d^r, a \in \{1, 2\}.$$

So $f(\ell(v)) = f \circ \pi_{u,v}^{-1}(\ell(u)) \neq g \circ \pi_{u,w}^{-1}(\ell(u)) = g(\ell(w))$. □

Lemma 8.7 (Soundness). *If G belongs to the NO case of (c, s, r)-UG Conjecture, $\mathcal{G}$ has an independent set of relative size μ and $d = O(\log 1/\mu)$ then $\mu \leq \text{poly}(s(n))$.*

Proof. Let $I_v : \mathsf{P}_{r,2d} \to \{0, 1\}$ be the indicator function of I restricted to vertices in $\mathcal{V}$ corresponding to $v \in V$. Let $J = \{v \in V : \mathbb{E}_{f \in \mathsf{P}_{r,2d}} I_v(f) \geq \mu/2\}$. Then we

have that $|J|/|V| \geq \mu/2$. For $v \in J$, let $L(v) = \{x \in \mathbb{F}_3^r : \mathrm{Inf}_x^{\leq k}(I_v) > \delta\}$ where $\delta = \mathrm{poly}(\mu), k = O(\log 1/\mu)$ are parameters from Theorem 5.5. Note that $|L(v)| \leq k/\delta$, since the sum of all degree k influences is at most k.

Claim 8.8. *Let $v, w \in J$ and $(u, v), (u, w) \in E$. Then there exists $a \in L(v), b \in L(w)$ such that $\pi_{u,v}(a) = \pi_{u,w}(b)$.*

Proof. Let $A(f) := I_v(f \circ \pi_{u,v}^{-1})$, $B(g) := I_w(g \circ \pi_{u,w}^{-1})$. Since I is an independent set, if $(v, f \circ \pi_{u,v}^{-1}), (w, g \circ \pi_{u,w}^{-1}) \in I$, then $T_{r,d}(f \circ \pi_{u,v}^{-1}, g \circ \pi_{u,w}^{-1}) = 0$, which gives that

$$\langle A, T_{r,d} B \rangle = 0 \tag{8.1.1}$$

From Theorem 5.5, there is some $c \in \mathbb{F}_3^r$ such that $\mathrm{Inf}_c^{\leq k}(A), \mathrm{Inf}_c^{\leq k}(B) > \delta$, which gives that $\pi_{u,v}^{-1}(c) \in L(v)$ and $\pi_{u,w}^{-1}(c) \in L(w)$. □

Now consider the randomized partial labeling L' to G, where for $v \in J$, $L'(v)$ is chosen randomly from $L(v)$ and for $u \in U$, choose a random neighbor $w \in J$ (if it exists), a random label $a \in L(w)$ and set $L(u) = \pi_{u,w}^{-1}(a)$. For any $v \in J$, any edge (u, v), the probability of it being satisfied by L' is $\mu^2/k^2 = \mathrm{poly}(\mu)$, because of Claim 8.8.

□

Proof of Theorem 8.4. The size of $\mathcal{G}$ denoted by N is at most $n3^{r^{O(d)}}$. Substituting $r = 2^{O(\sqrt{\log \log n})}, d = \log 1/\mu \leq O(\sqrt{\log \log n})$, we get that $N = \mathrm{poly}(n)$ and hence the reduction is polynomial time. □

9

Approximate Hypergraph Coloring

In this chapter, we give improved hardness results for the Approximate-k-Hypergraph-Coloring(c, C), for small constant c, and exponentialy better values of C than was previously known.

In Section 9.1, we prove the following theorem (joint work with Guruswami, Håstad, Harsha & Srinivasan [GHH+14]).

Theorem 9.1 (Approximate-3-Hypergraph-Coloring$(3, C)$). *Assuming* NP *does not have* $n^{2^{O(\log\log n/\log\log\log n)}}$ *time algorithms, there is no polynomial time algorithm which, when given as input a* 3*-uniform hypergraph* H *on* N *vertices can distinguish between the following:*

- H *is* 3 *colorable.*
- H *has no independent set of size* $N/2^{O(\log\log N/\log\log\log N)}$.

Prior to this result, the best inapproximability result for O(1)-colorable 3-uniform hypergraphs were as follows: Khot [Kho02b] showed that it is quasi-NP-hard to color a 3-colorable 3-uniform hypergraphs with $(\log\log N)^{1/9}$ colors and Dinur, Regev & Smyth [DRS05] showed that it is quasi-NP-hard to color a 2-colorable 3-uniform hypergraphs with $(\log\log N)^{1/3}$ colors (observe that $2^{O(\log\log N/\log\log\log N)}$ is exponentially larger than $(\log\log N)^{\Omega(1)}$). For 2-colorable 3-uniform hypergraphs, the result of

Dinur et al. [DRS05] only rules out colorability by $(\log \log N)^{\Omega(1)}$, while a recent result due to Khot & Saket [KS14b] shows that it is hard to find a δN-sized independent set in a given N-vertex 2-colorable 3-uniform hypergraph assuming the d-to-1 games conjecture. Our improved inapproximability result is obtained by adapting Khot's proof to the low-degree long code using the new noise function over $\mathbb{F}_3$. We remark that this result is not as strong as the next two, since the starting point is a multilayered smooth label cover instance instead of just label cover, which causes a blow-up in size and a corresponding deterioration in the parameters.

In Section 9.2, we show that the quadratic outer PCP verifier of Khot & Saket [KS14a] can in fact be combined with a 8-query test for the quadratic code based on the Guruswami et al. [GHH+14] inner PCP verifier to obtain a hardness result for 2-colorable 8-uniform hypergraphs. More precisely, we show the following.

Theorem 9.2 (Approximate-8-Hypergraph-Coloring$(2, C)$). *For every constant $\varepsilon > 0$ there is a reduction from* 3-SAT *on m variables to a 8-uniform hypergraph $\mathcal{G}$ on n vertices such that the following holds:*

1. *The running time of the reduction is $m^{\text{poly}(\log m)}$.*
2. *YES Case: If the* 3-SAT *instance is satisfiable then $\mathcal{G}$ is 2-colorable.*
3. *NO Case: If the* 3-SAT *instance is unsatisfiable then $\mathcal{G}$ does not have an independent set of relative size $2^{-(\log n)^{\frac{1}{20}-\varepsilon}}$.*

In Section 9.3, we use the technique of Guruswami et al. [GHH+14] for reducing the uniformity from 8 to 4 at the cost of increasing the number of colors from 2 to 4. We note that a similar trick can be performed in our setting to obtain the following result.

Theorem 9.3 (Approximate-4-Hypergraph-Coloring$(4, C)$). *For every constant $\varepsilon > 0$ there is a reduction from* 3-SAT *on m variables to a 4-uniform hypergraph $\mathcal{G}$ on n vertices such that the following holds:*

1. *The running time of the reduction is* $m^{\text{poly}(\log m)}$.

2. *YES Case: If the* 3-SAT *instance is satisfiable then* $\mathcal{G}$ *is* 4*-colorable.*

3. *NO Case: If the* 3-SAT *instance is unsatisfiable then* $\mathcal{G}$ *does not have an independent set of relative size* $2^{-(\log n)^{\frac{1}{20}-\varepsilon}}$.

We remark that the analyses of the inner verifier in Section 9.2 and Section 9.3 are simpler than the analyses of the corresponding inner verifiers in Guruswami et al. [GHH+14] and Khot & Saket [KS14a] results. Furthermore, in the language of covering complexity [*] introduced by Guruswami, Håstad & Sudan [GHS02], (the proof of) Theorem 9.3 demonstrates a Boolean CSP on 4 variables for which it is quasi-NP-hard to distinguish between covering number of 2 vs. $(\log n)^{\Omega(1)}$.

9.1 3-Colorable 3-Uniform Hypergraphs

This construction is an adaptation of Khot's construction [Kho02b] to the low-degree long code setting. We prove the theorem by a reduction from 3-SAT via the instances of the multilayered label cover problem obtained in Theorem 6.9. Let r, ℓ, η be parameters that will be determined later and let $I(\varphi)$ be an instance of the r-repeated ℓ-layered η-smooth label cover instance with constraint graph $G = (V_0, \ldots, V_{\ell-1}, \{E_{ij}\}_{0 \le i < j < \ell})$ obtained from the 3-SAT instance φ. We use the results from the preliminaries with the field set to $\mathbb{F}_3 = \{0, 1, 2\}$. For every layer i and every vertex $v \in V_i$, let $\{c_1, \cdots c_{(T+\ell-i)r}\}$ be the clauses corresponding to v where $T = \lceil l/\eta \rceil$ as in Definition 6.8. We construct polynomials $\{p_1, \cdots p_{(T+\ell-i)r}\}$ of degree at most 6 over $\mathbb{F}_3$ such that p_j depends only on variables in c_j with the following properties. Let $a \in \mathbb{F}_3^3$. If $a \notin \{0,1\}^3$ then $p_j(a) \neq 0$. Otherwise $p_j(a) = 0$ iff $c_j(a) = 1$. For a degree parameter d that we will

[*] The covering number of a CSP is the minimal number of assignments to the vertices so that each hyperedge is covered by at least one assignment.

determine later, for each vertex v define the subspace J_v as follows:

$$J_v := \left\{ \sum_i q_i p_i : q_i \in \mathsf{P}_{m_v,2d-6} \right\} \text{ where } m_v := m_i = 3(T+\ell-i)r + ir.$$

We now define the hypergraph H produced by the reduction. The vertices of H — denoted $V(H)$ — are obtained by replacing each $v \in G$ by a block $\mathcal{B}_v$ of $N_v := |\mathsf{P}_{m_v,2d}/J_v|$. vertices, which we identify with elements of $\mathsf{P}_{m_v,2d}/J_v$. Let N denote $|V(H)| = \sum_v N_v$.

We think of a 3-coloring of $V(H)$ as a map from $V(H)$ to $\mathbb{F}_3$. Given a coloring $A : V(H) \to 3$, we denote by $A_v : \mathsf{P}_{m_v,2d}/J_v \to \mathbb{F}_3$ the restriction of A to the block $\mathcal{B}_v$. Let $A'_v : \mathsf{P}_{m_v,2d} \to 3$ denote the lift of A_v as defined in Fact 1.

The (weighted) edge set $E(H)$ of H is specified implicitly by the following PCP verifier.

3-Color 3-Uniform Test(d):

1. Choose two layers $0 \leq i < j < \ell$ uniformly at random and then choose a uniformly random edge $(u, v) \in E_{ij}$. Let π denote $\pi_{uv} : \mathbb{F}_3^{m_u} \to \mathbb{F}_3^{m_v}$.

2. Choose $p \in \mathsf{P}_{m_u,d}$, $g \in \mathsf{P}_{m_u,2d}$ and $f \in \mathsf{P}_{m_v,2d}$ independently and uniformly at random and let $g' := p^2 + 1 - g - f \circ \pi$.

3. Accept if and only if $A'_v(f), A'_u(g), A'_u(g')$ are not all equal.

The hyperedges in the 3-uniform case straddle both sides of the corresponding edge (u, v) in the label cover instance. Hence, if constructed from the bipartite label cover, the corresponding 3-uniform hypergraph will also be bipartite and hence always 2-colorable irrespective of the label cover instance. Using the multilayered construction gets around this problem.

Lemma 9.4 (Completeness). *If $\varphi \in$ 3-SAT, then there is proof $A : V(H) \to \mathbb{F}_3$ which the verifier accepts with probability* 1. *In other words, the hypergraph H is* 3*-colorable.*

Proof. Since $\varphi \in$ 3-SAT, Theorem 6.9 tells us that there are labelings $L_i : V_i \to \{0,1\}^{m_i}$ for $0 \leq i < \ell$ which satisfy all the constraints in $I(\varphi)$. For $\forall i, v \in V_i$, we set $A_v : \mathsf{P}_{m_v,2d}/J_v \to \mathbb{F}_3$ such that its lift $A'_v = \mathsf{LC}_{2d}(L_i(v))$. This is possible since A'_v is folded over J_v. For any edge (u,v) between layers i,j, with labels $L_i(u) = a, L_j(v) = b$ such that $\pi(a) = b$, $(A'_v(f), A'_u(g), A'_u(g')) = (f(b), g(a), g'(a))$. The lemma follows by observing that $g'(a) + g(a) + f(b) \neq 0$ always (since $p^2(a) + 1 \neq 0$). □

Lemma 9.5 (Soundness). *Let $\ell = 32/\delta^2$. If $\varphi \notin$ 3-SAT and H contains a independent set of size $\delta|V(H)|$, then*

$$\delta^5/2^9 \leq 2^{-\Omega(r)} \cdot 3^d + \eta \cdot 3^d + \exp(-3^{\Omega(d)}).$$

Proof. Let $A : V(H) \to \{0,1\}$ be the characteristic function of the independent set of fractional size exactly δ. We have that $\forall v, \mathbb{E}_{g \in \mathsf{P}_{m_v,2d}/J_v}[A_v(g)] = \mathbb{E}_{g \in \mathsf{P}_{m_v,2d}}[A'_v(g)]$ where A'_v is the lift of A_v. Define

$$Q(u,v) := \mathop{\mathbb{E}}_{f,g,p}\left[A'_v(f)A'_u(g)A'_u(p^2 + 1 - f \circ \pi - g)\right].$$

Observe that $\mathbb{E}_{i,j,u,v}[Q(u,v)] = 0$ as A corresponds to an independent set. Using Lemma 3.2, we have the following Fourier expansion of Q:

$$Q(u,v) = \sum_{\alpha,\beta,\gamma} \widehat{A}'_v(\alpha)\widehat{A}'_u(\beta)\widehat{A}'_u(\gamma) \mathop{\mathbb{E}}_{f,g,p}\left[\chi_\alpha(f)\chi_\beta(g)\chi_\gamma(p^2 + 1 - f \circ \pi - g)\right], \tag{9.1.1}$$

where the summation is over $\alpha \in \Lambda_{m_v,2d}$, $\beta, \gamma \in \Lambda_{m_u,2d}$ and Λ is as defined in Lemma 3.2. From the orthonormality of characters, the non-zero terms satisfy $\beta = \gamma$

and $\alpha = \pi_3(\beta)$. Substituting in (9.1.1), we get

$$Q(u,v) = \sum_{\beta} \underbrace{\widehat{A}'_u(\beta)^2 \widehat{A}'_v(\pi_3(\beta)) \mathop{\mathbb{E}}_{p} \left[\chi_\beta(p^2+1)\right]}_{\xi_{u,v}(\beta)}. \tag{9.1.2}$$

Claim 9.6. *If $\ell = 32/\delta^2$, there exists layers $0 \le i < j < \ell$ such that $\mathbb{E}_{(u,v)\in E_{ij}}[\xi_{u,v}(0)] \ge \delta^5/2^9$.*

Proof. Since A' has fractional size δ, there exists a set S of vertices of fractional size $\delta/2$ such that $\forall v \in S, \widehat{A}'_v(0) = \mathbb{E}_f[A'_v(f)] \ge \delta/2$. Furthermore, there exists $\delta\ell/4$ layers, in which the fractional size of $S_i := S \cap V_i$ in layer V_i is at least $\delta/4$. Since $\ell = 32/\delta^2$, we obtain from Theorem 6.9 that there exists layers i, j such that the fraction of edges in E_{ij} between S_i and S_j is at least $\delta' = \delta^2/64$. From above, we have that

$$\mathop{\mathbb{E}}_{(u,v)\in E_{ij}}[\xi_{u,v}(0)] \ge \delta' \cdot (\delta/2)^3 \ge \delta^5/2^9. \quad \square$$

For the rest of the proof, layers i, j will be fixed as given by Claim 9.6. To analyze the expression in (9.1.2), we consider the following breakup of $\Lambda_{m_i,2d} \setminus \{0\}$ for every $(u,v) \in E_{ij}$: $\textsc{far} := \{\beta \in \Lambda_{m_i,2d} : \Delta(\beta, (\mathsf{P}_{m_i,2d})^\perp) \ge 3^{d/2}\}$, $\textsc{near}_1 := \{\beta \in \Lambda_{m_i,2d} \setminus \textsc{far} : \beta \ne 0 \text{ and } \pi_3(\beta) \notin (\mathsf{P}_{m_v,2d})^\perp\}$ and $\textsc{near}_0 := \{\beta \in \Lambda_{m_i,2d} \setminus \textsc{far} : \beta \ne 0 \text{ and } \pi_3(\beta) \in (\mathsf{P}_{m_v,2d})^\perp\}$. In Claims 9.7, 9.8 and 9.9, we bound the absolute values of the sum of $\mathbb{E}_{u,v}[\xi_{u,v}(\beta)]$ for β in $\textsc{far}$, $\textsc{near}_0$ and $\textsc{near}_1$ respectively.

Claim 9.7. $\left|\mathbb{E}_{(u,v)\in E_{ij}}\left[\sum_{\beta\in\textsc{far}} \xi_{u,v}(\beta)\right]\right| \le \exp(-3^{\Omega(d)})$.

Claim 9.8. $\left|\mathbb{E}_{(u,v)\in E_{ij}}\left[\sum_{\beta\in\textsc{near}_1} \xi_{u,v}(\beta)\right]\right| \le 2^{-\Omega(r)} \cdot 3^d$.

Claim 9.9. $\left|\mathbb{E}_{(u,v)\in E_{ij}}\left[\sum_{\beta\in\textsc{near}_0} \xi_{u,v}(\beta)\right]\right| \le \eta \cdot 3^d$.

Combined with Claim 9.6, this exhausts all terms in the expansion (9.1.2). Lemma 9.5 now follows from Claims 9.6–9.9. $\square$

We now proceed to the proofs of Claims 9.7, 9.8 and 9.9.

Proof of Claim 9.7.

$$\left|\mathop{\mathbb{E}}_{(u,v)\in E_{ij}}\left[\sum_{\beta\in\textsc{far}}\xi_{u,v}(\beta)\right]\right| \le \mathop{\mathbb{E}}_{(u,v)\in E_{Ij}}\left[\sum_{\beta\in\textsc{far}}|\widehat{A'_u}(\beta)|^2\cdot|\widehat{A'_v}(\pi_3(\beta))|\cdot\left|\mathop{\mathbb{E}}_{p}\left[\omega^{\langle\beta,p^2+1\rangle}\right]\right|\right].$$

The quantity $\langle\beta,p^2\rangle$ is analyzed in Section 4.3. Let z be a uniformly random $\mathbb{F}_3$ element. By Lemmas 4.4 and 4.7, we get that the statistical distance between the distributions of $\langle\beta,p^2+1\rangle$ and z is $\exp(-3^{\Omega(d)})$. Since the $\mathbb{E}_z[\omega^z] = 0$, we have that $\left|\mathbb{E}_p\left[\omega^{\langle\beta,p^2+1\rangle}\right]\right| \le \exp(-3^{\Omega(d)})$. The claim follows since $\left|\widehat{A'_v}(\alpha)\right| \le 1$ for any α and $\sum_\beta|\widehat{A'_u}(\beta)|^2 \le 1$. □

Proof of Claim 9.8. It suffices to bound the following for proving the claim.

$$\begin{aligned}
&\mathop{\mathbb{E}}_{(u,v)\in E_{ij}}\left[\sum_{\beta\in\textsc{near}_1}|\widehat{A'_u}(\beta)|^2\cdot|\widehat{A'_v}(\pi_3(\beta))|\right]\\
&\le \mathop{\mathbb{E}}_{(u,v)\in E_{ij}}\left[\sqrt{\sum_{\beta\in\textsc{near}_1}|\widehat{A'_u}(\beta)|^2\cdot|\widehat{A'_v}(\pi_3(\beta))|^2}\sqrt{\sum_{\beta\in\textsc{near}_1}|\widehat{A'_u}(\beta)|^2}\right] \qquad [\text{ by Cauchy-Schwarz }]\\
&\le \sqrt{\mathop{\mathbb{E}}_{(u,v)\in E_{ij}}\left[\sum_{\beta\in\textsc{near}_1}|\widehat{A'_u}(\beta)|^2\cdot|\widehat{A'_v}(\pi_3(\beta))|^2\right]} \qquad [\text{ by Jensen's inequality }].
\end{aligned}$$

We bound the above using a Fourier decoding argument. For every vertex $v \in V_i \cup V_j$, pick a random β according to $|\widehat{A'_v}(\beta)|^2$ (note $\sum_\beta|\widehat{A'_v}(\beta)|^2 \le 1$) and assign a random labeling to v from the support of β. It is easy to see that the fraction of edges that are satisfied by this labeling is lower bounded by the LHS of the expression bellow. We get the inequality using the soundness of the multilayered label cover from Theorem 6.9.

$$\frac{1}{3^d}\mathop{\mathbb{E}}_{(u,v)\in E_{ij}}\left[\sum_{\beta\in\textsc{near}_1}|\widehat{A'_v}(\pi_3(\beta))|^2|\widehat{A'_u}(\beta)|^2\right] \le 2^{-\Omega(r)}. \quad \square$$

Proof of Claim 9.9. We bound this sum using the smoothness property of the label cover instance.

$$\mathop{\mathbb{E}}_{(u,v)\in E_{ij}}\left[\sum_{\beta\in \textsc{near}_0}|\widehat{A'_u}(\beta)|^2\cdot|\widehat{A'_v}(\pi_3(\beta))|\right]\leq \mathop{\mathbb{E}}_{u\in V_i}\left[\sum_{\beta\notin \textsc{far}\cup\{0\}}\Pr_{v:(u,v)\in E_{ij}}\left[\pi_3(\beta)\in(\mathsf{P}_{2d}^{m_v})^{\perp}\right]\cdot|\widehat{A'_u}(\beta)|^2\right].$$

We now argue that for every u and $\beta\notin \textsc{far}\cup\{0\}$, $\Pr_{(u,v)\in E_{ij}}\left[\pi_3(\beta)\in(\mathsf{P}_{m_v,2d})^{\perp}\right]$ is at most $3^d\cdot\eta$. This combined with the fact that $\sum_\beta|\widehat{A'_u}(\beta)|^2\leq 1$ yields the claim. For every $u\in V_i$ and β such that $0\neq|\operatorname{support}(\beta)|=\Delta(\beta,(\mathsf{P}_{m_u,2d})^{\perp})\leq 3^{d/2}$, by the smoothness property (Theorem 6.9), we have that with probability at least $1-3^d\eta$, we have

$$\forall a\neq a'\in\operatorname{support}(\beta),\pi(a)\neq\pi(a'). \tag{9.1.3}$$

When (9.1.3) holds, we have $\pi_3(\beta)\neq 0$. Now since $|\operatorname{support}(\pi_3(\beta))|\leq|\operatorname{support}(\beta)|\leq 3^{d/2}$ and non-zero polynomials in $(\mathsf{P}_{m_v,2d})^{\perp}$ has support at least 3^d, we can further conclude that $\pi_3(\beta)\notin(\mathsf{P}_{m_v,2d})^{\perp}$ whenever (9.1.3) holds. □

Proof of Theorem 9.1. Given the completeness (Lemma 9.4) and soundness (Lemma 9.5), we only need to fix parameters. Let n be the size of the 3-SAT instance and N the size of the hypergraph produced by the reduction.

Let $d=C_1\log\log(1/\delta')$, $\eta=(\delta')^5/C_2$ and $r=C_3\log(1/\delta')$ for large enough constants C_1,C_2,C_3 and parameter $\delta'\in(0,1)$ to be determined shortly. By Lemma 9.5, if H has an independent set of size δN, then $\delta^5/2^9\leq 3^d\cdot 2^{-\Omega(r)}+3^d\cdot\eta+\exp(-3^{\Omega(d)})<(\delta')^5/2^9$ for large enough C_1,C_2,C_3. Hence, H has no independent sets of $\delta' N$.

The hypergraph H produced by the reduction is of size $N=\ell n^{(1+1/\eta)\ell r}3^{((1+1/\eta)\ell r)^{O(d)}}$. Setting $\ell=C_4/(\delta')^2$ and $\log(1/\delta')=\Theta(\log\log n/\log\log\log n)$, we get that $N=n^{2^{O(\log\log n/\log\log\log n)}}$. Since $\log\log n=\Theta(\log\log N)$, we also get that $1/\delta'=2^{\Theta(\log\log N/\log\log\log N)}$. This completes the proof of Theorem 9.1. □

9.2 2-Colorable 8-Uniform Hypergraphs

In this section we prove Theorem 9.2. Our reduction starts from the label cover instances given by Theorem 7.4. Let (U, V, E, Π) be an instance of the label cover. We will construct a hypergraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. For $v \in V$, let $\mathcal{H}_v \subseteq \mathbb{F}_2^{m\times m}$ be the dual of the subspace of the set matrices that are symmetric and which satisfies the constraint C_v. The set of vertices $\mathcal{V}$ will be the same as $V \times \left(\mathbb{F}_2^{m\times m}/\mathcal{H}_v\right)$. Any 2-coloring of $\mathcal{G}$ is a collection of functions $A'_v : \mathbb{F}_2^{m\times m}/\mathcal{H}_v \to \{0,1\}$ for $v \in V$. For any such function, we can uniquely extend it to get $A_v : \mathbb{F}_2^{m\times m} \to \{0,1\}$ which is constant over cosets of $\mathcal{H}_v$. This method is called folding and it ensures that A_v satisfies the following: if $\alpha \in \mathbb{F}_2^{m\times m}$ is such that $\widehat{A}_v(\alpha)$ is non-zero, then α is symmetric and satisfies C_v.

The set of edges $\mathcal{E}$ will be defined by the test mentioned below, which checks whether a supposed 2-coloring $A'_v : \mathbb{F}_2^{m\times m}/\mathcal{H}_v \to \{0,1\}$ is valid. There is an edge in $\mathcal{E}$ between any set of vertices in $\mathcal{V}$ that are queried together by the test. The test will be querying the extended functions A_v at matrices in $\mathbb{F}_2^{m\times m}$ instead of A'_v. So a query to A_v at $X \in \mathbb{F}_2^{m\times m}$ corresponds to a query to A'_v at the coset of $\mathcal{H}_v$ that contains X.

2-Colorable 8-Uniform Test $\mathcal{T}_{2,8}$:

1. Choose $u \in U$ uniformly at random and $v, w \in V$ uniformly and independently at random from the neighbors of u. Let $\pi, \sigma : \mathbb{F}_2^{m\times m} \to \mathbb{F}_2^{r\times r}$ be the projections corresponding to the edges $(u, v), (u, w)$ respectively. Uniformly and independently at random choose $X_1, X_2, Y_1, Y_2 \in \mathbb{F}_2^{m\times m}$ and $\overline{x}, \overline{y}, \overline{z}, \overline{x}', \overline{y}', \overline{z}' \in \mathbb{F}_2^m$ and $F \in \mathbb{F}_2^{r\times r}$. Let $\overline{e}_m \in \mathbb{F}_2^m$ be the vector with only the m^{th} entry 1 and the rest is 0.

2. Accept if and only if the following 8 values are not all equal :

$$A_v(X_1) \quad A_v(X_3) \quad \text{where } X_3 := X_1 + \overline{x} \otimes \overline{y} + F \circ \pi$$

$$A_v(X_2) \quad A_v(X_4) \quad \text{where } X_4 := X_2 + (\overline{x} + \overline{e}_m) \otimes \overline{z} + F \circ \pi$$

$$A_w(Y_1) \quad A_w(Y_3) \quad \text{where } Y_3 := Y_1 + \overline{x}' \otimes \overline{y}' + F \circ \sigma + \overline{e}_m \otimes \overline{e}_m$$

$$A_w(Y_2) \quad A_w(Y_4) \quad \text{where } Y_4 := Y_2 + (\overline{x}' + \overline{e}_m) \otimes \overline{z}' + F \circ \sigma + \overline{e}_m \otimes \overline{e}_m$$

9.2.1 YES Case

Let $\overline{y}_v \otimes \overline{y}_v$ for $v \in V$ and $\overline{x}_u \otimes \overline{x}_u$ for $u \in U$ be a perfectly satisfying labeling of the label cover instance. That is, for every $(u, v) \in E, \pi_{u,v}(\overline{y}_v \otimes \overline{y}_v) = \overline{x}_u \otimes \overline{x}_u$. Such a labeling is guaranteed by the YES instance of label cover, with the additional property that the m^{th} coordinate of $\overline{y}_v$ is 1. Consider the following 2-coloring of $\mathcal{G}$: for each $v \in V$, $A_v(X) := \langle X, \overline{y}_v \otimes \overline{y}_v \rangle$. Note that such a function is constant over cosets of $\mathcal{H}_v$. Let

$$x_1 := \langle X_1, \overline{y}_v \otimes \overline{y}_v \rangle \qquad x_2 := \langle X_2, \overline{y}_v \otimes \overline{y}_v \rangle$$

$$y_1 := \langle Y_1, \overline{y}_w \otimes \overline{y}_w \rangle \qquad y_2 := \langle Y_2, \overline{y}_w \otimes \overline{y}_w \rangle$$

and $f := \langle F, \overline{x}_u \otimes \overline{x}_u \rangle$. Note that $\langle F \circ \pi_{u,v}, \overline{y}_v \otimes \overline{y}_v \rangle = \langle F, \pi_{u,v}(y_v \otimes y_v) \rangle = \langle F, \overline{x}_u \otimes \overline{x}_u \rangle$, and $\langle \overline{e}_m \otimes \overline{e}_m, \overline{y}_v \otimes \overline{y}_v \rangle = \langle \overline{e}_m, \overline{y}_v \rangle = 1$. Using these, the assignments to the 8 query locations are:

$$x_1 \qquad x_1 + \langle \overline{y}_v, \overline{x} \rangle \langle \overline{y}_v, \overline{y} \rangle + f$$

$$x_2 \qquad x_2 + (\langle \overline{y}_v, \overline{x} \rangle + 1) \langle \overline{y}_v, \overline{z} \rangle + f$$

$$y_1 \qquad y_1 + \langle \overline{y}_w, \overline{x}' \rangle \langle \overline{y}_w, \overline{y}' \rangle + f + 1$$

$$y_2 \qquad y_2 + (\langle \overline{y}_w, \overline{x}' \rangle + 1) \langle \overline{y}_w, \overline{z}' \rangle + f + 1$$

It is easy to see at least one of the 4 rows are always not equal. Hence A is a valid 2-coloring of $\mathcal{G}$.

9.2.2 NO Case

Suppose the reduction was applied to a NO instance of label cover. Let k and δ be the parameters specified by Theorem 7.4.

Lemma 9.10. *If there is an independent set in $\mathcal{G}$ of relative size s then*

$$s^8 \leq \delta + \frac{1}{2^{k/2+1}}.$$

Proof. The proof of the lemma is similar to Section 8.2 in Khot & Saket [KS14a]. Consider any set $A \subseteq \mathcal{V}$ of fractional size s. For every $v \in V$, let $A_v : \mathbb{F}_2^{m\times m} \to \{0,1\}$ be the indicator function that is extended such that it is constant over cosets of $\mathcal{H}_v$. A is an independent set if and only if

$$\Theta := \mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{X_i,Y_i\in\mathcal{T}_{2,8}} \prod_{i=1}^{4} A_v(X_i)A_w(Y_i) = 0. \tag{9.2.1}$$

Now we do the Fourier expansion and take expectations over X_1, X_2, Y_1, Y_2 to obtain the following:

$$\begin{aligned}\Theta = \mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\alpha_1,\alpha_2\\ \beta_1,\beta_2\in\mathbb{F}_2^{m\times m}}} \mathop{\mathbb{E}}_{F,\overline{x},\overline{x}'} \Bigg[&\widehat{A}_v(\alpha_1)^2 \mathop{\mathbb{E}}_{\overline{y}} \left[\chi_{\alpha_1}(\overline{x}\otimes\overline{y})\right] \chi_{\alpha_1}(F\circ\pi) \\ &\widehat{A}_v(\alpha_2)^2 \mathop{\mathbb{E}}_{\overline{z}} \left[\chi_{\alpha_2}((\overline{x}+\overline{e}_m)\otimes\overline{z})\right] \chi_{\alpha_2}(F\circ\pi) \\ &\widehat{A}_w(\beta_1)^2 \mathop{\mathbb{E}}_{\overline{y}'} \left[\chi_{\beta_1}(\overline{x}'\otimes\overline{y}')\right] \chi_{\beta_1}(F\circ\sigma)\chi_{\beta_1}(\overline{e}_m\otimes\overline{e}_m) \\ &\underbrace{\widehat{A}_w(\beta_2)^2 \mathop{\mathbb{E}}_{\overline{z}'} \left[\chi_{\beta_2}((\overline{x}'+\overline{e}_m)\otimes\overline{z}')\right] \chi_{\beta_2}(F\circ\sigma)\chi_{\beta_2}(\overline{e}_m\otimes\overline{e}_m)}_{=:\mathrm{Term}_{u,v,w}(\alpha_1,\alpha_2,\beta_1,\beta_2)}\Bigg]\end{aligned}$$

Note that since $F \in \mathbb{F}_2^{r\times r}$ is chosen uniformly at random,

$$\mathop{\mathbb{E}}_{F} \chi_{\alpha_1}(F\circ\pi)\chi_{\alpha_2}(F\circ\pi)\chi_{\beta_1}(F\circ\sigma)\chi_{\beta_2}(F\circ\sigma) = \mathop{\mathbb{E}}_{F}(-1)^{\langle\pi(\alpha_1+\alpha_2),F\rangle+\langle\sigma(\beta_1+\beta_2),F\rangle}$$

is zero unless $\pi(\alpha_1+\alpha_2) = \sigma(\beta_1+\beta_2)$. Let $\nu(\alpha) := (-1)^{\langle\alpha,\overline{e}_m\otimes\overline{e}_m\rangle}$. Now taking expectations over $\overline{x},\overline{y},\overline{z},\overline{x}',\overline{y}',\overline{z}'$, and noting that $\langle\alpha, x\otimes y\rangle = \langle\alpha x, y\rangle$, we obtain

$$\begin{aligned}\mathrm{Term}_{u,v,w}(\alpha_1,\alpha_2,\beta_1,\beta_2) = (-1)^{\nu(\beta_1+\beta_2)}\widehat{A}_v(\alpha_1)^2\widehat{A}_v(\alpha_2)^2\widehat{A}_w(\beta_1)^2\widehat{A}_w(\beta_2)^2 \\ \Pr_{\overline{x}}[\alpha_1\overline{x}=0\wedge\alpha_2\overline{x}=\alpha_2 e_m]\cdot \\ \Pr_{\overline{x}'}[\beta_1\overline{x}'=0\wedge\beta_2\overline{x}'=\beta_2 e_m]\end{aligned} \tag{9.2.2}$$

when $\pi(\alpha_1+\alpha_2) = \sigma(\beta_1+\beta_2)$ and 0 otherwise. Define:

$$\Theta_0 = \mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\mathsf{rank}(\alpha_1+\alpha_2),\mathsf{rank}(\beta_1+\beta_2)\le k\\ \pi(\alpha_1+\alpha_2)=\sigma(\beta_1+\beta_2)\\ \nu(\beta_1+\beta_2)=0}} \mathrm{Term}_{u,v,w}(\alpha_1,\alpha_2,\beta_1,\beta_2) \tag{9.2.3}$$

$$\Theta_1 = \mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\mathsf{rank}(\alpha_1+\alpha_2),\mathsf{rank}(\beta_1+\beta_2)\le k\\ \pi(\alpha_1+\alpha_2)=\sigma(\beta_1+\beta_2)\\ \nu(\beta_1+\beta_2)=1}} \mathrm{Term}_{u,v,w}(\alpha_1,\alpha_2,\beta_1,\beta_2) \tag{9.2.4}$$

$$\Theta_2 = \mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\max\{\mathsf{rank}(\alpha_1+\alpha_2),\mathsf{rank}(\beta_1+\beta_2)\}>k\\ \pi(\alpha_1+\alpha_2)=\sigma(\beta_1+\beta_2)}} \mathrm{Term}_{u,v,w}(\alpha_1,\alpha_2,\beta_1,\beta_2) \tag{9.2.5}$$

We lower bound Θ_0 by s^8, upper bound $|\Theta_1|$ by δ and $|\Theta_2|$ by $1/2^{k/2+1}$ below. Along with (9.2.1), this will prove Lemma 9.10.

Lower Bound on Θ_0: Note that all terms in Θ_0 are positive. Now consider the term corresponding to $\alpha_1=\alpha_2=\beta_1=\beta_2=0$.

$$\mathop{\mathbb{E}}_{u,v,w}\widehat{A}_v^4(0)\widehat{A}_w^4(0) = \mathop{\mathbb{E}}_{u}\left(\mathop{\mathbb{E}}_{v}\widehat{A}_v^4(0)\right)^2 \ge \left(\mathop{\mathbb{E}}_{uv}\widehat{A}_v(0)\right)^8 \ge s^8. \tag{9.2.6}$$

Upper Bound on $|\Theta_1|$: We can upper bound $|\Theta_1|$ by

$$\mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\mathsf{rank}(\alpha_1+\alpha_2),\mathsf{rank}(\beta_1+\beta_2)\le k,\\ \pi(\alpha_1+\alpha_2)=\sigma(\beta_1+\beta_2),\\ \nu(\beta_1+\beta_2)=1}} \widehat{A}_v^2(\alpha_1)\widehat{A}_v^2(\alpha_2)\widehat{A}_w^2(\beta_1)\widehat{A}_w^2(\beta_2). \tag{9.2.7}$$

Consider the following strategy for labeling vertices $u \in U$ and $v \in V$. For $u \in U$, pick a random neighbor v, choose (α_1, α_2) with probability $\widehat{A}_v^2(\alpha_1)\widehat{A}_v^2(\alpha_2)$ and set its label to $\pi(\alpha_1 + \alpha_2)$. For $w \in V$, choose (β_1, β_2) with probability $\widehat{A}_w^2(\beta_1)\widehat{A}_w^2(\beta_2)$ and set its label to $\beta_1 + \beta_2$. Since A_w is folded, both β_1 and β_2 are symmetric and satisfies C_v. Since these constraints are homogeneous, $\beta_1 + \beta_2$ is also symmetric and satisfies C_v. Also π maps symmetric matrices to symmetric matrices. Note that (9.2.7) gives the probability that a random edge (u, w) of the label cover is satisfied by this labeling. Hence (9.2.7) and $|\Theta_1|$ are upper bounded by δ.

Upper Bound on $|\Theta_2|$: Note that if the $\mathsf{rank}(\alpha) > k$, for any fixed b, $\Pr_{\overline{x}}[\alpha x = b] \le 1/2^{k+1}$. All terms in Θ_2 has $\max\{\mathsf{rank}(\alpha_1), \mathsf{rank}(\alpha_2), \mathsf{rank}(\beta_1), \mathsf{rank}(\beta_2)\} > k/2$. From (9.2.2) we have that, for any fixed choice of u, v, w each term in Θ_2 has absolute value at most $1/2^{k/2+1}$. Since A, B are $\{0, 1\}$ valued functions, sum of their squared coefficients is upper bounded by 1 (i.e. Parseval's inequality). Thus $|\Theta_2| \le 1/2^{k/2+1}$.

□

Proof of Theorem 9.2. We already saw in Section 9.2.1 that an YES instance of label cover is mapped to a 2-colorable hypergraph. Since $k = (\log N)^{1/8-2\varepsilon}$ and $\delta = 2^{-(\log N)^{1/4-2\varepsilon}}$, $s \le 2^{-(\log N)^{1/8-3\varepsilon}}$. Also the number of vertices in $\mathcal{G}$,

$$n \le N2^{m^2} \le N \cdot 2^{(\log N)^{10/4+2\varepsilon}}.$$

From Lemma 9.10 and above, a NO instance of label cover is mapped to a hypergraph $\mathcal{G}$ that has no independent set of relative size $2^{-(\log n)^{1/20-4\varepsilon}}$. □

9.3 4-Colorable 4-Uniform Hypergraphs

In this section, we modify the reduction in the previous section, so that the uniformity of the hypergraph produced is decreased to 4 at the cost of increasing the number of colors required in the YES case to 4. This method was proposed by Guruswami et al. [GHH+14]. The hypergraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ constructed will have vertices

$$\mathcal{V} = V \times \left(\mathbb{F}_2^{m\times m} \times \mathbb{F}_2^{m\times m}/\mathcal{H}_v \times \mathcal{H}_v\right).$$

Any 4-coloring of $\mathcal{G}$ can be expressed as a collection of functions

$$A'_v : \left(\mathbb{F}_2^{m\times m} \times \mathbb{F}_2^{m\times m}/\mathcal{H}_v \times \mathcal{H}_v\right) \to \{0,1\}^2, \text{ for } v \in V.$$

We can uniquely extend such functions to get $A_v : \mathbb{F}_2^{m\times m} \times \mathbb{F}_2^{m\times m} \to \{0,1\}^2$ which is constant over cosets of $\mathcal{H}_v \times \mathcal{H}_v$. This ensures that A satisfies the following: if $\alpha = (\alpha_1, \alpha_2) \in \mathbb{F}_2^{m\times m} \times \mathbb{F}_2^{m\times m}$ is such that $\widehat{A}(\alpha)$ is non-zero, then α_1, α_2 are both symmetric and satisfies C_v. The set of edges $\mathcal{E}$ will be defined by the test mentioned below.

4-Colorable 4-Uniform Test:

1. Sample v, w and $\{X_i, Y_i\}_{i=1}^4$ from the distribution $\mathcal{T}_{2,8}$ as described by the test in the previous section.

2. Accept if and only if the following 4 values are not all equal :

$$A_v(X_1, X_2) \quad A_v(X_3, X_4) \quad A_w(Y_1, Y_2) \quad A_w(Y_3, Y_4)$$

9.3.1 YES Case

Given a perfectly satisfying labeling $\overline{y}_v \otimes \overline{y}_v$ for $v \in V$ and $\overline{x}_u \otimes \overline{x}_u$ for $u \in U$, we define the following 4-coloring for $\mathcal{G}$: for each $v \in V$,

$$A_v(X_1, X_2) := (\langle X_1, \overline{y}_v \otimes \overline{y}_v \rangle, \langle X_2, \overline{y}_v \otimes \overline{y}_v \rangle) .$$

Note that such a function is constant over cosets of $\mathcal{H}_v$. Using the arguments from Section 9.2.1, it is easy to see that A is a valid 4-coloring of $\mathcal{G}$.

9.3.2 NO Case

The analysis of the NO case is similar to Section 9.2.2.

10

Covering CSPs

In this chapter, we describe our hardness results for Covering CSPs (joint work with Bhangale & Harsha [BHV15]).

The notion of *covering complexity* was introduced by Guruswami, Håstad & Sudan [GHS02] and more formally by Dinur & Kol [DK13] to obtain a better understanding of the complexity of hypergraph coloring problems. Let P be a predicate and Φ an instance of a constraint satisfaction problem (CSP) over n variables, where each constraint in Φ is a constraint of type P over the n variables and their negations. We will refer to such CSPs as P-CSPs. The *covering number* of Φ, denoted by $\nu(\Phi)$, is the smallest number of assignments to the variables such that each constraint of Φ is satisfied by at least one of the assignments, in which case we say that the set of assignments *covers* the instance Φ. If c assignments cover the instance Φ, we say that Φ is c-coverable or equivalently that the set of assignments form a c-covering for Φ. The covering number is a generalization of the notion of chromatic number (to be more precise, the logarithm of the the chromatic number) to all predicates in the following sense. Suppose P is the not-all-equal predicate NAE and the instance Φ has no negations in any of its constraints, then the covering number $\nu(\Phi)$ is exactly $\lceil \log \chi(G_\Phi) \rceil$ where G_Φ is the underlying constraint graph of the instance Φ.

Cover-P refers to the problem of finding the covering number of a given P-CSP

instance. Finding the exact covering number for most interesting predicates P is NP-hard. We therefore study the problem of approximating the covering number. In particular, we would like to study the complexity of the following problem, denoted by Covering-P-CSP(c, s), for some $1 \leq c < s \in \mathbb{N}$: "given a c-coverable P-CSP instance Φ, find an s-covering for Φ". Similar problems have been studied for the Max-CSP setting: "for $0 < s < c \leq 1$, "given a c-satisfiable P-CSP instance Φ, find an s-satisfying assignment for Φ". Max-CSPs and Cover-CSPs, as observed by Dinur & Kol [DK13], are very different problems. For instance, if P is an odd predicate, i.e, if for every assignment x, either x or its negation $x + \overline{1}$ satisfies P, then any P-CSP instance Φ has a trivial two covering, any assignment and its negation. Thus, 3-LIN and 3-CNF*, being odd predicates, are easy to cover though they are hard predicates in the Max-CSP setting. The main result of Dinur & Kol is that the 4-LIN predicate, in contrast to the above, is hard to cover: for every constant $t \geq 2$, Covering-4-LIN-CSP$(2, t)$ is NP-hard. In fact, their arguments show that Covering-4-LIN-CSP$(2, \Omega(\log \log \log n))$ is quasi-NP-hard.

Having observed that odd predicate based CSPs are easy to cover, Dinur and Kol proceeded to ask the question "are all non-odd-predicate CSPs hard to cover?". In a partial answer to this question, they showed that assuming a covering variant of the unique games conjecture Covering-UGC(c) (Conjecture 10.8), if a predicate P is not odd and there is a balanced pairwise independent distribution on its support, then for all constants k, Covering-P-CSP$(2c, k)$ is NP-hard (here, c is a fixed constant that depends on the covering variant of the unique games conjecture Covering-UGC(c)).

Our first result states that assuming the same covering variant of unique games conjecture Covering-UGC(c) of Dinur & Kol [DK13], one can in fact show the covering hardness of *all* non-odd predicates P over *any* constant-sized alphabet $[q]$. The notion

*3-LIN : $\{0,1\}^3 \to \{0,1\}$ refers to the 3-bit predicate defined by 3-LIN$(x_1, x_2, x_3) := x_1 \oplus x_2 \oplus x_3$ while 3-CNF : $\{0,1\}^3 \to \{0,1\}$ refers to the 3-bit predicate defined by 3-CNF$(x_1, x_2, x_3) := x_1 \vee x_2 \vee x_3$

of odd predicate can be extended to any alphabet in the following natural way: a predicate $P \subseteq [q]^k$ is odd if for all assignments $x \in [q]^k$, there exists $a \in [q]$ such that the assignment $x + \overline{a}$ satisfies P.

Theorem 10.1 (Covering hardness of non-odd predicates). *Assuming* Covering-UGC(c)*, for any constant-sized alphabet* $[q]$*, any constant* $k \in \mathbb{N}$ *and any* non-odd *predicate* $P \subseteq [q]^k$*, for all constants* $t \in \mathbb{N}$*, the* Covering-P-CSP($2cq, t$) *problem is NP-hard.*

Since odd predicates $P \subseteq [q]^k$ are trivially coverable with q assignments, the above theorem, gives a *full characterization of hard-to-cover predicates* over any constant sized alphabet (modulo the covering variant of the unique games conjecture): a predicate is hard to cover iff it is not odd.

We then ask if we can prove similar covering hardness results under more standard complexity assumptions (such as NP≠P or the exponential-time hypothesis (ETH)). Though we are not able to prove that every non-odd predicate is hard under these assumptions, we give sufficient conditions on the predicate P for the corresponding approximate covering problem to be quasi-NP-hard. Recall that $2k$-LIN $\subseteq \{0,1\}^{2k}$ is the predicate corresponding to the set of odd parity strings in $\{0,1\}^{2k}$.

Theorem 10.2 (NP-hardness of Covering). *Let* $k \geq 2$*. Let* $P \subseteq 2k$-LIN *be any* $2k$*-bit predicate such there exists distributions* $\mathcal{P}_0, \mathcal{P}_1$ *supported on* $\{0,1\}^k$ *with the following properties:*

1. *the marginals of* $\mathcal{P}_0$ *and* $\mathcal{P}_1$ *on all* k *coordinates is uniform,*
2. *every* $a \in \text{support}(\mathcal{P}_0)$ *has even parity and every* $b \in \text{support}(\mathcal{P}_1)$ *has odd parity and furthermore, both* $a \cdot b, b \cdot a \in P$ *(where* $\cdot$ *stands for concatenation of strings).*

Then, unless NP $\subseteq DTIME(2^{\text{poly}\log n})$*, for all* $\varepsilon \in (0, 1/2]$, Covering-P-CSP$(2, \Omega(\log\log n))$ *is not solvable in polynomial time.*

Furthermore, the YES and NO instances of Covering-P-CSP$(2, \Omega(\log\log n))$ *satisfy the following properties.*

- *YES Case : There are* 2 *assignments such that each of them covers* $1 - \varepsilon$ *fraction of the constraints and they together cover the instance.*
- *NO Case : Even the* $2k$-LIN-CSP *instance with the same constraint graph as the given instance is not* $\Omega(\log\log n)$*-coverable.*

The furthermore clause in the soundness guarantee is in fact a strengthening for the following reason: if two predicates P, Q satisfy $P \subseteq Q$ and Φ is a c-coverable P-CSP instance, then the Q-CSP instance $\Phi_{P\to Q}$ obtained by taking the constraint graph of Φ and replacing each P constraint with the weaker Q constraint, is also c-coverable.

The following is a simple corollary of the above theorem.

Corollary 10.3. *Let* $k \geq 2$ *be even,* $x, y \in \{0,1\}^k$ *be distinct strings having even and odd parity respectively and* $\overline{x}, \overline{y}$ *denote the complements of* x *and* y *respectively. For any predicate* P *satisfying*

$$2k\text{-LIN} \supseteq P \supseteq \{x \cdot y,\ x \cdot \overline{y},\ \overline{x} \cdot y,\ \overline{x} \cdot \overline{y},\ y \cdot x,\ y \cdot \overline{x},\ \overline{y} \cdot x,\ \overline{y} \cdot \overline{x}\},$$

unless NP $\subseteq DTIME(2^{\text{poly}\log n})$*, the problem* Covering-P-CSP$(2, \Omega(\log\log n))$ *is not solvable in polynomial time.*

This corollary implies the covering hardness of 4-LIN predicate proved by Dinur & Kol [DK13] by setting $x := 00$ and $y := 01$. With respect to the covering hardness of 4-LIN, we note that we can considerably simplify the proof of Dinur & Kol and in fact

obtain a even stronger soundness guarantee (see Theorem below). The stronger soundness guarantee in the theorem below states that there are no large ($\geq 1/\text{poly}\log n$ fractional sized) independent sets in the constraint graph and hence, even the 4-NAE-CSP instance[†] with the same constraint graph as the given instance is not coverable using $\Omega(\log\log n)$ assignments. Both the Dinur-Kol result and the above corollary only guarantee (in the soundness case) that the 4-LIN-CSP instance is not coverable.

Theorem 10.4 (Hardness of Covering 4-LIN). *Assuming that NP* $\not\subseteq DTIME(2^{\text{poly}\log n})$*, for all* $\varepsilon \in (0,1)$*, there does not exist a polynomial time algorithm that can distinguish between* 4-LIN-CSP *instances of the following two types:*

- *YES Case : There are* 2 *assignments such that each of them covers* $1-\varepsilon$ *fraction of the constraints, and they together cover the entire instance.*
- *NO Case : The largest independent set in the constraint graph of the instance is of fractional size at most* $1/\text{poly}\log n$*.*

10.1 Preliminaries

We will denote the set $\{0,1,\cdots q-1\}$ by $[q]$. For $a \in [q], \bar{a} \in [q]^k$ is the element with a in all the k coordinates (where k and q will be implicit from the context).

Definition 10.5 (P-CSP). *For a predicate* $P \subseteq [q]^k$*, an instance of* P*-CSP is given by a (hyper)graph* $G=(V,E)$*, referred to as the* constraint graph*, and a literals function* $L : E \to [q]^k$*, where* V *is a set of variables and* $E \subseteq V^k$ *is a set of constraints. An assignment* $f : V \to [q]$ *is said to* cover *a constraint* $e = (v_1,\cdots,v_k) \in E$*, if* $(f(v_1),\cdots,f(v_k)) + L(e) \in P$*, where addition is coordinate-wise modulo* q*. A set of assignments* $F = \{f_1,\cdots,f_c\}$ *is said to* cover (G,L)*, if for every* $e \in E$*, there is some*

[†]The k-NAE predicate over k bits is given by k-NAE $= \{0,1\}^k \setminus \{\bar{0},\bar{1}\}$.

$f_i \in F$ *that covers* e *and* F *is said to be a* c-covering *for* G. G *is said to be* c-coverable *if there is a* c*-covering for* G. *If* L *is not specified then it is the constant function which maps* E *to* $\bar{0}$.

Definition 10.6 (Covering-P-CSP(c, s)). *For* $P \subseteq [q]^k$ *and* $c, s \in \mathbb{N}$, *the* Covering-P-CSP(c, s) *problem is, given a* c*-coverable instance* $(G = (V, E), L)$ *of* P-CSP, *find an* s*-covering.*

Definition 10.7 (Odd). *A predicate* $P \subseteq [q]^k$ *is* odd *if* $\forall x \in [q]^k, \exists a \in [q], x + \bar{a} \in P$, *where addition is coordinate-wise modulo* q.

For odd predicates the covering problem is *trivially solvable*, since any CSP instance on such a predicate is q-coverable by the q translates of any assignment, i.e., $\{x + \bar{a} \mid a \in [q]\}$ is a q-covering for any assignment $x \in [q]^k$.

Our characterization of hardness of covering CSPs is based on the following conjecture due to Dinur & Kol [DK13].

Conjecture 10.8 (Covering-UGC(c)). *There exists* $c \in \mathbb{N}$ *such that for every sufficiently small* $\delta > 0$ *there exists* $L \in \mathbb{N}$ *such that the following holds. Given a an instance* $G = (U, V, E, [L], [L], \{\pi_e\}_{e \in E})$ *of* Unique-Game *it is NP-hard to distinguish between the following two cases:*

- *YES case: There exist* c *assignments such that for every vertex* $u \in U$, *at least one of the assignments satisfies all the edges touching* u.
- *NO case: Every assignment satisfies at most* δ *fraction of the edge constraints.*

10.2 A Characterization of Hard-to-cover CSPs

In this section, we prove the following theorem, which in turn implies Theorem 10.1 (see below for proof).

Theorem 10.9. *Let* $[q]$ *be any constant sized alphabet and* $k \geq 2$ *and for* $b \in [q]$, $\bar{b} := (b, \cdots, b) \in [q]^k$ *. Recall that* NAE $:= [q]^k \setminus \{\bar{b} \mid b \in [q]\}$. *Let* $P \subseteq [q]^k$ *be a predicate such that there exists* $a \in$ NAE *and* NAE $\supset P \supseteq \{a + \bar{b} \mid b \in [q]\}$. *Assuming* Covering-UGC(c), *for every sufficiently small constant* $\delta > 0$ *it is NP-hard to distinguish between* P-CSP *instances* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ *of the following two cases:*

- *YES Case :* $\mathcal{G}$ *is* $2c$*-coverable.*
- *NO Case :* $\mathcal{G}$ *does not have an independent set of fractional size* δ.

Proof of Theorem 10.1. Let Q be an arbitrary non odd predicate, i.e, $Q \subseteq [q]^k \setminus \{h + \bar{b} \mid b \in [q]\}$ for some $h \in [q]^k$. Consider the predicate $Q' \subseteq [q]^k$ defined as

$$Q' := \{x \in [q]^k : \exists y \in Q, \forall i \in [k], x_i = y_i - h\}.$$

Observe that $Q' \subseteq$ NAE. Given any Q'-CSP instance Φ with literals function $L(e) = \overline{0}$, consider the Q-CSP instance $\Phi_{Q' \to Q}$ with literals function M given by $M(e) := \overline{h}, \forall e$. It has the same constraint graph as Φ. Clearly, Φ is c-coverable iff $\Phi_{Q' \to Q}$ is c-coverable. Thus, it suffices to prove the result for any predicate $Q' \subseteq$ NAE with literals function $L(e) = \overline{0}$[‡]. We will consider two cases, both of which will follow from Theorem 10.9.

Suppose the predicate Q' satisfies $Q' \supseteq \{a + \bar{b} \mid b \in [q]\}$ for some $a \in [q]^k$. Then this predicate Q' satisfies the hypothesis of Theorem 10.9 and the theorem follows if we show that the soundness guarantee of Theorem 10.9 implies that in Theorem 10.1. Any instance in the NO case of Theorem 10.9, is not $t := \log_q(1/\delta)$-coverable even on the NAE-CSP instance with the same constraint graph. This is because any t-covering for the NAE-CSP instance gives a coloring of the constraint graph using q^t

[‡]This observation [DK13] that the cover-Q problem for any non-odd predicate Q is equivalent to the cover-Q' problem where $Q' \subseteq$ NAE shows the centrality of the NAE predicate in understanding the covering complexity of any non-odd predicate.

colors, by choosing the color of every variable to be a string of length t and having the corresponding assignments in each position in $[t]$. Hence the Q'-CSP instance is also not t-coverable.

Suppose $Q' \not\supseteq \{a + \bar{b} \mid b \in [q]\}$ for all $a \in [q]^k$. Then consider the predicate $P = \{a + \bar{b} \mid a \in Q', b \in [q]\} \subseteq \text{NAE}$. Notice that P satisfies the conditions of Theorem 10.9 and if the P-CSP instance is t-coverable then the Q'-CSP instance is qt-coverable. Hence an YES instance of Theorem 10.9 maps to a $2cq$-coverable Q-CSP instance and NO instance maps to an instance with covering number at least $\log_q(1/\delta)$.

□

We now prove Theorem 10.9 by giving a reduction from an instance $G = (U, V, E, [L], [L], \{\pi_e\}_{e \in E})$ of Unique-Game as in Definition 6.5, to an instance $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ of a P-CSP for any predicate P that satisfies the conditions mentioned. As stated in the introduction, we adapt the long-code test of Bansal & Khot [BK10] for proving the hardness of finding independent sets in almost k-partite k-uniform hypergraphs to our setting. The set of variables $\mathcal{V}$ is $V \times [q]^{2L}$. Any assignment to $\mathcal{V}$ is given by a set of functions $f_v : [q]^{2L} \to [q]$, for each $v \in V$. The set of constraints $\mathcal{E}$ is given by the following test which checks whether f_v's are long codes of a good labeling to V. There is a constraint corresponding to all the variables that are queried together by the test.

Long Code Test $\mathcal{T}_1$:

1. Choose $u \in U$ uniformly and k neighbors $w_1, \ldots, w_k \in V$ of u uniformly and independently at random.

2. Choose a random matrix X of dimension $k \times 2L$ as follows. Let X^i denote the i^{th} column of X. Independently for each $i \in [L]$, choose (X^i, X^{i+L}) uniformly

at random from the set

$$S := \left\{(y, y') \in [q]^k \times [q]^k \mid y \in \{a + \bar{b} \mid b \in [q]\} \vee y' \in \{a + \bar{b} \mid b \in [q]\}\right\}. \tag{10.2.1}$$

3. Let $x_1, \cdots, x_k$ be the rows of matrix X. Accept iff

$$(f_{w_1}(x_1 \circ \pi_{uw_1}), f_{w_2}(x_2 \circ \pi_{uw_2}), \cdots, f_{w_k}(x_k \circ \pi_{uw_k})) \in P,$$

where $x \circ \pi$ is the string defined as $(x \circ \pi)(i) := x_{\pi(i)}$ for $i \in [L]$ and $(x \circ \pi)(i) := x_{\pi(i-L)+L}$ otherwise.

Lemma 10.10 (Completeness). *If the* Unique-Game *instance G is c-coverable then the P-*CSP *instance $\mathcal{G}$ is $2c$-coverable.*

Proof. Let $\ell_1, \ldots, \ell_c : U \cup V \to [L]$ be a c-covering for G as described in Definition 6.5. We will show that the $2c$ assignments given by $f_v^i(x) := x_{\ell_i(v)}, g_v^i(x) := x_{\ell_i(v)+L}, i = 1, \ldots, c$ form a $2c$-covering of $\mathcal{G}$. Consider any $u \in U$ and let ℓ_i be the labeling that covers all the edges incident on u. For any $(u, w_j)_{j \in \{1,\cdots,k\}} \in E$ and X chosen by the long code test $\mathcal{T}_1$, the vector $(f_{w_1}^i(x_1 \circ \pi_{uw_1}), \cdots, f_{w_k}^i(x_k \circ \pi_{uw_k}))$ gives the $\ell_i(u)^{\text{th}}$ column of X. Similarly the above expression corresponding to g^i gives the $(\ell_i(u) + L)^{\text{th}}$ column of the matrix X. Since, for all $i \in [L]$, either i^{th} column or $(i + L)^{\text{th}}$ column of X contains element from $\{a + \bar{b} \mid b \in [q]\} \subseteq P$, either $(f_{w_1}^i(x_1 \circ \pi_{uw_1}), \cdots, f_{w_k}^i(x_k \circ \pi_{uw_k})) \in P$ or $(g_{w_1}^i(x_1 \circ \pi_{uw_1}), \cdots, g_{w_k}^i(x_k \circ \pi_{uw_k})) \in P$. Hence the set of $2c$ assignments $\{f_v^i, g_v^i\}_{i \in \{1,\cdots,c\}}$ covers all constraints in $\mathcal{G}$. □

To prove soundness, we show that the set S, as defined in Equation (10.2.1), is connected, so that Theorem 2.2 is applicable.

Claim 10.11. *Consider $S \subseteq [q]^k \times [q]^k$ as a subset of $([q]^2)^k$ as follows: the element $(y, y') \in S$ is mapped to the element $((y_1, y'_1), \cdots, (y_k, y'_k)) \in ([q]^2)^k$. Let $\Omega = [q]^2$. The set $S \subset \Omega^k$ is connected (as in Definition 2.1).*

Proof. Consider any $x := (x^1, x^2), y := (y^1, y^2) \in S \subset [q]^k \times [q]^k$. Suppose both $x^1, y^1 \in \{a + \bar{b} \mid b \in [q]\}$, then it is easy to come up with a sequence of strings belonging to S, starting with x and ending with y such that consecutive strings differ in at most 1 coordinate,. Now suppose $x^1, y^2 \in \{a + \bar{b} \mid b \in [q]\}$. First we come up with a sequence from (x^1, x^2) to (x^1, y^2), and then another sequence for (x^1, y^2) to (y^1, y^2). This can be done since in the definition of S, we are only constraining one of x^1 or x^2 to be in $\{a + \bar{b} : b \in [q]\}$. □

Lemma 10.12 (Soundness). *For every constant $\delta > 0$, there exists a constant s such that, if G is at most s-satisfiable then $\mathcal{G}$ does not have an independent set of size δ.*

Proof. Let $I \subseteq \mathcal{V}$ be an independent set of fractional size δ in the constraint graph. For every variable $v \in V$, let $f_v : [q]^{2L} \to \{0, 1\}$ be the indicator function of the independent set restricted to the vertices that correspond to v. For a vertex $u \in U$, let $N(u) \subseteq V$ be the set of neighbors of u and define $f_u(x) := \mathbb{E}_{w \in N(u)}[f_w(x \circ \pi_{uw})]$. Since I is an independent set, we have

$$0 = \mathop{\mathbb{E}}_{u, w_i, \ldots, w_k} \mathop{\mathbb{E}}_{X \sim \mathcal{T}_1} \left[\prod_{i=1}^{k} f_{w_i}(x_i \circ \pi_{uw_i}) \right] = \mathop{\mathbb{E}}_{u} \mathop{\mathbb{E}}_{X \sim \mathcal{T}_1} \left[\prod_{i=1}^{k} f_u(x_i) \right]. \tag{10.2.2}$$

Since the bipartite graph (U, V, E) is left regular and $|I| \geq \delta |V|$, we have $\mathbb{E}_{u,x}[f_u(x)] \geq \delta$. By an averaging argument, for at least $\frac{\delta}{2}$ fraction of the vertices $u \in U$, $\mathbb{E}_x[f_u(x)] \geq \frac{\delta}{2}$. Call a vertex $u \in U$ *good* if it satisfies this property. A string $x \in [q]^{2L}$ can be thought as an element from $([q]^2)^L$ by grouping the pair of coordinates x_i, x_{i+L}. Let $\overline{x} \in ([q]^2)^L$ denotes this grouping of x, i.e., j^{th} coordinate of $\overline{x}$ is $(x_j, x_{j+L}) \in [q]^2$. With this grouping, the function f_u can be viewed as $f_u : ([q]^2)^L \to \{0, 1\}$. From

Equation (10.2.2), we have that for any $u \in U$,

$$\mathop{\mathbb{E}}_{X \sim \mathcal{T}_1} \left[\prod_{i=1}^{k} f_u(\overline{x}_i) \right] = 0.$$

By Claim 10.11, for all $j \in [L]$ the tuple $((\overline{x}_1)_j, \ldots, (\overline{x}_k)_j)$ (corresponding to columns (X^j, X^{j+L}) of X) is sampled from a distribution whose support is a connected set. Hence for a good vertex $u \in U$, we can apply Theorem 2.2 with $\varepsilon = \underline{\Gamma}(\delta/2)/2$ to get that there exists $j \in [L], d \in \mathbb{N}, \tau > 0$ such that $\mathrm{Inf}_j^{\leq d}(f_u) > \tau$. We will use this fact to give a randomized labeling for G. Labels for vertices $w \in V, u \in U$ will be chosen uniformly and independently from the sets

$$\mathrm{Label}(w) := \left\{ i \in [L] \mid \mathrm{Inf}_i^{\leq d}(f_w) \geq \frac{\tau}{2} \right\}, \mathrm{Label}(u) := \left\{ i \in [L] \mid \mathrm{Inf}_i^{\leq d}(f_u) \geq \tau \right\}.$$

By the above argument (using Theorem 2.2), we have that for a good vertex u, $\mathrm{Label}(u) \neq \emptyset$. Furthermore, since the sum of degree d influences is at most d, the above sets have size at most $2d/\tau$. Now, for any $j \in \mathrm{Label}(u)$, we have

$$\begin{aligned} \tau < \mathrm{Inf}_j^{\leq d}[f_u] &= \sum_{S : j \in S, |S| \leq d} \|f_{u,S}\|^2 = \sum_{S : j \in S, |S| \leq d} \left\| \mathop{\mathbb{E}}_{w \in N(u)} \left[f_{w, \pi_{uw}^{-1}(S)} \right] \right\|^2 \quad \text{(By Definition.)} \\ &\leq \sum_{S : j \in S, |S| \leq d} \mathop{\mathbb{E}}_{w \in N(u)} \left\| f_{w, \pi_{uw}^{-1}(S)} \right\|^2 = \mathop{\mathbb{E}}_{w \in N(u)} \mathrm{Inf}_{\pi_{uw}^{-1}(j)}^{\leq d}[f_w]. \quad \text{(By Convexity of square.)} \end{aligned}$$

Hence, by another averaging argument, there exists at least $\frac{\tau}{2}$ fraction of neighbors w of u such that $\mathrm{Inf}_{\pi_{uw}^{-1}(j)}^{\leq d}(f_w) \geq \frac{\tau}{2}$ and hence $\pi_{uw}^{-1}(j) \in \mathrm{Label}(w)$. Therefore, for a good vertex $u \in U$, at least $\frac{\tau}{2}\frac{\tau}{2d}$ fraction of edges incident on u are satisfied in expectation. Also, at least $\frac{\delta}{2}$ fraction of vertices in U are good, it follows that the expected fraction of edges that are satisfied by this random labeling is at least $\frac{\delta}{2}\frac{\tau}{2}\frac{\tau}{2d}$. Choosing $s < \frac{\delta}{2}\frac{\tau}{2}\frac{\tau}{2d}$ completes the proof. □

10.3 Some NP-hardness Results

In this section, we prove Theorem 10.2. We give a reduction from an instance of a LC, $G = (U, V, E, [L], [R], \{\pi_e\}_{e \in E})$ as in Definition 6.5, to a P-CSP instance $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ for any predicate P that satisfies the conditions mentioned in Theorem 10.2. The reduction and proof is similar to that of Dinur & Kol [DK13]. The main difference is that they used a test and invariance principle very specific to the 4-LIN predicate, while we show that a similar analysis can be performed under milder conditions on the test distribution.

We assume that $R = dL$ and $\forall i \in [L], e \in E, |\pi_e^{-1}(i)| = d$. This is done just for simplifying the notation and the proof does not depend upon it. The set of variables $\mathcal{V}$ is $V \times \{0,1\}^{2R}$. Any assignment to $\mathcal{V}$ is given by a set of functions $f_v : \{0,1\}^{2R} \to \{0,1\}$, for each $v \in V$. The set of constraints $\mathcal{E}$ is given by the following test which checks whether f_v's are long codes of a good labeling to V.

Long Code Test $\mathcal{T}_2$:

1. Choose $u \in U$ uniformly and $v, w \in V$ neighbors of u uniformly and independently at random. For $i \in [L]$, let $B_{uv}(i) := \pi_{uv}^{-1}(i), B'_{uv}(i) := R + \pi_{uv}^{-1}(i)$ and similarly for w.

2. Choose matrices X, Y of dimension $k \times 2dL$ as follows. For $S \subseteq [2dL]$, we denote by $X|_S$ the submatrix of X restricted to the columns S. Independently for each $i \in [L]$, choose $c_1 \in \{0,1\}$ uniformly and

   (a) if $c_1 = 0$, choose $\left(X|_{B_{uv}(i) \cup B'_{uv}(i)}, Y|_{B_{uw}(i) \cup B'_{uw}(i)}\right)$ from $\mathcal{P}_0^{\otimes 2d} \otimes \mathcal{P}_1^{\otimes 2d}$,

   (b) if $c_1 = 1$, choose $\left(X|_{B_{uv}(i) \cup B'_{uv}(i)}, Y|_{B_{uw}(i) \cup B'_{uw}(i)}\right)$ from $\mathcal{P}_1^{\otimes 2d} \otimes \mathcal{P}_0^{\otimes 2d}$.

3. Perturb X, Y as follows. Independently for each $i \in [L]$, choose $c_2 \in \{*, 0, 1\}$ as follows: $\Pr[c_2 = *] = 1 - 2\varepsilon$, and $\Pr[c_2 = 1] = \Pr[c_2 = 0] = \varepsilon$. Perturb the ith matrix block $\left(X|_{B_{uv}(i) \cup B'_{uv}(i)}, Y|_{B_{uw}(i) \cup B'_{uw}(i)}\right)$ as follows:

   (a) if $c_2 = *$, leave the matrix block $\left(X|_{B_{uv}(i) \cup B'_{uv}(i)}, Y|_{B_{uw}(i) \cup B'_{uw}(i)}\right)$ unperturbed,

   (b) if $c_2 = 0$, choose $\left(X|_{B'_{uv}(i)}, Y|_{B'_{uw}(i)}\right)$ uniformly from $\{0,1\}^{k \times d} \times \{0,1\}^{k \times d}$,

   (c) if $c_2 = 1$, choose $\left(X|_{B_{uv}(i)}, Y|_{B_{uw}(i)}\right)$ uniformly from $\{0,1\}^{k \times d} \times \{0,1\}^{k \times d}$.

4. Let $x_1, \cdots, x_k$ and $y_1, \cdots, y_k$ be the rows of the matrices X and Y respectively. Accept if

$$(f_v(x_1), \cdots, f_v(x_k), f_w(y_1), \cdots, f_w(y_k)) \in P.$$

Lemma 10.13 (Completeness). *If G is an YES instance of* LC*, then there exists f, g such that each of them covers $1 - \varepsilon$ fraction of $\mathcal{E}$ and they together cover all of $\mathcal{E}$.*

Proof. Let $\ell : U \cup V \to [L] \cup [R]$ be a labeling to G that satisfies all the constraints. Consider the assignments $f_v(x) := x_{\ell(v)}$ and $g_v(x) := x_{R+\ell(v)}$ for each $v \in V$. First consider the assignment f. For any $(u, v), (u, w) \in E$ and $x_1, \cdots, x_k, y_1, \cdots, y_k$ chosen by the long code test $\mathcal{T}_2$, $(f_v(x_1), \cdots, f_v(x_k)), (f_w(y_1), \cdots, f_w(y_k))$ gives the $\ell(v)^{\text{th}}$ and $\ell(w)^{\text{th}}$ column of the matrices X and Y respectively. Since $\pi_{uv}(\ell(v)) = \pi_{uw}(\ell(w))$, they are jointly distributed either according to $\mathcal{P}_0 \otimes \mathcal{P}_1$ or $\mathcal{P}_1 \otimes \mathcal{P}_0$ after Step 2. The probability that these rows are perturbed in Step 3c is at most ε. Hence with probability $1 - \varepsilon$ over the test distribution, f is accepted. A similar argument shows that the test accepts g with probability $1 - \varepsilon$. Note that in Step 3, the columns given by f, g, are never re-sampled uniformly together. Hence they together cover $\mathcal{G}$.

□

Now we will show that if G is a NO instance of LC then no t assignments can cover the $2k$-LIN-CSP with constraint hypergraph $\mathcal{G}$. For the rest of the analysis, we will use $+1, -1$ instead of the symbols $0, 1$. Suppose for contradiction, there exist t assignments $f_1, \cdots, f_t : \{\pm 1\}^{2R} \to \{\pm 1\}$ that form a t-cover to $\mathcal{G}$. The probability that all the t assignments are rejected in Step 4 is

$$\mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{\mathcal{T}_2} \left[\prod_{i=1}^{t} \frac{1}{2} \left(\prod_{j=1}^{k} f_{i,v}(x_j) f_{i,w}(y_j) + 1 \right) \right] = \frac{1}{2^t} + \frac{1}{2^t} \sum_{\emptyset \subset S \subseteq \{1,\cdots,t\}} \mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{\mathcal{T}_2} \left[\prod_{j=1}^{k} f_{S,v}(x_j) f_{S,w}(y_j) \right]. \tag{10.3.1}$$

where $f_{S,v}(x) := \prod_{i \in S} f_{i,v}(x)$. Since the t assignments form a t-cover, the LHS in Equation (10.3.1) is 0 and hence, there exists an $S \neq \emptyset$ such that

$$\mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{\mathcal{T}_2} \left[\prod_{j=1}^{k} f_{S,v}(x_j) f_{S,w}(y_j) \right] \leq -1/(2^t - 1). \tag{10.3.2}$$

The following lemma shows that this is not possible if t is not too large, thus proving that there does not a exist t-cover.

Lemma 10.14 (Soundness). *Let $c_0 \in (0,1)$ be the constant from Theorem 6.6 and $S \subseteq \{1, \cdots, t\}, |S| > 0$. If G is at most s-satisfiable then*

$$\mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{X,Y \in \mathcal{T}_2} \left[\prod_{i=1}^{k} f_{S,v}(x_i) f_{S,w}(y_i) \right] \geq -O(k s^{c_0/8}) - 2^{O(k)} \frac{s^{(1-3c_0)/8}}{\varepsilon^{3/2c_0}}.$$

We will continue the proof of Lemma 10.14, after the proof of Theorem 10.2.

Proof of Theorem 10.2. Using Theorem 6.6, the size of the CSP instance $\mathcal{G}$ produced by the reduction is $N = n^r 2^{2^{O(r)}}$ and the parameter $s \leq 2^{-d_0 r}$. Setting $r = \Theta(\log \log n)$, gives that $N = 2^{\text{poly}(\log n)}$ for a constant k. Lemma 10.14 and Equation 10.3.2 imply that

$$O(k s^{c_0/8}) + 2^{O(k)} \frac{s^{(1-3c_0)/8}}{\varepsilon^{3/2c_0}} \geq \frac{1}{2^t - 1}.$$

Since k is a constant, this gives that $t = \Omega(\log \log n)$. □

Proof of Lemma 10.14. Notice that for a fixed u, the distribution of X and Y have identical marginals. Hence the value of the above expectation, if calculated according to a distribution which is the direct product of the marginals, is positive. We will first show that the expectation can change by at most $O(ks^{c_0/8})$ in moving to an *attenuated* version of the functions (see Claim 10.15). Then we will show that the error incurred by changing the distribution to the product distribution of the marginals has absolute value at most $2^{O(k)}\frac{s^{(1-3c_0)/8}}{\varepsilon^{3/2c_0}}$ (see Claim 10.16). This is done by showing that there is a labeling to G that satisfies an s fraction of the constraints if the error is more than $2^{O(k)}\frac{s^{(1-3c_0)/8}}{\varepsilon^{3/2c_0}}$.

For the rest of the analysis, we write f_v and f_w instead of $f_{S,v}$ and $f_{S,w}$ respectively. Let $f_v = \sum_{\alpha\subseteq[2R]}\widehat{f_v}(\alpha)\chi_\alpha$ be the Fourier decomposition of the function and for $\gamma \in (0,1)$, let $T_{1-\gamma}f_v := \sum_{\alpha\subseteq[2R]}(1-\gamma)^{|\alpha|}\widehat{f_v}(\alpha)\chi_\alpha$. The following claim is similar to a lemma of Dinur & Kol [DK13, Lemma 4.11]. The only difference in the proof is that, we use the smooth projections property of Theorem 6.6 (which was shown by Håstad [Hås01, Lemma 6.9]).

Claim 10.15. *Let $\gamma := s^{(c_0+1)/4}\varepsilon^{1/c_0}$ where c_0 is the constant from Theorem 6.6.*

$$\left|\mathop{\mathbb{E}}_{u,v,w}\mathop{\mathbb{E}}_{T_2}\left[\prod_{i=1}^{k} f_v(x_i)f_w(y_i)\right] - \mathop{\mathbb{E}}_{u,v,w}\mathop{\mathbb{E}}_{T_2}\left[\prod_{i=1}^{k} T_{1-\gamma}f_v(x_i)T_{1-\gamma}f_w(y_i)\right]\right| \leq O(ks^{c_0/8}).$$

Fix u, v, w chosen by the test. Recall that we thought of f_v as having domain $\prod_{i\in[L]}\Omega_i$ where $\Omega_i = \{0,1\}^{2d}$ corresponds to the set of coordinates in $B_{uv}(i)\cup B'_{uv}(i)$. Since the grouping of coordinates depends on u, we define $\overline{\mathrm{Inf}}_i^u[f_v] := \mathrm{Inf}_i[f_v]$ where $i \in [L]$ for explicitness. We will think of f_v as $f_v : \prod_{i\in L}\Omega_i \to \mathbb{R}$ where $\Omega_i = \{0,1\}^d$ consists of the d coordinates j such that $\pi_{uv}(j) = i$. An Efron-Stein decomposition of $f : \prod_{i\in L}\Omega_i \to \mathbb{R}$ over the uniform distribution over $\{0,1\}^{dL}$, can be obtained from

the Fourier decomposition as

$$f_\beta(x) = \sum_{\alpha \subseteq [dL] : \pi(\alpha) = \beta} \widehat{f}(\alpha)\chi_\alpha. \tag{10.3.3}$$

From Equation (10.3.3),

$$\overline{\mathrm{Inf}}_i^u[f_v] = \sum_{\alpha \subseteq [2dL] : i \in \widetilde{\pi}_{uv}(\alpha)} \widehat{f}_v(\alpha)^2,$$

where $\widetilde{\pi}_{uv}(\alpha) := \{i \in [L] : \exists j \in [R], (j \in \alpha \vee j + R \in \alpha) \wedge \pi_{uv}(j) = i\}$.

Claim 10.16. *Let* $\tau_{u,v,w} := \sum_{i \in [L]} \overline{\mathrm{Inf}}_i^u[T_{1-\gamma} f_v] \cdot \overline{\mathrm{Inf}}_i^u[T_{1-\gamma} f_w]$.

$$\mathop{\mathbb{E}}_{u,v,w} \left| \mathop{\mathbb{E}}_{\mathcal{T}_2} \left[\prod_{i=1}^k T_{1-\gamma} f_v(x_i) T_{1-\gamma} f_w(y_i) \right] - \mathop{\mathbb{E}}_{\mathcal{T}_2} \left[\prod_{i=1}^k T_{1-\gamma} f_v(x_i) \right] \mathop{\mathbb{E}}_{\mathcal{T}_2} \left[\prod_{i=1}^k T_{1-\gamma} f_w(y_i) \right] \right| \leq 2^{O(k)} \sqrt{\frac{\mathbb{E}_{u,v,w} \tau_{u,v,w}}{\gamma}}.$$

We defer the proofs of Claim 10.16 and Claim 10.15 to later sections. From Claim 10.16 and Claim 10.15 and using the fact the the marginals of the test distribution $\mathcal{T}_2$ on $(x_1, \ldots, x_k)$ is the same as marginals on $(y_1, \ldots, y_k)$, for $\gamma := s^{(c_0+1)/4} \varepsilon^{1/c_0}$, we get

$$\mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{X,Y \in \mathcal{T}_2} \left[\prod_{i=1}^k f_v(x_i) f_w(y_i) \right] \geq -O(k s^{c_0/8}) - 2^{O(k)} \sqrt{\frac{\mathbb{E}_{u,v,w} \tau_{u,v,w}}{\gamma}} + \mathop{\mathbb{E}}_u \left(\mathop{\mathbb{E}}_v \mathop{\mathbb{E}}_{\mathcal{T}_2} \left[\prod_{i=1}^k T_{1-\gamma} f_v(x_i) \right] \right)^2. \tag{10.3.4}$$

If $\tau_{u,v,w}$ in expectation is large, there is a standard way of decoding the assignments to a labeling to the label cover instance, as shown in Claim 10.17.

Claim 10.17. *If G is an at most s-satisfiable instance of* LC *then*

$$\mathop{\mathbb{E}}_{u,v,w} \tau_{u,v,w} \leq \frac{s}{\gamma^2}.$$

Substituting above into Equation (10.3.4) proves Lemma 10.14. □

We now prove Claim 10.17, Claim 10.16 and Claim 10.15.

Proof of Claim 10.17. Note that $\sum_{\alpha\subseteq[2R]}(1-\gamma)^{|\alpha|}\widehat{f}_v(\alpha)^2 \leq 1$. We will give a randomized labeling to the LC instance. For each $v \in V$, choose a random $\alpha \subseteq [2R]$ with probability $(1-\gamma)^{|\alpha|}\widehat{f}_v(\alpha)^2$ and assign a uniformly random label j in α to v; if the label $j \geq R$, change the label to $j - R$ and with the remaining probability assign an arbitrary label. For $u \in U$, choose a random neighbor $w \in V$ and a random $\beta \subseteq [2R]$ with probability $(1-\gamma)^{|\beta|}\widehat{f}_w(\beta)^2$, choose a random label ℓ in β and assign the label $\widetilde{\pi}_{uw}(\ell)$ to u. With the remaining probability, assign an arbitrary label. The fraction of edges satisfied by this labeling is at least

$$\mathop{\mathbb{E}}_{u,v,w}\sum_{i\in[L]}\sum_{(\alpha,\beta):i\in\widetilde{\pi}_{uv}(\alpha),i\in\widetilde{\pi}_{uw}(\beta)}\frac{(1-\gamma)^{|\alpha|+|\beta|}}{|\alpha|\cdot|\beta|}\widehat{f}_v(\alpha)^2\widehat{f}_w(\beta)^2.$$

Using the fact that $1/r \geq \gamma(1-\gamma)^r$ for every $r > 0$ and $\gamma \in [0,1]$, we lower bound $\frac{1}{|\alpha|}$ and $\frac{1}{|\beta|}$ by $\gamma(1-\gamma)^{|\alpha|}$ and $\gamma(1-\gamma)^{|\beta|}$ respectively. The above is then lower bounded by

$$\gamma^2\mathop{\mathbb{E}}_{u,v,w}\sum_{i\in[L]}\left(\sum_{\alpha:i\in\widetilde{\pi}_{uv}(\alpha)}(1-\gamma)^{2|\alpha|}\widehat{f}_v(\alpha)^2\right)\left(\sum_{\beta:i\in\widetilde{\pi}_{uw}(\beta)}(1-\gamma)^{2|\beta|}\widehat{f}_w(\beta)^2\right) = \gamma^2\mathop{\mathbb{E}}_{u,v,w}\tau_{u,v,w}.$$

Since G is at most s-satisfiable, the labeling can satisfy at most s fraction of constraints and the above equation is upper bounded by s. □

Proof of Claim 10.16. It is easy to check that $\sum_{i\in[L]}\overline{\mathrm{Inf}}_i^u[T_{1-\gamma}f_v] \leq 1/\gamma$ (see Wenner [Wen13, Lemma 1.13]). For any u, v, w, since the test distribution satisfies the condi-

tions of Theorem 2.4, we get

$$\left|\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{i=1}^{k} T_{1-\gamma}f_v(x_i)T_{1-\gamma}f_w(y_i)\right] - \mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{i=1}^{k} T_{1-\gamma}f_v(x_i)\right]\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{i=1}^{k} T_{1-\gamma}f_w(y_i)\right]\right| \leq 2^{O(k)}\sqrt{\frac{\tau_{u,v,w}}{\gamma}}.$$

The claim follows by taking expectation over u, v, w and using the concavity of square root. □

Proof of Claim 10.15. We will add the $T_{1-\gamma}$ operator to one function at a time and upper bound the absolute value of the error incurred each time by $O(s^{c_0/8})$. The total error is at most $2k$ times the error in adding $T_{1-\gamma}$ to one function. Hence, it suffices to prove the following

$$\left|\mathop{\mathbb{E}}_{u,v,w}\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\prod_{i=1}^{k} f_v(x_i)f_w(y_i)\right] - \mathop{\mathbb{E}}_{u,v,w}\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[\left(\prod_{i=1}^{k-1} f_v(x_i)f_w(y_i)\right) f_v(x_k)T_{1-\gamma}f_w(y_k)\right]\right| \leq O(s^{c_0/8}). \tag{10.3.5}$$

Recall that X, Y denote the matrices chosen by test $\mathcal{T}_2$. Let Y_{-k} be the matrix obtained from Y by removing the kth row and $F_{u,v,w}(X, Y_{-k}) := \left(\prod_{i=1}^{k-1} f_v(x_i)f_w(y_i)\right) f_v(x_k)$. Then, (10.3.5) can be rewritten as

$$\left|\mathop{\mathbb{E}}_{u,v,w}\mathop{\mathbb{E}}_{\mathcal{T}_2}\left[F_{u,v,w}(X, Y_{-k})\left(I - T_{1-\gamma}\right)f_w(y_k)\right]\right| \leq O(s^{c_0/8}). \tag{10.3.6}$$

Let U be the operator that maps functions on the variable y_k, to one on the variables (X, Y_{-k}) defined by

$$(Uf)(X, Y_{-k}) := \mathop{\mathbb{E}}_{y_k|X,Y_{-k}} f(y_k).$$

Let $G_{u,v,w}(X, Y_{-k}) := \left(U(I - T_{1-\gamma})f_w\right)(X, Y_{-k})$. Note that $\mathbb{E}_{y\in\{0,1\}^{2R}} G_{u,v,w}(y) = 0$. For the rest of the analysis, fix u, v, w chosen by the test. We will omit the subscript u, v, w from now on for notational convenience. The domain of G can be thought of as $(\{0,1\}^{2k-1})^{2dL}$ and the test distribution on any row is independent across the blocks $\{B_{uv}(i) \cup B'_{uv}(i)\}_{i\in[L]}$. We now think of G as having domain $\prod_{i\in[L]} \Omega_i$ where

$\Omega_i = (\{0,1\}^{2k-1})^{2d}$ corresponds to the set of rows in $B_{uv}(i) \cup B'_{uv}(i)$. Let the following be the Efron-Stein decomposition of G with respect to $\mathcal{T}_2$,

$$G(X, Y_{-k}) = \sum_{\alpha \subseteq [L]} G_\alpha(X, Y_{-k}).$$

The following technical claim follows from a result similar to [DK13, Lemma 4.7] and then using [Mos10, Proposition 2.12]. We defer its proof to Section 10.3.1.

Claim 10.18. *For $\alpha \subseteq [L]$*

$$\|G_\alpha\|^2 \leq (1-\varepsilon)^{|\alpha|} \sum_{\beta \subseteq [2R] : \widetilde{\pi}_{uw}(\beta) = \alpha} \left(1 - (1-\gamma)^{2|\beta|}\right) \widehat{f}_w(\beta)^2 \tag{10.3.7}$$

where $\widetilde{\pi}_{uw}(\beta) := \{i \in [L] : \exists j \in [R], (j \in \beta \vee j + R \in \beta) \wedge \pi_{uv}(j) = i\}$.

Substituting the Efron-Stein decomposition of G, F into the LHS of (10.3.6) gives

$$\begin{aligned}
\left| \mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{\mathcal{T}_2} \left[F_{u,v,w}(X, Y_{-k}) \left(I - T_{1-\gamma}\right) f_w(y_k) \right] \right| &= \left| \mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{\mathcal{T}_2} F(X, Y_{-k}) G(X, Y_{-k}) \right| \\
&\overset{\substack{\text{(By orthonormality of} \\ \text{Efron-Stein decomposition)}}}{=} \left| \mathop{\mathbb{E}}_{u,v,w} \sum_{\alpha \subseteq [L]} \mathop{\mathbb{E}}_{\mathcal{T}_2} F_\alpha(X, Y_{-k}) G_\alpha(X, Y_{-k}) \right| \\
&\overset{\text{(By Cauchy-Schwarz inequality)}}{\leq} \mathop{\mathbb{E}}_{u,v,w} \sqrt{\sum_{\alpha \subseteq [L]} \|F_\alpha\|^2} \cdot \sqrt{\sum_{\alpha \subseteq [L]} \|G_\alpha\|^2} \\
&\overset{\left(\text{Using } \sum_{\alpha \subseteq [L]} \|F_\alpha\|^2 = \|F\|_2^2 = 1\right)}{\leq} \mathop{\mathbb{E}}_{u,v,w} \sqrt{\sum_{\alpha \subseteq [L]} \|G_\alpha\|^2}.
\end{aligned}$$

Using concavity of square root and substituting for $\|G_\alpha\|^2$ from Equation (10.3.7), we get that the above is upper bounded by

$$\sqrt{\sum_{\alpha \subseteq [L]} \sum_{\substack{\beta \subseteq [2R]: \\ \widetilde{\pi}_{uw}(\beta) = \alpha}} \underbrace{\mathop{\mathbb{E}}_{u,v,w} (1-\varepsilon)^{|\alpha|} \left(1 - (1-\gamma)^{2|\beta|}\right) \widehat{f}_w(\beta)^2}_{=: \mathrm{Term}_{u,w}(\alpha,\beta)}}.$$

We will now break the above summation into three different parts and bound each part separately.

$$\Theta_0 := \mathop{\mathbb{E}}_{u,w} \sum_{\alpha,\beta:|\alpha|\geq\frac{1}{\varepsilon s^{c_0/4}}} \mathrm{Term}_{u,w}(\alpha,\beta), \quad \Theta_1 := \mathop{\mathbb{E}}_{u,w} \sum_{\substack{\alpha,\beta:|\alpha|<\frac{1}{\varepsilon s^{c_0/4}} \\ |\beta|\leq\frac{2}{s^{1/4}\varepsilon^{1/c_0}}}} \mathrm{Term}_{u,w}(\alpha,\beta),$$

$$\Theta_2 := \mathop{\mathbb{E}}_{u,w} \sum_{\substack{\alpha,\beta:|\alpha|<\frac{1}{\varepsilon s^{c_0/4}} \\ |\beta|>\frac{2}{s^{1/4}\varepsilon^{1/c_0}}}} \mathrm{Term}_{u,w}(\alpha,\beta).$$

Upper Bounding Θ_0: When $|\alpha| > \frac{1}{\varepsilon s^{c_0/4}}$, $(1-\varepsilon)^{|\alpha|} < s^{c_0/4}$. Also since f_w is $\{+1,-1\}$ valued, sum of squares of Fourier coefficient is 1. Hence $|\Theta_0| < s^{c_0/4}$.

Upper Bounding Θ_1: When $|\beta| \leq \frac{2}{s^{1/4}\varepsilon^{1/c_0}}$,

$$1-(1-\gamma)^{2|\beta|} \leq 1-\left(1-\frac{4}{s^{1/4}\varepsilon^{1/c_0}}\gamma\right) = \frac{4}{s^{1/4}\varepsilon^{1/c_0}}\gamma = 4s^{c_0/4}.$$

Again since the sum of squares of Fourier coefficients is 1, $|\Theta_1| \leq 4s^{c_0/4}$.

Upper Bounding Θ_2: From the smooth projections property of Theorem 6.6, we have that for any $v \in V$ and β with $|\beta| > \frac{2}{s^{1/4}\varepsilon^{1/c_0}}$, the probability that $|\widetilde{\pi}_{uv}(\beta)| < 1/\varepsilon s^{c_0/4}$, for a random neighbor u, is at most $\varepsilon s^{c_0/4}$. Hence $|\Theta_2| \leq s^{c_0/4}$. □

10.3.1 Proof of Claim 10.18

We will be reusing the notation introduced in the long code test $\mathcal{T}_2$. We denote the $k \times 2d$ dimensional matrix $X|_{B(i)\cup B'(i)}$ by X^i and $Y|_{B(i)\cup B'(i)}$ by Y^i. Also by X^i_j, we mean the jth row of the matrix X^i and Y^i_{-k} is the first $k-1$ rows of Y^i. The spaces of the random variables X^i, X^i_j, Y^i_{-k} will be denoted by $\mathcal{X}^i, \mathcal{X}^i_j, \mathcal{Y}^i_{-k}$.

Before we proceed to the proof of claim, we need a few definitions and lemmas related to correlated spaces defined by Mossel [Mos10].

Definition 10.19. *Let $(\Omega_1 \times \Omega_2, \mu)$ be a finite correlated space, the correlation between Ω_1 and Ω_2 with respect to μ us defined as*

$$\rho(\Omega_1, \Omega_2; \mu) := \max_{\substack{f:\Omega_1\to\mathbb{R}, \mathbb{E}[f]=0, \mathbb{E}[f^2]\leq 1 \\ g:\Omega_2\to\mathbb{R}, \mathbb{E}[g]=0, \mathbb{E}[g^2]\leq 1}} \mathbb{E}_{(x,y)\sim\mu}[|f(x)g(y)|].$$

Definition 10.20 (Markov Operator). *Let $(\Omega_1 \times \Omega_2, \mu)$ be a finite correlated space, the Markov operator, associated with this space, denoted by U, maps a function $g : \Omega_2 \to \mathbb{R}$ to functions $Ug : \Omega_1 \to \mathbb{R}$ by the following map:*

$$(Ug)(x) := \mathbb{E}_{(X,Y)\sim\mu}[g(Y) \mid X = x].$$

The following results (from [Mos10]) provide a way to upper bound correlation of a correlated spaces.

Lemma 10.21 ([Mos10, Lemma 2.8]). *Let $(\Omega_1 \times \Omega_2, \mu)$ be a finite correlated space. Let $g : \Omega_2 \to \mathbb{R}$ be such that $\mathbb{E}_{(x,y)\sim\mu}[g(y)] = 0$ and $\mathbb{E}_{(x,y)\sim\mu}[g(y)^2] \leq 1$. Then, among all functions $f : \Omega_1 \to \mathbb{R}$ that satisfy $\mathbb{E}_{(x,y)\sim\mu}[f(x)^2] \leq 1$, the maximum value of $|\mathbb{E}[f(x)g(y)]|$ is given as:*

$$|\mathbb{E}[f(x)g(y)]| = \sqrt{\mathbb{E}_{(x,y)\sim\mu}[(Ug(x))^2]}.$$

Proposition 10.22 ([Mos10, Proposition 2.11]). *Let $(\prod_{i=1}^n \Omega_i^{(1)} \times \prod_{i=1}^n \Omega_i^{(2)}, \prod_{i=1}^n \mu_i)$ be a product correlated spaces. Let $g : \prod_{i=1}^n \Omega_i^{(2)} \to \mathbb{R}$ be a function and U be the Markov operator mapping functions form space $\prod_{i=1}^n \Omega_i^{(2)}$ to the functions on space $\prod_{i=1}^n \Omega_i^{(1)}$. If $g = \sum_{S\subseteq[n]} g_S$ and $Ug = \sum_{S\subseteq[n]} (Ug)_S$ be the Efron-Stein decom-*

position of g and Ug respectively then,

$$(Ug)_S = U(g_S)$$

i.e. the Efron-Stein decomposition commutes with Markov operators.

Proposition 10.23 ([Mos10, Proposition 2.12]). *Assume the setting of Proposition 10.22 and furthermore assume that $\rho(\Omega_i^{(1)}, \Omega_i^{(2)}; \mu_i) \leq \rho$ for all $i \in [n]$, then for all g it holds that*

$$\|U(g_S)\|_2 \leq \rho^{|S|} \|g_S\|_2.$$

We will prove the following claim.

Claim 10.24. *For each $i \in [L]$,*

$$\rho\left(\mathcal{X}^i \times \mathcal{Y}^i_{-k}, \mathcal{Y}^i_k; \mathcal{T}_2^i\right) \leq \sqrt{1-\varepsilon}.$$

Before proving this claim, first let's see how it leads to the proof of Claim 10.18.

Proof of Claim 10.18. Proposition 10.22 shows that the Markov operator U commutes with taking the Efron-Stein decomposition. Hence, $G_\alpha := (U((I - T_{1-\gamma})f_w))_\alpha = U((I - T_{1-\gamma})(f_w)_\alpha)$, where $(f_w)_\alpha$ is the Efron-Stein decomposition of f_w w.r.t the marginal distribution of $\mathcal{T}_2$ on $\prod_{i=1}^L \mathcal{Y}_k^i$ which is a uniform distribution. Therefore, $(f_w)_\alpha = \sum_{\substack{\beta \subseteq [2R], \\ \widetilde{\pi}_{uw}(\beta) = \alpha}} \hat{f}_w(\beta)\chi_\beta$. Using Proposition 10.23 and Claim 10.24, we have

$$\begin{aligned}\|G_\alpha\|_2^2 = \|U((I - T_{1-\gamma})(f_w)_\alpha)\|_2^2 &\leq (\sqrt{1-\varepsilon})^{2|\alpha|} \|(I - T_{1-\gamma})(f_w)_\alpha\|_2^2 \\ &= (1-\varepsilon)^{|\alpha|} \sum_{\beta \subseteq [2R] : \widetilde{\pi}_{uw}(\beta) = \alpha} \left(1 - (1-\gamma)^{2|\beta|}\right) \hat{f}_w(\beta)^2,\end{aligned}$$

where the norms are with respect to the marginals of $\mathcal{T}_2$ in the corresponding spaces.

□

Proof of Claim 10.24. Recall the random variable $c_2 \in \{*, 0, 1\}$ defined in Step 3 of test $\mathcal{T}_2$. Let g and f be the functions that satisfies $\mathbb{E}[g] = \mathbb{E}[f] = 0$ and $\mathbb{E}[g^2], \mathbb{E}[f^2] \leq 1$ such that $\rho\left(\mathcal{X}^i \times \mathcal{Y}^i_{-k}, \mathcal{Y}^i_k; \mathcal{T}^i_2\right) = \mathbb{E}[|fg|]$. Define the *Markov Operator*

$$Ug(X^i, Y^i_{-k}) = \mathop{\mathbb{E}}_{(\tilde{X},\tilde{Y})\sim\mathcal{T}^i_2}[g(\tilde{Y}_k) \mid (\tilde{X}, \tilde{Y}_{-k}) = (X^i, Y^i_{-k})].$$

By Lemma 10.21, we have

$$\begin{aligned}
\rho\left(\mathcal{X}^i \times \mathcal{Y}^i_{-k}, \mathcal{Y}^i_k; \mathcal{T}^i_2\right)^2 &\leq \mathop{\mathbb{E}}_{\mathcal{T}^i_2}[Ug(X^i, Y^i_{-k})^2] \\
&= (1-2\varepsilon) \mathop{\mathbb{E}}_{\mathcal{T}^i_2}[Ug(X^i, Y^i_{-k})^2 \mid c_2 = *] + \varepsilon \mathop{\mathbb{E}}_{\mathcal{T}^i_2}[Ug(X^i, Y^i_{-k})^2 \mid c_2 = 0] + \\
&\quad \varepsilon \mathop{\mathbb{E}}_{\mathcal{T}^i_2}[Ug(X^i, Y^i_{-k})^2 \mid c_2 = 1] \\
&\leq (1-2\varepsilon) + \varepsilon \mathop{\mathbb{E}}_{\mathcal{T}^i_2}[Ug(X^i, Y^i_{-k})^2 \mid c_2 = 0] + \varepsilon \mathop{\mathbb{E}}_{\mathcal{T}^i_2}[Ug(X^i, Y^i_{-k})^2 \mid c_2 = 1],
\end{aligned}$$

where the last inequality uses the fact that $\mathbb{E}_{\mathcal{T}^i_2}[Ug(X^i, Y^i_{-k})^2 \mid c_2 = *] = \mathbb{E}[g^2]$ which is at most 1. Consider the case when $c_2 = 0$. By definition, we have

$$\mathop{\mathbb{E}}_{\mathcal{T}^i_2}[Ug(X^i, Y^i_{-k})^2 \mid c_2 = 0] = \mathop{\mathbb{E}}_{\binom{X^i,}{Y^i_{-k}}\sim\mathcal{T}^i_2} \left(\mathop{\mathbb{E}}_{(\tilde{X},\tilde{Y})\sim\mathcal{T}^i_2}[g(\tilde{Y}_k) \mid (\tilde{X}, \tilde{Y}_{-k}) = (X^i, Y^i_{-k}) \wedge c_2 = 0] \right)^2.$$

Under the conditioning, for any fixed value of X^i, Y^i_{-k}, the value of $\tilde{Y}_k|_{B'(i)}$ is a uniformly random string whereas $\tilde{Y}_k|_{B(i)}$ is a fixed string (since the $parity$ of all columns in $B(i)$ is 1). Let $\mathcal{U}$ be the uniform distribution on $\{-1, +1\}^d$ and $\mathcal{P}(X^i, Y^i_{-k}) \in$

$\{+1,-1\}^d$ denotes the column wise parities of $\begin{bmatrix} X^i|_{B(i)} \\ Y^i_{-k}|_{B(i)} \end{bmatrix}$.

$$\begin{aligned}
\mathop{\mathbb{E}}_{\mathcal{T}_2^i}[Ug(X^i,Y^i_{-k})^2 \mid c_2=0] &= \mathop{\mathbb{E}}_{X^i,Y^i_{-k}\sim\mathcal{T}_2^i}\left(\mathop{\mathbb{E}}_{(\tilde{X},\tilde{Y})\sim\mathcal{T}_2^i}\left[g(\tilde{Y}_k) \mid {}^{(\tilde{X},\tilde{Y}_{-k})=(X^i,Y^i_{-k})\wedge}_{c_2=0}\right]\right)^2 \\
&= \mathop{\mathbb{E}}_{\substack{X^i,Y^i_{-k}\sim\mathcal{T}_2^i,\\ z=\mathcal{P}(X^i,Y^i_{-k})}}\left(\mathop{\mathbb{E}}_{r\sim\mathcal{U}}[g(-z,r)]\right)^2 \\
&= \mathop{\mathbb{E}}_{z\sim\mathcal{U}}\left(\mathop{\mathbb{E}}_{r\sim\mathcal{U}}[g(z,r)]\right)^2 \quad \text{(Since marginal on } z \text{ is uniform)} \\
&= \mathop{\mathbb{E}}_{z\sim\mathcal{U}}\left(\mathop{\mathbb{E}}_{r\in\mathcal{U}}\sum_{\alpha\subseteq B(i)\cup B'(i)}\hat{g}(\alpha)\chi_\alpha(z,r)\right)^2 \\
&= \mathop{\mathbb{E}}_{z\sim\mathcal{U}}\left(\sum_{\alpha\subseteq B(i)\cup B'(i)}\hat{g}(\alpha)\mathop{\mathbb{E}}_{r\in\mathcal{U}}[\chi_\alpha(z,r)]\right)^2 \\
&= \mathop{\mathbb{E}}_{z\sim\mathcal{U}}\left(\sum_{\alpha\subseteq B(i)}\hat{g}(\alpha)\chi_\alpha(z)\right)^2 = \sum_{\alpha\subseteq B(i)}\hat{g}(\alpha)^2.
\end{aligned}$$

Similarly we have,

$$\mathop{\mathbb{E}}_{\mathcal{T}_2^i}[Ug(X^i,Y^i_{-k})^2 \mid c_2=1] = \sum_{\alpha\subseteq B'(i)}\hat{g}(\alpha)^2.$$

Now we can bound the correlation as follows:

$$\begin{aligned}
\rho\left(\mathcal{X}^i\times\mathcal{Y}^i_{-k},\mathcal{Y}^i_k;\mathcal{T}_2^i\right)^2 &\leq (1-2\varepsilon)+\varepsilon\sum_{\alpha\subseteq B(i)}\hat{g}(\alpha)^2+\varepsilon\sum_{\alpha\subseteq B'(i)}\hat{g}(\alpha)^2 \\
&\leq (1-2\varepsilon)+\varepsilon\sum_{\alpha\subseteq B(i)\cup B'(i)}\hat{g}(\alpha)^2 \quad \text{(Using } \hat{g}(\varphi)=\mathbb{E}[g]=0\text{)} \\
&\leq (1-\varepsilon). \qquad \text{(Using } \mathbb{E}[g^2]\leq 1 \text{ and Parseval's Identity)}
\end{aligned}$$

□

10.4 Strong Hardness of 4-LIN

In this section, we prove Theorem 10.4. We give a reduction from an instance of LC, $G = (U, V, E, [L], [R], \{\pi_e\}_{e\in E})$ as in Definition 6.5, to a 4-LIN-CSP instance $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The set of variables $\mathcal{V}$ is $V \times \{0,1\}^{2R}$. Any assignment to $\mathcal{V}$ is given by a set of functions $f_v : \{0,1\}^{2R} \to \{0,1\}$, for each $v \in V$. The set of constraints $\mathcal{E}$ is given by the following test which checks whether f_v's are long codes of a good labeling to V.

Long Code Test $\mathcal{T}_3$:

1. Choose $u \in U$ uniformly and neighbors $v, w \in V$ of u uniformly and independently at random.

2. Choose x, x', z, z' uniformly and independently from $\{0,1\}^{2R}$ and y from $\{0,1\}^{2L}$. Choose $(\eta, \eta') \in \{0,1\}^{2L} \times \{0,1\}^{2L}$ as follows: Independently for each $i \in [L]$, $(\eta_i, \eta_{L+i}, \eta'_i, \eta'_{L+i})$ is set to

   (a) $(0,0,0,0)$ with probability $1 - 2\varepsilon$,

   (b) $(1,0,1,0)$ with probability ε and

   (c) $(0,1,0,1)$ with probability ε.

3. For $y \in \{0,1\}^{2L}$, let $y \circ \pi_{uv} \in \{0,1\}^{2R}$ be the string such that $(y \circ \pi_{uv})_i := y_{\pi_{uv}(i)}$ for $i \in [R]$ and $(y \circ \pi_{uv})_i := y_{\pi_{uv}(i-R)+L}$ otherwise. Given $\eta \in \{0,1\}^{2L}, z \in \{0,1\}^{2R}$, the string $\eta \circ \pi_{uv} \cdot z \in \{0,1\}^{2R}$ is obtained by taking coordinate-wise product of $\eta \circ \pi_{uv}$ and z. Accept iff

$$f_v(x)+f_v(x+y\circ\pi_{uv}+\eta\circ\pi_{uv}\cdot z)+f_w(x')+f_w(x'+y\circ\pi_{uw}+\eta'\circ\pi_{uw}\cdot z'+1) = 1 \pmod 2. \tag{10.4.1}$$

   (Here by addition of strings, we mean the coordinate-wise sum modulo 2.)

Lemma 10.25 (Completeness). *If G is an YES instance of* LC*, then there exists f, g such that each of them covers $1-\varepsilon$ fraction of $\mathcal{E}$ and they together cover all of $\mathcal{E}$.*

Proof. Let $\ell : U \cup V \to [L] \cup [R]$ be a labeling to G that satisfies all the constraints. Consider the assignments given by $f_v(x) := x_{\ell(v)}$ and $g_v(x) := x_{R+\ell(v)}$ for each $v \in V$. On input f_v, for any pair of edges $(u,v),(u,w) \in E$, and $x, x', z, z', \eta, \eta', y$ chosen by the long code test $\mathcal{T}_3$, the LHS in (10.4.1) evaluates to

$$x_{\ell(v)}+x_{\ell(v)}+y_{\ell(u)}+\eta_{\ell(u)}z_{\ell(v)}+x'_{\ell(w)}+x'_{\ell(w)}+y_{\ell(u)}+\eta'_{\ell(u)}z'_{\ell(w)}+1 = \eta_{\ell(u)}z_{\ell(v)}+\eta'_{\ell(u)}z'_{\ell(w)}+1.$$

Similarly for g_v, the expression evaluates to $\eta_{L+\ell(u)}z_{R+\ell(v)} + \eta'_{L+\ell(u)}z'_{R+\ell(w)} + 1$. Since $(\eta_i, \eta'_i) = (0,0)$ with probability $1-\varepsilon$, each of f, g covers $1-\varepsilon$ fraction of $\mathcal{E}$. Also for $i \in [L]$ whenever $(\eta_i, \eta'_i) = (1,1)$, $(\eta_{L+i}, \eta'_{L+i}) = (0,0)$ and vice versa. So one of the two evaluations above is $1 \pmod 2$. Hence the pair of assignment f, g cover $\mathcal{E}$. □

Lemma 10.26 (Soundness). *Let c_0 be the constant from Theorem 6.6. If G is at most s-satisfiable with $s < \frac{\delta^{10/c_0+5}}{4}$, then any independent set in $\mathcal{G}$ has fractional size at most δ.*

Proof. Let $I \subseteq \mathcal{V}$ be an independent set of fractional size δ in the constraint graph $\mathcal{G}$. For every variable $v \in V$, let $f_v : \{0,1\}^{2R} \to \{0,1\}$ be the indicator function of the independent set restricted to the vertices that correspond to v. Since I is an independent set, we have

$$\mathop{\mathbb{E}}_{u,v,w} \mathop{\mathbb{E}}_{\substack{x,x',\\ z,z',\\ \eta,\eta',y}} [f_v(x) f_v(x + y \circ \pi_{uv} + \eta \circ \pi_{uv} \cdot z) f_w(x') f_w(x' + y \circ \pi_{uw} + \eta' \circ \pi_{uw} \cdot z' + 1)] = 0. \tag{10.4.2}$$

For $\alpha \subseteq [2R]$, let $\pi^{\oplus}_{uv}(\alpha) \subseteq [2L]$ be the set containing elements $i \in [2L]$ such that if $i < L$ there are an odd number of $j \in [R] \cap \alpha$ with $\pi_{uv}(j) = i$ and if $i \geq L$ there are an odd number of $j \in ([2R] \setminus [R]) \cap \alpha$ with $\pi_{uv}(j-R) = i - L$. It is easy to see that

$\chi_\alpha(y \circ \pi_{uw}) = \chi_{\pi^{\oplus}_{uv}(\alpha)}(y)$. Expanding f_v in the Fourier basis and taking expectation over x, x' and y, we get that

$$\mathop{\mathbb{E}}_{u,v,w} \sum_{\alpha,\beta\subseteq[2R]:\pi^{\oplus}_{uv}(\alpha)=\pi^{\oplus}_{uw}(\beta)} \widehat{f}_v(\alpha)^2 \widehat{f}_w(\beta)^2 (-1)^{|\beta|} \mathop{\mathbb{E}}_{z,z',\eta,\eta'} [\chi_\alpha(\eta \circ \pi_{uv} \cdot z)\chi_\beta(\eta' \circ \pi_{uw} \cdot z')] = 0. \tag{10.4.3}$$

Now the expectation over z, z' simplifies as

$$\mathop{\mathbb{E}}_{u,v,w} \sum_{\alpha,\beta\subseteq[2R]:\pi^{\oplus}_{uv}(\alpha)=\pi^{\oplus}_{uw}(\beta)} \underbrace{\widehat{f}_v(\alpha)^2 \widehat{f}_w(\beta)^2 (-1)^{|\beta|} \Pr_{\eta,\eta'}[\alpha \cdot (\eta \circ \pi_{uv}) = \beta \cdot (\eta' \circ \pi_{uw}) = \bar{0}]}_{=:\mathrm{Term}_{u,v,w}(\alpha,\beta)} = 0, \tag{10.4.4}$$

where we think of α, β as the characteristic vectors in $\{0,1\}^{2R}$ of the corresponding sets. We will now break up the above summation into different parts and bound each part separately. For a projection $\pi : [R] \to [L]$, define $\widetilde{\pi}(\alpha) := \{i \in [L] : \exists j \in [R], (j \in \alpha \vee j + R \in \alpha) \wedge (\pi(j) = i)\}$. We need the following definitions.

$$\Theta_0 := \mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\alpha,\beta:\\ \pi^{\oplus}_{uv}(\alpha)=\pi^{\oplus}_{uw}(\beta)=\emptyset}} \mathrm{Term}_{u,v,w}(\alpha,\beta),$$

$$\Theta_1 := \mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\alpha,\beta:\\ \pi^{\oplus}_{uv}(\alpha)=\pi^{\oplus}_{uw}(\beta)\neq\emptyset,\\ \max\{|\alpha|,|\beta|\}\leq 2/\delta^{5/c_0}}} \mathrm{Term}_{u,v,w}(\alpha,\beta),$$

$$\Theta_2 := \mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\alpha,\beta:\\ \pi^{\oplus}_{uv}(\alpha)=\pi^{\oplus}_{uw}(\beta)\neq\emptyset,\\ \max\{|\widetilde{\pi}_{uv}(\alpha)|,|\widetilde{\pi}_{uw}(\beta)|\}\geq 1/\delta^5}} \mathrm{Term}_{u,v,w}(\alpha,\beta),$$

$$\Theta_3 := \mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\alpha,\beta:\\ \pi^{\oplus}_{uv}(\alpha)=\pi^{\oplus}_{uw}(\beta)\neq\emptyset,\\ \max\{|\alpha|,|\beta|\}>2/\delta^{5/c_0},\\ \max\{|\widetilde{\pi}_{uv}(\alpha)|,|\widetilde{\pi}_{uw}(\beta)|\}<1/\delta^5}} \mathrm{Term}_{u,v,w}(\alpha,\beta).$$

Lower Bounding Θ_0: If $\pi^{\oplus}_{uw}(\beta) = \emptyset$, then $|\beta|$ is even. Hence, all the terms in Θ_0 are positive and

$$\Theta_0 \geq \mathop{\mathbb{E}}_{u,v,w} \mathrm{Term}_{u,v,w}(0,0) = \mathop{\mathbb{E}}_{u} \left(\mathop{\mathbb{E}}_{v} \widehat{f}_v(0)^2 \right)^2 \geq \left(\mathop{\mathbb{E}}_{u,v} \widehat{f}_v(0) \right)^4 = \delta^4.$$

Upper Bounding Θ_1: Consider the following strategy for labeling vertices $u \in U$ and $v \in V$. For $u \in U$, pick a random neighbor v, choose α with probability $\widehat{f}_v(\alpha)^2$ and set its label to a random element in $\widetilde{\pi}_{uv}(\alpha)$. For $w \in V$, choose β with probability $\widehat{f}_w(\beta)^2$ and set its label to a random element of β. If the label $j \geq R$, change the label to $j - R$. The probability that a random edge (u, w) of the label cover is satisfied by this labeling is

$$\begin{aligned}\mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\alpha,\beta: \\ \widetilde{\pi}_{uv}(\alpha) \cap \widetilde{\pi}_{uw}(\beta) \neq \emptyset}} \widehat{f}_v(\alpha)^2 \widehat{f}_w(\beta)^2 \frac{1}{|\widetilde{\pi}_{uv}(\alpha)| \cdot |\beta|} &\geq \mathop{\mathbb{E}}_{u,v,w} \sum_{\substack{\alpha,\beta: \\ \pi^{\oplus}_{uv}(\alpha) = \pi^{\oplus}_{uw}(\beta) \neq \emptyset \\ \max\{|\alpha|,|\beta|\} \leq 2/\delta^{5/c_0}}} \widehat{f}_v(\alpha)^2 \widehat{f}_w(\beta)^2 \frac{\delta^{10/c_0}}{4} \\ &\geq |\Theta_1| \cdot \frac{\delta^{10/c_0}}{4}.\end{aligned}$$

Since the instance is at most s-satisfiable, the above is upper bounded by s. Choosing $s < \frac{\delta^{10/c_0+5}}{4}$, will imply $|\Theta_1| \leq \delta^5$.

Upper Bounding Θ_2: Suppose $|\widetilde{\pi}_{uv}(\alpha)| \geq 1/\delta^5$, then note that

$$\Pr_{\eta,\eta'}[\alpha \cdot (\eta \circ \pi_{uv}) = \beta \cdot (\eta' \circ \pi_{uw}) = 0] \leq \Pr_{\eta}[\alpha \cdot (\eta \circ \pi_{uv}) = 0] \leq (1-\varepsilon)^{|\widetilde{\pi}_{uv}(\alpha)|} \leq (1-\varepsilon)^{1/\delta^5}.$$

Since the sum of squares of Fourier coefficients of f is less than 1 and ε is a constant, we get that $|\Theta_2| \leq 1/2^{\Omega(1/\delta^5)} < O(\delta^5)$.

Upper Bounding Θ_3: From the smooth projections property of Theorem 6.6, we have that for any $v \in V$ and $\alpha \subseteq [2R]$ with $|\alpha| > 2/\delta^{5/c_0}$, the probability that

$|\widetilde{\pi}_{uv}(\alpha)| < 1/\delta^5$, for a random neighbor u of v, is at most δ^5. Hence $|\Theta_3| \leq \delta^5$.

On substituting the above bounds in Equation (10.4.4), we get that $\delta^4 - O(\delta^5) \leq 0$ which gives a contradiction for small enough δ. Hence there is no independent set in $\mathcal{G}$ of size δ. □

Proof of Theorem 10.4. From Theorem 6.6, the size of the CSP instance $\mathcal{G}$ produced by the reduction is $N = n^r 2^{2^{O(r)}}$ and the parameter $s \leq 2^{-d_0 r}$. Setting $r = \Theta(\log \log n)$, gives that $N = 2^{\text{poly}(\log n)}$ and the size of the largest independent set $\delta = 1/\text{poly}(\log n) = 1/\text{poly}(\log N)$. □

References

[ADFS04] Noga Alon, Irit Dinur, Ehud Friedgut, and Benny Sudakov. *Graph products, fourier analysis and spectral techniques*. Geometric & Functional Analysis GAFA, 14(5):913–940, 2004. doi:10.1007/s00039-004-0478-3.

[ALM+98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. *Proof verification and the hardness of approximation problems*. J. ACM, 45(3):501–555, May 1998. (Preliminary version in *33rd FOCS*, 1992). eccc:TR98-008, doi:10.1145/278298.278306.

[AS98] Sanjeev Arora and Shmuel Safra. *Probabilistic checking of proofs: A new characterization of NP*. J. ACM, 45(1):70–122, January 1998. (Preliminary version in *33rd FOCS*, 1992). doi:10.1145/273865.273901.

[BGH+12] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer. *Making the long code shorter*. In *Proc. 53rd IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 370–379. 2012. arXiv:1111.0405, doi:10.1109/FOCS.2012.83.

[BGS98] Mihir Bellare, Oded Goldreich, and Madhu Sudan. *Free bits, PCPs, and nonapproximability—towards tight results*. SIAM J. Computing, 27(3):804–915, June 1998. (Preliminary version in *36th FOCS*, 1995). eccc:TR95-024, doi:10.1137/S0097539796302531.

[BHV15] Amey Bhangale, Prahladh Harsha, and Girish Varma. *A Characterization of Hard-to-cover CSPs*. In David Zuckerman, ed., *30th Conference on Computational Complexity (CCC 2015)*, volume 33 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 280–303. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2015. doi:http://dx.doi.org/10.4230/LIPIcs.CCC.2015.280.

[BK97] Avrim Blum and David R. Karger. *An $\tilde{O}(n^{3/14})$-coloring algorithm for 3-colorable graphs*. Inf. Process. Lett., 61(1):49–53, 1997.

[BK10] Nikhil Bansal and Subhash Khot. *Inapproximability of hypergraph vertex cover and applications to scheduling problems*. In Samson Abramsky, Cyril Gavoille, Claude Kirchner, Friedhelm Meyer auf der Heide, and Paul G. Spirakis, eds., *Proc. 37th International Colloquium of Automata, Languages and Programming (ICALP), Part I*, volume 6198 of *LNCS*, pages 250–261. Springer, 2010. doi:10.1007/978-3-642-14165-2_22.

[BKS+10] Arnab Bhattacharyya, Swastik Kopparty, Grant Schoenebeck, Madhu Sudan, and David Zuckerman. *Optimal testing of Reed-Muller codes*. In *Proc. 51st IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 488–497. 2010. arXiv:0910.0641, doi:10.1109/FOCS.2010.54.

[Blu94] Avrim Blum. *New approximation algorithms for graph coloring*. Journal of the ACM, 41:470–516, 1994.

[DG14] Irit Dinur and Venkatesan Guruswami. *PCPs via low-degree long code and hardness for constrained hypergraph coloring*. In *Proc. 54th IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 340–349. 2014. eccc:TR13-122, doi:10.1109/FOCS.2013.44.

[DGJ+10] Ilias Diakonikolas, Parikshit Gopalan, Ragesh Jaiswal, Rocco A. Servedio, and Emanuele Viola. *Bounded independence fools halfspaces*. SIAM J. Computing, 39(8):3441–3462, 2010. (Prelimimary version in *50th FOCS*, 2009). arXiv:0902.3757, doi:10.1137/100783030.

[DGKR05] Irit Dinur, Venkatesan Guruswami, Subhash Khot, and Oded Regev. *A new multilayered PCP and the hardness of hypergraph vertex cover*. SIAM J. Computing, 34(5):1129–1146, 2005. (Preliminary version in *35th STOC*, 2003). arXiv:cs.CC/0304026, doi:10.1137/S0097539704443057.

[DHSV15] Irit Dinur, Prahladh Harsha, Srikanth Srinivasan, and Girish Varma. *Derandomized Graph Product Results Using the Low Degree Long Code*. In Ernst W. Mayr and Nicolas Ollinger, eds., *32nd International Symposium on Theoretical Aspects of Computer Science (STACS 2015)*, volume 30 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 275–287. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2015. doi:http://dx.doi.org/10.4230/LIPIcs.STACS.2015.275.

[DK13] Irit Dinur and Gillat Kol. *Covering CSPs*. In *Proc. 28th IEEE Conference on Computational Complexity*, pages 207–218. 2013. eccc:TR12-088, doi:10.1109/CCC.2013.29.

[DMR09] Irit Dinur, Elchanan Mossel, and Oded Regev. *Conditional hardness for approximate coloring*. SIAM J. Computing, 39(3):843–873, 2009. (Preliminary version in *38th STOC*, 2006). arXiv:cs/0504062, doi:10.1137/07068062X.

[DRS05] Irit Dinur, Oded Regev, and Clifford D. Smyth. *The hardness of 3-uniform hypergraph coloring*. Combinatorica, 25(5):519–535, 2005. (Preliminary version in *43rd FOCS*, 2002). doi:10.1007/s00493-005-0032-4.

[DS10] Irit Dinur and Igor Shinkar. *On the conditional hardness of coloring a 4-colorable graph with super-constant number of colors*. In Maria J. Serna, Ronen Shaltiel, Klaus Jansen, and José D. P. Rolim, eds., *Proc. 13th*

International Workshop on Randomization and Approximation Techniques in Computer Science (APPROX), volume 6302 of *LNCS*, pages 138–151. Springer, 2010.

[Fei98] Uriel Feige. *A threshold of* ln *n for approximating set cover*. J. ACM, 45(4):634–652, July 1998. (Preliminary version in *28th STOC*, 1996). doi:10.1145/285055.285059.

[FK00] Uriel Feige and Joe Kilian. *Zero-knowledge and the chromatic number*. J. Computer and System Sciences, 60(2):337–353, April 2000. (Preliminary version in 12*th Conference on Computational Complexity*, 1997).

[FKN02] Ehud Friedgut, Gil Kalai, and Assaf Naor. *Boolean functions whose fourier transform is concentrated on the first two levels*. Advances in Applied Mathematics, 29(3):427 – 437, 2002. doi:http://dx.doi.org/10.1016/S0196-8858(02)00024-6.

[GHH+14] Venkat Guruswami, Prahladh Harsha, Johan Håstad, Srikanth Srinivasan, and Girish Varma. *Super-polylogarithmic hypergraph coloring hardness via low-degree long codes*. In *Proc.* 46*th ACM Symp. on Theory of Computing (STOC)*, pages 614–623. 2014. arXiv:1311.7407, doi:10.1145/2591796.2591882.

[GHS02] Venkatesan Guruswami, Johan Håstad, and Madhu Sudan. *Hardness of approximate hypergraph coloring*. SIAM J. Computing, 31(6):1663–1686, 2002. (Preliminary Version in *41st FOCS*, 2000). doi:10.1137/S0097539700377165.

[GK04] Venkatesan Guruswami and Sanjeev Khanna. *On the hardness of 4-coloring a 3-colorable graph*. SIAM J. Discrete Math., 18(1):30–40, 2004.

[GL15] Venkatesan Guruswami and Euiwoong Lee. *Strong inapproximability results on balanced rainbow-colorable hypergraphs*. In *Proc.* 26*th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 822–836. 2015. doi:10.1137/1.9781611973730.56.

[GOWZ10] Parikshit Gopalan, Ryan O'Donnell, Yi Wu, and David Zuckerman. *Fooling functions of halfspaces under product distributions*. In *Computational Complexity (CCC), 2010 IEEE 25th Annual Conference on*, pages 223–234. June 2010. doi:10.1109/CCC.2010.29.

[Hås01] Johan Håstad. *Some optimal inapproximability results*. J. ACM, 48(4):798–859, July 2001. (Preliminary version in *29th STOC*, 1997). doi:10.1145/502090.502098.

[HSS13] Elad Haramaty, Amir Shpilka, and Madhu Sudan. *Optimal testing of multivariate polynomials over small prime fields*. SIAM J. Computing, 42(2):536–562, 2013. (Preliminary version in *52nd FOCS*, 2011). eccc:TR11-059, doi:10.1137/120879257.

[Kho02a] Subhash Khot. *Hardness results for approximate hypergraph coloring*. In *Proc. 34th ACM Symp. on Theory of Computing (STOC)*, pages 351–359. 2002. doi:10.1145/509907.509962.

[Kho02b] ———. *Hardness results for coloring 3-colorable 3-uniform hypergraphs*. In *Proc. 43rd IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 23–32. 2002. doi:10.1109/SFCS.2002.1181879.

[Kho02c] ———. *On the power of unique 2-prover 1-round games*. In *Proc. 34th ACM Symp. on Theory of Computing (STOC)*, pages 767–775. 2002. doi:10.1145/509907.510017.

[KKMO07] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. *Optimal inapproximability results for MAX-CUT and other 2-variable CSPs?* SIAM J. Computing, 37(1):319–357, 2007. (Preliminary version in *45th FOCS*, 2004). eccc:TR05-101, doi:10.1137/S0097539705447372.

[KLS00] Sanjeev Khanna, Nathan Linial, and Shmuel Safra. *On the hardness of approximating the chromatic number*. Combinatorica, 20(3):393–415, 2000. doi:10.1007/s004930070013.

[KM13] Daniel M. Kane and Raghu Meka. *A PRG for Lipschitz functions of polynomials with applications to sparsest cut*. In *Proc. of the 45th ACM Symp. on Theory of Computing*, STOC '13, pages 1–10. ACM, New York, NY, USA, 2013. doi:10.1145/2488608.2488610.

[KMS98] David R. Karger, Rajeev Motwani, and Madhu Sudan. *Approximate graph coloring by semidefinite programming*. J. ACM, 45(2):246–265, 1998. doi:10.1145/274787.274791.

[KR08] Subhash Khot and Oded Regev. *Vertex cover might be hard to approximate to within 2-ε*. J. Computer and System Sciences, 74(3):335–349, 2008. (Preliminary version in *18th IEEE Conference on Computational Complexity*, 2003). doi:10.1016/j.jcss.2007.06.019.

[KS14a] Subhash Khot and Rishi Saket. *Hardness of coloring 2-colorable 12-uniform hypergraphs with $\exp(\log^{\Omega(1)} n)$ colors*. In *Proc. 55th IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 206–215. 2014. doi:10.1109/FOCS.2014.30.

[KS14b] ———. *Hardness of finding independent sets in 2-colorable and almost 2-colorable hypergraphs*. In *Proc. 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1607–1625. 2014. arXiv:1308.3247.

[KT14] Ken ichi Kawarabayashi and Mikkel Thorup. *Coloring 3-colorable graphs with $o(n^{1/5})$ colors*. In Ernst W. Mayr and Natacha Portier, eds., *STACS 2014*, volume 25 of *LIPIcs*, pages 458–469. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2014. doi:10.4230/LIPIcs.STACS.2014.458.

[LN97] Rudolf Lidl and Harald Niederreiter. *Finite Fields*, volume 2 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1997. doi:10.1017/CBO9780511525926.

[Mos08] Elchanan Mossel. *Gaussian bounds for noise correlation of functions and tight analysis of long codes.* In *Proc. 49th IEEE Symp. on Foundations of Comp. Science (FOCS)*, pages 156–165. 2008. arXiv:math/0703683, doi:10.1109/FOCS.2008.44.

[Mos10] ———. *Gaussian bounds for noise correlation of functions*. 19(6):1713–1756, 2010. (Preliminary version in *49th FOCS*, 2008). arXiv:math/0703683, doi:10.1007/s00039-010-0047-x.

[Raz98] Ran Raz. *A parallel repetition theorem*. SIAM J. Computing, 27(3):763–803, June 1998. (Preliminary version in *27th STOC*, 1995). doi:10.1137/S0097539795280895.

[Var14] Girish Varma. *Reducing uniformity in khot-saket hypergraph coloring hardness reductions*. CoRR, abs/1408.0262, 2014.

[Wen13] Cenny Wenner. *Circumventing d-to-1 for approximation resistance of satisfiable predicates strictly containing parity of width at least four*. Theory of Computing, 9(23):703–757, 2013. (Preliminary version in *APPROX*, 2012). eccc:TR12-145, doi:10.4086/toc.2013.v009a023.

[Wig83] Avi Wigderson. *Improving the performance guarantee for approximate graph coloring*. J. ACM, 30(4):729–735, October 1983. doi:10.1145/2157.2158.

A template that can be used to format a TIFR PhD thesis with this look and feel can be found online at https://github.com/geevi/tifr-thesis. It is a modified version of https://github.com/suchow/Dissertate, by Jordan Suchow.