Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 177 (2010)

Bypassing UGC from some Optimal Geometric Inapproximability Results

Venkatesan Guruswami* Prasad Raghavendra† Rishi Saket‡ Yi Wu§

January 24, 2015

Abstract

The Unique Games conjecture (UGC) has emerged in recent years as the starting point for several optimal inapproximability results. While for *none* of these results a reverse reduction to Unique Games is known, the assumption of bijective projections in the Label Cover instance nevertheless seems critical in these proofs. In this work we bypass the UGC assumption in inapproximability results for two geometric problems, obtaining a tight NP-hardness result in each case.

The first problem, known as L_p Subspace Approximation, is a generalization of the classic least squares regression problem. Here, the input consists of a set of points $S = \{a_1, \ldots, a_m\} \subseteq \mathbb{R}^n$ and a parameter k (possibly depending on n). The goal is to find a k-dimensional subspace H of $\mathbb{R}^n$ that minimizes the ℓ_p norm of the Euclidean distances to the points in S. For $p = 2$, $k = n-1$, this reduces to the least squares regression problem, while for $p = \infty$, $k = 0$ it reduces to the problem of finding a ball of minimum radius enclosing all the points. We show that for any fixed p ($2 < p < \infty$), and for $k = n-1$, it is NP-hard to approximate this problem to within a factor of $\gamma_p - \varepsilon$ for constant $\varepsilon > 0$, where γ_p is the pth norm of a standard Gaussian random variable. This matches the γ_p approximation algorithm obtained by Deshpande, Tulsiani and Vishnoi [11] who also showed the same hardness result under the Unique Games Conjecture.

The second problem we study is the related L_p Quadratic Grothendieck Maximization Problem, considered by Kindler, Naor and Schechtman [26]. Here, the input is a multilinear quadratic form $\sum_{i,j=1}^{n} a_{ij}x_ix_j$ and the goal is to maximize the quadratic form over the ℓ_p unit ball, namely all x with $\sum_{i=1}^{n} |x_i|^p = 1$. The problem is polynomial time solvable for $p = 2$. We show that for any constant p ($2 < p < \infty$), it is NP-hard to approximate the quadratic form to within a factor of $\gamma_p^2 - \varepsilon$ for any $\varepsilon > 0$. The same hardness factor was shown under the UGC in [26]. We also obtain a γ_p^2-approximation algorithm for the problem using the convex relaxation of the problem defined by [26]. A γ_p^2 approximation algorithm has also been independently obtained by Naor and Schechtman [29].

These are the *first* approximation thresholds, proven under P $\neq$ NP, that involve the Gaussian random variable in a fundamental way. Note that the problem statements themselves have no mention of Gaussians.

*Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213. guruswami@cmu.edu. Research supported in part by a Packard Fellowship.

†College of Computing, Georgia Institute of Technology, Atlanta, GA. raghavendra@cc.gatech.edu.

‡Center for Computational Intractability, Princeton University, Princeton, NJ. rsaket@cs.princeton.edu. Research supported in part by NSF 0832797. Some of this work was done when the author was at Carnegie Mellon University.

§IBM Almaden Research Center, San Jose, CA. wuyi@us.ibm.com. Some of this work was done when the author was at Carnegie Mellon University.

ISSN 1433-8092

Contents

1 Introduction

The Unique Games Conjecture of Khot [21] asserts that a certain binary constraint satisfaction problem is hard to approximate over a large enough alphabet. The conjecture has been shown in recent years to imply optimal hardness results for various important combinatorial optimization problems such as Maximum Cut [22], Vertex Cover [23] and more generally, constraint satisfaction problems [32] and ordering problems [17]. However, arguably there has been little progress towards proving the conjecture. On the contrary, recent algorithmic results have disproved some stronger variants of the conjecture [3], and solved the Unique Games problem on special classes of instances like expanders [5] and more generally graphs with few "bad" eigenvalues [18, 8]. Moreover, while the Unique Games Conjecture is known to imply optimal inapproximability results for Maximum Cut, Vertex Cover and several other problems, the converse is unknown in each case. (In other words, we only know Unique-Games hardness results, but no "Unique-Games completeness" results.) This leaves the possibility open that while the implications of the conjecture are true, the conjecture itself is false.

For all these reasons, it is a worthwhile endeavor to investigate if the optimal inapproximability results obtained via the Unique Games Conjecture can be shown without appealing to the conjecture. In this work, we consider two geometric problems for which optimal inapproximability results based on the Unique Games Conjecture have been shown previously, and obtain the same hardness results unconditionally, i.e., without appealing to the conjecture.

1.1 Our Main Results

L_p Subspace Approximation Problem The first problem we consider is the L_p Subspace Approximation Problem for $2 \leqslant p < \infty$ – a natural generalization of the least squares regression problem, the low rank matrix approximation problem and the problem of computing radii of point sets. Here the input consists of a set of points $S = \{a_1, \dots, a_m\} \subseteq \mathbb{R}^n$, and an integer $1 \leqslant k \leqslant n$. The goal is to find a k-dimensional subspace H of $\mathbb{R}^n$ that minimizes the ℓ_p norm of the Euclidean distances to the points in S. Formally, the goal is to compute:

$$\mathrm{Sub}_p(S,k) = \min_{H \subseteq \mathbb{R}^n : \dim(H) = k} \left(\sum_{i=1}^{m} \mathrm{dist}(H, a_i)^p \right)^{1/p}, \tag{1}$$

where dist(,) is the usual ℓ_2 distance between a subspace and a point. Informally, it is the problem of determining how *close* a given set of points is from lying in a smaller subspace, where the measure of closeness to a subspace is the ℓ_p norm of the tuple of Euclidean distances of the set of points from the subspace. Such problems arise naturally in classification of large data sets for applications in machine learning and data mining. As an algorithmic question, it is a generalization of various special cases for different values of p such as Low rank matrix approximation ($p = 2$) or Computing the radii of point sets ($p = \infty$). Generally speaking, optimizing loss functions with small p is more robust againist outlier points while optimizing with a large p gives stronger theoretical guarantee. We refer the reader to [11] for a more comprehensive discussion of these connections.

In this work we focus on the hardness of approximating the L_p Subspace Approximation Problem for the case when $k = n - 1$, i.e., the problem of finding a hyperplane that is closest to the set of points in the measure defined above. Let γ_p denote the p'th norm of a normal random variable. Recently, Deshpande, Tulsiani and Vishnoi [11] obtained a γ_p approximation for the problem, and showed a matching hardness assuming the Unique Games Conjecture. Bypassing the need for the UGC, we obtain a $(1-\varepsilon)\gamma_p$ hardness of approximation unconditionally. The following theorem is proved in Section 4.

Theorem 1.1. *For any given p ($2 < p < \infty$) the L_p Subspace Approximation Problem is* NP*-hard to approximate within a factor of* $(1-\varepsilon)\gamma_p$ *for any* $\varepsilon > 0$.

L_p **Grothendieck Problem.** The second problem we consider is that of maximizing a multilinear quadratic form over the unit ℓ_p ball in $\mathbb{R}^n$ for constant p, $2 \leqslant p < \infty$. Formally, the input to the problem is a symmetric $n \times n$ matrix $A = (a_{ij})$ with zero diagonal entries, the goal is to compute the following quantity,

$$\mathrm{Val}_p(A) := \max\left\{ \sum_{i,j=1}^{n} a_{ij} x_i x_j \;\middle|\; \sum_i |x_i|^p \leqslant 1 \right\}, \tag{2}$$

We refer to this problem as the L_p Quadratic Grothendieck Maximization Problem. In the case where $p = 2$, $\mathrm{Val}_2(A)$ is nothing but the maximum eigenvalue of the matrix A and hence is computationally tractable. The case $p = \infty$ is commonly referred to as the Grothendieck problem and has been extensively studied in mathematics and computer science for its applications to combinatorial optimization, graph theory and correlation clustering [30, 2, 1, 10]. The case when $2 < p < \infty$ has applications towards studying spin glass systems in physics (See [26]).

Kindler, Naor and Schechtman [26] obtained a $(\gamma_p^2 - \varepsilon)$-hardness for every ε for the L_p-Grothendieck problem assuming the UGC, and also exhibited an *almost* matching $\frac{p}{e} + 30\log p$-approximation algorithm. Bypassing the Unique Games Conjecture, we obtain a γ_p^2 hardness unconditionally for the problem. We also obtain an approximation algorithm that *exactly* matches the hardness result for every p.

Theorem 1.2. *For any constant $p > 2$, it is* NP*-hard to approximate* $\mathrm{Val}_p(A)$ *to within* $\gamma_p^2 - \varepsilon$ *for any* $\varepsilon > 0$*, where A is a symmetric matrix with all diagonal entries zero.*

Theorem 1.3. *There is a polynomial time algorithm to approximate* $\mathrm{Val}_p(A)$ *to within* γ_p^2 *for any symmetric matrix A with all diagonal entries equal to zero.*

A γ_p^2 approximation for $\mathrm{Val}_p(A)$ has also been independently obtained by Naor and Schechtman [29] as part of a more general result. Theorems 1.2 and 1.3 are proved in Sections 5 and 6 respectively.

1.2 Discussion of Our Results

The above mentioned NP-hardness results are noteworthy for the following reasons.

First, the inapproximability factors for both problems are irrational numbers arising from the Gaussian distribution, although neither of the problems involve the Gaussian distribution directly. Inapproximability factors arising from properties of the Gaussian distribution have previously been obtained for other problems — such as Maximum Cut [22] — using the Unique Games Conjecture, reductions based on which naturally involve the Gaussian distribution via analytic tools such as the Invariance Principle [28].

Second, the inapproximability factors obtained in each case arise directly from a semidefinite program for the problem. Again, the optimality of semidefinite programs has been a recurring theme in UGC hardness results, while this result is among the few *NP-hardness* results that highlight this phenomenon. In addition, almost all these few known tight NP-hardness of approximation results have ratios being "nice" numbers, such as $2/3$ for Majority function on three variables and $1/2$ for Max 3CSPs [36]. In comparison, our thresholds are irrational and involve the Gaussian Distribution in a fundamental way. These properties of the thresholds suggest that it is unlikely an alternative simple combinatorial algorithm will achieve the same approximation ratio; at least to analyze such an algorithm one would have to define Gaussian Distribution and (probably) introduce the central limit theorem which is not common for combinatorial algorithms.

Third, the reductions in this work are based on a *dictatorship* test which quantitatively utilizes the Central Limit Theorem, i.e. the distribution of a sum of significant number of independent Bernoulli random variables is close to a Gaussian distribution. This is precisely the reason for the appearance of the Gaussian distribution. A key ingredient in our reductions is the smooth version of Label Cover which enables us to devise a more sophisticated *decoding* procedure which can be combined with the dictatorship test. It is pertinent to note that a couple of the (few) previous results using smooth versions of Label Cover, on hardness of learning intersection of halfspaces [25] and monomials [14], have also used analysis based on versions of the Central Limit Theorem. Our results imply that for many geometric (and possibly other combinatorial optimization) problems, using Unique Games Conjecture is not necessary and the smooth label cover (which is NP-hard) suffices in its place. Technically speaking, compared with previous work such as [14] (which use smoothness of the Label Cover to bound the fourth moments in the invariance analysis), we believe the decoding technique based on smooth Label Cover in this work is conceptually simpler and maybe useful for bypassing UGC from other problems.

2 Motivation and Related Work

2.1 L_p Subspace Approximation Problem

Algorithmically various special cases of this problem have been well studied. For $p = 2$ it reduces to the problem of determining a rank k approximation B to an $n \times m$ matrix A with respect to the Frobenius norm, which can be computed in polynomial time by using Singular Value Decomposition of A [15]. Efficient $(1+\varepsilon)$ approximations have been given various cases such as: for $p = 1$ and constant k by Feldman *et al.* [13]; $p = \infty$ and constant k by Har-Peled and Varadarajan [19]; and for general p and constant k by Shyamalkumar and Varadarajan [34] and Deshpande and Varadarajan [12]. On the other hand, the problem can be approximated to within $O(\sqrt{\log m})$ for any value of k for $p = \infty$ as shown by Varadarajan *et al.* [35] building on the work of Nemirovski *et al.* [30].

On the complexity front, Brieden, Gritzman and Klee [9] showed that the problem is NP-hard to solve optimally for $k = n-1$ and $p = \infty$. Subsequently, the problem was shown to be NP-hard to approximate within $(\log m)^{\delta}$ for $k \leqslant n - n^{\varepsilon}$ for any $0 < \varepsilon < 1$ and $p = \infty$ [35].

In more recent work, Deshpande,Tulsiani and Vishnoi [11] gave a $\sqrt{2}\gamma_p$ approximation for this problem for any k and any $p \geqslant 2$, and a γ_p approximation factor when $k = n-1$. Assuming the Unique Games Conjecture they also prove that the problem is hard approximate within a factor of $(1-\varepsilon)\gamma_p$ when $k = n-1$.

2.2 L_p Quadratic Grothendieck Maximization Problem

The special case of the problem when $p = \infty$ (maximizing over the hypercube), has been extensively studied. The problem is known to admit an $O(\log n)$ approximation [31, 30, 27, 10]. On the other hand, it was shown to be NP-hard to approximate within some constant factor in [2] and [10]. In [4], Arora *et al.* gave the best known inapproximability factor of $(\log n)^c$ for some $c > 0$ for this problem.

The L_p Quadratic Grothendieck Maximization Problem for constant p such that $2 < p < \infty$ has received attention more recently in the work of Kindler, Naor and Schechtman [26]. They exhibit an algorithm to approximate $\mathrm{Val}_p(A)$ to within a factor of $\frac{p}{e} + 30\log p$ and also show a Unique Games Conjecture [21] based inapproximability factor of $\gamma_p^2 - \varepsilon$ for all $\varepsilon > 0$. Here γ_p denotes the pth norm of a standard Gaussian variable. Note that while asymptotically (i.e. as $p \to \infty$) the upper and lower bounds both tend to $\frac{p}{e}(1+o(1))$, for a fixed constant p ($2 < p < \infty$), there remained a gap between them.

2.3 Overview of the Techniques

In the next few paragraphs we give an informal description of the techniques used in proving the results of the paper and the new ingredients employed to build upon the work of [26] and [11].

2.4 NP-Hardness Reductions

Both hardness of approximation results shown in this work are obtained via reductions from the *Smooth Label Cover* problem which is a variant of the well-known Label Cover problem. The Label cover problem, a CSP with constraints on pairs of variables, is the starting point of a majority of reductions showing hardness of approximation. In a label cover instance, the variables take values over a finite domain $[R] = \{1 \dots R\}$. Every constraint π between two variables u, v is such that the value assigned to one of the variables (say u) determines the value of the other (say v). The goal is to find an assignment satisfying the maximum number of edges. Very strong hardness of approximation results are known for the label cover problem, and these in turn are the starting point of almost all hardness of approximation results.

Unique Games is a special case of Label cover where every constraint is a bijection, i.e., the value of either variable determines the other. While constraints being bijections is convenient for showing hardness results, the difficulty of solving unique games is still conjectural.

Smooth label cover is a variant of label cover while while having properties similar to unique games, can still be shown to be NP-hard to approximate. It was first introduced in [20] for proving hardness results in hypergraph coloring and subsequently utilized for other applications in [24, 25, 14]. Roughly speaking, in a smooth label cover instance, for every constraint u, v, while u determines the value of v, the value of v almost always *disambiguates* the other. More precisely, for every vertex u, and given two or constantly many choices of labels of u, and a random neighbour of v, the value assigned to v disambiguates between the choices. Note that a unique games instance is always a smooth label cover instance.

Now we describe the reduction from Smooth Label Cover to L_p Subspace Approximation Problem. Analogous to the construction of [11], we formulate the label cover instance as a system of linear equations. More precisely, there is a coordinate for every label of every vertex of the Smooth Label Cover instance. The constraints between the labels of vertices are translated to linear equations between the corresponding variables. Let V denote the vector space of solutions to this linear system. The L_p Subspace Approximation instance we construct will lie entirely in this subspace V. This method of enforcing constraints of label cover is known as *folding* [16]. Not every vector in V corresponds to a labelling of the Smooth label cover instance. In particular, every vertex can be assigned at most one label – the vector corresponding to a labelling is necessarily *sparse*. Hence, the goal of solving Smooth Label Cover reduces to the problem of finding a *sparse* vector in this linear space V.

The instance will consist of a set of points S inside V, such that if $\langle w, x \rangle - \theta = 0$ is a subspace close in the appropriate norm (L_p norm of the Euclidean distances) to S then w is a *sparse* vector in V. To this end, for every vertex v in the label cover, we include 2^R points in V as follows: set coordinates corresponding to every other vertex to 0, set the coordinates corresponding to v randomly in $\{-1, 1\}$, and finally project the resulting vector in to V. This construction follows closely the construction in [11], except for the *folding* step where the vector is projected in to the linear space V. Folding could not be used in [11] since unique games does not have perfect completeness (optimal solution need not satisfy all constraints) and hence cannot be translated in to a system of linear equations.

The crucial property that is used in the analysis is the Berry-Essen theorem. For a real vector $v \in \mathbb{R}^R$, $\langle v, x \rangle$ over a random $x \in \{-1, 1\}^R$, is distributed as a $\{-1, 1\}$ random variable if v corresponds to a valid assignment. Hence, if v corresponds to a valid assignment, then the L_p norm of $\langle v, x \rangle$ is at most 1. On the other hand, by Berry-Essen theorem, if v is *not sparse* then the $\langle v,x\rangle$ is distributed like a Gaussian, thereby having a L_p^{th} norm of γ_p. This lies at the core of the γ_p-hardness result for L_p subspace approximation. The difficulty in the analysis is to show the converse that, every vector w, which is somewhat sparse ($\langle w,x\rangle$ not distributed as a Gaussian) corresponds to a *good* labelling of the label cover instance. Smoothness of the label cover instance plays a crucial role in showing this fact.

For the L_p Quadratic Grothendieck Problem our overall approach is similar to that of [26]: use the quadratic form to simulate a *Long Code test* on the vertices of the Smooth Label Cover (or Unique Games in [26]). As before, the coordinates are the union of Long Codes for each vertex of Smooth Label Cover instance. The quadratic form is given in terms of the Fourier coefficients of the various Long Codes. In our case however, the quadratic form differs from that of [26] in order to avoid the usage of the Cauchy-Schwartz inequality in the analysis which requires the uniqueness property of constraints afforded by an instance of Unique Games. More specifically, the quadratic form in our construction simulates the *dictatorship* and *consistency* tests on Long Codes. The dictatorship test yields a small set each of *influential* labels for a significant fraction of vertices. This is combined with the consistency test to obtain a *good* labeling to the instance of Smooth Label Cover. Our analysis crucially depends on the *smoothness* property of the constraints of the instance which roughly stated is: for any vertex, given a small set of labels, most of the constraints involving the vertex restricted to that set of labels appear structurally similar to constraints of an instance of Unique Games. Formally, the projection constraints $\pi : [M] \to [N]$ incident on any given vertex v of a Smooth Label Cover instance form a *hash family*: for any pair $a \neq b \in [N]$, one has $\pi(a) \neq \pi(b)$ with high probability for a random π incident on v. This "local uniqueness" property of the Label Cover constraints enables us to design a two step decoding procedure which extracts the aforementioned good labeling to the instance.

2.5 Approximating the L_p Quadratic Grothendieck Maximization Problem

Our algorithm is essentially a simplification of the techniques in [26]. We define the following convex relaxation for $\mathrm{Val}_p(A)$:

$$\mathrm{Vec}_p(A) = \max\left\{\sum_{i,j=1}^{n} a_{ij}\langle v_i, v_j\rangle \ : \ \{v_1,\ldots,v_n\} \subseteq L_2,\ \sum_{i=1}^{n} \|v_i\|_2^p \leqslant 1\right\}. \tag{3}$$

As observed in [26] the above convex program can be solved in polynomial time to arbitrary small precision. We directly show that $\mathrm{Vec}_p(A)$ is a γ_p^2 approximation to $\mathrm{Val}_p(A)$. This can be easily derived from the following fact: there exist mean zero Gaussian random variables h_i for $i = 1,\ldots,n$ such that $\mathbb{E}[h_ih_j] = \langle v_i, v_j\rangle$. Writing $\mathrm{Vec}_p(A)$ as $\mathbb{E}\left[\sum_{i,j=1}^n a_{ij}h_ih_j\right]$ and normalizing each variable by $\left(\sum_{k=1}^n |h_k|^p\right)^{1/p}$ yields the desired approximation. The details appear in the full version included as an appendix. This differs from the proof of [26] which obtains a slightly weaker approximation via a truncation based rounding algorithm. A generalization of our approximation for convex bodies has been obtained independently by Naor and Schechtman [29] and for the unit ℓ_p ($p > 2$) ball it essentially gives the same result as ours. Our proof also yields a polynomial time rounding algorithm to compute a solution $\{x_i\}_{i=1}^n$ which approximates $\mathrm{Val}_p(A)$ to within a factor of $\gamma_p^2(1+\delta)$ for arbitrarily small $\delta > 0$.

3 Preliminaries

We begin this section by first formally defining the two problems that we study.

Definition 3.1. *The L_p* Subspace Approximation Problem*, which we denote by* Subspace(k,p) *where k is a parameter (possibly depending on n) is: given a set of points $S = \{a_1, \ldots, a_m\} \subseteq \mathbb{R}^n$, to compute the following quantity,*

$$\mathrm{Sub}_p(S,k) = \min_{H \subseteq \mathbb{R}^n : \dim(H) = k} \left(\sum_{i=1}^{m} \mathrm{dist}(H, a_i)^p \right)^{1/p}, \tag{4}$$

where the minimum is taken over all k-dimensional subspaces of $\mathbb{R}^n$ and $\mathrm{dist}(H,a)$ is the minimum Euclidean distance between a and any point in H.

Definition 3.2. *The L_p* Quadratic Grothendieck Maximization Problem *which we denote as* QM(p) *for $1 \leqslant p < \infty$ is: given a symmetric matrix $A \in \mathbb{R}^{n \times n}$ with diagonal entries all zero, to compute the following quantity,*

$$\mathrm{Val}_p(A) := \max \left\{ \sum_{i,j=1}^{n} a_{ij} x_i x_j \;\middle|\; \sum_{i=1}^{n} |x_i|^p \leqslant 1 \right\}. \tag{5}$$

We will denote by γ_p the pth norm of a standard Gaussian random variable. Formally, for any $p \geqslant 0$, $\gamma_p := (\mathbb{E}[|g|^p])^{1/p}$, where g is a Gaussian random variable with mean 0 and variance 1. The analysis of the dictatorship tests in our reductions, requires lower bounds on the norms of sums of independent Bernoulli variables. The following lemma, proved in [26] (as Lemma 2.5) gives us the required bound.

Lemma 3.3. *Let $X_1, \ldots, X_n$ be independent Bernoulli random variables such that $\mathbb{E}[X_i] = 0$ for all $1 \leqslant i \leqslant n$ and $\sum_{j=1}^n \mathbb{E}[X_j^2] = 1$. Assume that for some $\tau \in (0, e^{-4})$, we have $\sum_{j=1}^n \mathbb{E}[|X_j|^3] \leqslant \tau$. Then for every $p \geqslant 1$,*

$$\left(\mathbb{E}\left[\Big| \sum_{j=1}^{n} X_j \Big|^p \right] \right)^{1/p} \geqslant \gamma_p \cdot \left(1 - 4\tau (\log(1/\tau))^{p/2} \right).$$

3.1 Smooth Label Cover

Our reductions require a special variant of the usual Label Cover problem, which is formally defined as follows.

Definition 3.4. *An instance of* Smooth Label Cover $\mathscr{L}(G(V,E), N, M, \{\pi^{e,v} | e \in E, v \in e\})$ *consists of a regular connected (undirected) graph $G(V,E)$ with vertex set V and edge set E. Every edge $e = (v_1, v_2)$ is associated with projection functions $\{\pi^{e,v_i}\}_{i=1}^2$ where $\pi^{e,v_i} : [M] \to [N]$. A vertex labeling is a mapping defined on $L : V \to [M]$. A labeling L satisfies edge $e = (v_1, v_2)$ if $\pi^{e,v_1}(L(v_1)) = \pi^{e,v_2}(L(v_2))$. The goal is to find a labeling which satisfies the maximum number of edges.*

The following theorem states the hardness of approximation for the Smooth Label Cover problem and also describes the various structural properties, including *smoothness*, that are satisfied by the hard instances. A proof of the theorem is included in Appendix A.

Theorem 3.5. *There exists a constant $c_0 > 0$ such that for any constant integer parameters $J, R \geqslant 1$, it is* NP*-hard to distinguish between the following two cases for a* Smooth Label Cover *instance $\mathscr{L}(G(V,E), N, M, \{\pi^{e,v} | e \in E, v \in e\})$ with $M = 7^{(J+1)R}$ and $N = 2^R 7^{JR}$:*

- *(YES Case). There is a labeling that satisfies every edge.*
- *(NO Case). Every labeling satisfies less than a fraction $2^{-c_0 R}$ of the edges.*

In addition, the instance $\mathscr{L}$ satisfies the following properties:

- *(Smoothness) For any vertex $w \in V$, $\forall i, j \in [M], i \neq j$,* $\mathbf{Pr}_{e\sim w}[\pi^{e,w}(i) = \pi^{e,w}(j)] \leqslant 1/J$, *where the probability is over a randomly chosen edge incident on w.*
- *The degree of the (regular) graph G, which we denote by d is a constant depending only on R and J.*
- *For any vertex v, edge e incident on v, and any element $i \in [N]$, we have $|(\pi^{e,v})^{-1}(i)| \leqslant t := 4^R$; i.e., there are at most $t = 4^R$ elements in $[M]$ that are mapped to the same element in $[N]$.*
- *(Weak Expansion)For any $\delta > 0$, let $V' \subseteq V$ and $|V'| = \delta \cdot |V|$, then the number of edges among the vertices in $|V'|$ is at least $(\delta^2/2)|E|$.*

Note on notation. In the following sections, the parameter n need not denote the size of the instances and its definition will be made clear at the beginning of each section.

4 Hardness reduction for $\mathrm{Subspace}(\dim - 1, p)$

In this section, we describe the NP-hardness reduction from Smooth Label Cover to the L_p Subspace Approximation Problem for a fixed $p > 2$. Specifically, we will show the following.

Theorem 4.1. *For any fixed $p > 2$ and $\varepsilon > 0$, there is polynomial time reduction from an instance $\mathscr{L}$ of* Smooth Label Cover *with appropriately chosen parameters J and R to a set of points $S \subseteq \mathbb{R}^n$ as an instance of* $\mathrm{Subspace}(n-1, p)$ *such that,*

- *(Completeness) If $\mathscr{L}$ is a YES instance, then* $\mathrm{Sub}_p(n-1, S) = 1$.
- *(Soundness) If $\mathscr{L}$ is a NO instance, then* $\mathrm{Sub}_p(n-1, S) \geqslant \gamma_p(1-\varepsilon)$.

The above implies that it is NP*-hard to approximate* $\mathrm{Subspace}(n-1, p)$ *within a factor of $(1-\varepsilon)\gamma_p$ for all $\varepsilon > 0$.*

Let the Smooth Label Cover instance be $\mathscr{L}(G(V,E), N, M, \{\pi^{e,v} | e \in E, v \in e\})$. We choose the parameters J and R as part of the analysis in Section 4.4. For convenience let $n := |V|$. Note that n does not correspond to the dimension of the point set constructed in the reduction. We do not explicitly calculate the dimension, but use the notation dim to denote it. The set of points constructed is an instance of $\mathrm{Subspace}(\dim - 1, p)$. The Euclidean distance of a point from a dimension $(\dim - 1)$ subspace, i.e., a hyperplane through the origin, is the same as the magnitude of the dot product of (the vector defining) that point with the unit normal vector to the subspace. Therefore the problem $\mathrm{Subspace}(\dim - 1, p)$ is the same as computing a unit vector which minimizes the sum of the pth powers of the dot products of the given points with the vector. Our reduction will follow this latter formulation, with the goal being to compute such a unit normal vector. The reduction proceeds in two steps: the first step yields a preliminary instance consisting of a set of points and the second step applies a *folding* operation to generate the final instance.

For notational convenience, in this section we will represent vectors with boldface characters.

4.1 Step 1: Preliminary Instance $\mathscr{A}_{prel}$

We begin by constructing the set of coordinates over which the instance is defined. For any vertex $v \in V$, let $\mathscr{P}_v$ be the set of coordinates $\{(v,i) \mid i \in [M]\}$, and $\mathscr{P} = \cup_{v\in V} \mathscr{P}_v$. In other words, $\mathscr{P}$ contains a coordinate for every label of every vertex. The instance $\mathscr{A}_{prel}$ will be over the space $\mathbb{R}^{\mathscr{P}}$ consisting of points constructed as follows.

For every vertex $v \in V$, let the set X^v be the set of all points in $\mathbb{R}^{\mathscr{P}}$ which are zero in the coordinates *not* corresponding to v i.e. $\mathscr{P} \setminus \mathscr{P}_v$ and take the values $\{-1,1\}$ in the M coordinates $\mathscr{P}_v$ corresponding to v. More formally,

$$X^v := \{\mathbf{x} \in \mathbb{R}^{\mathscr{P}} \mid \forall i \in [M],\ \mathbf{x}(v',i) \in \{-1,1\} \text{ if } v' = v \text{ and } 0 \text{ otherwise}\}.$$

The instance $\mathscr{A}_{prel}$ consists of the point set $X := \cup_{v \in V} X^v$.

Consider a vector $\mathbf{b} \in \mathbb{R}^{\mathscr{P}}$. For any vertex $v \in V$, define $\mathbf{b}_v$ to be the vector which is same as $\mathbf{b}$ in the M coordinates $\mathscr{P}_v$ and zero in rest of the coordinates. It is easy to see that for any $v \in V$,

$$\mathbb{E}_{\mathbf{x}\in\{-1,1\}^{\mathscr{P}_v}} \left[\langle \mathbf{b}_v, \mathbf{x}\rangle^2\right] = \|\mathbf{b}_v\|_2^2 = \sum_{j=1}^{M} \mathbf{b}(v,j)^2. \tag{6}$$

Using the above definition, given $\mathscr{A}_{prel}$ as an instance, the problem of $\text{Subspace}(\dim -1, k)$ is equivalent to computing a a unit normal vector $\mathbf{b}$ that minimizes $\mathbb{E}_{v\in V}\left[\mathbb{E}_{\mathbf{x}\in\{-1,1\}^{\mathscr{P}_v}}\left[|\langle \mathbf{b}_v, \mathbf{x}\rangle|^p\right]\right]$. More formally, $\mathscr{A}_{prel}$ as an instance of $\text{Subspace}(\dim -1, k)$ is equivalent to the following optimization problem:

$$\begin{aligned} & \min \left(\mathbb{E}_{v\in V}\left[\mathbb{E}_{\mathbf{x}\in\{-1,1\}^{\mathscr{P}_v}}\left[|\langle \mathbf{b}_v, \mathbf{x}\rangle|^p\right]\right]\right)^{1/p} \\ \text{subject to,} \quad & \mathbb{E}_{v\in V}\left[\|\mathbf{b}_v\|_2^2\right] = 1. \end{aligned}$$

In the next step, we use folding to implicitly induce additional constraints on the structure of the vector $\mathbf{b}$, which incorporates the projection constraints of the edges and enables a good solution $\mathbf{b}$ to be decoded into a good labeling of the Smooth Label Cover instance $\mathscr{L}$.

4.2 Step 2: Folding and Final Instance $\mathscr{A}_{final}$

For any edge $e = (u,v)$ and element $j \in [N]$, define the vector $\mathbf{h}_j^e$ as follows,

$$\mathbf{h}_j^e(w,i) = \begin{cases} 1 & \text{if } w = u \text{ and } i \in (\pi^{e,u})^{-1}(j) \\ -1 & \text{if } w = v \text{ and } i \in (\pi^{e,v})^{-1}(j) \\ 0 & \text{otherwise.} \end{cases}$$

The above implies that for any vector $\mathbf{b} \in \mathbb{R}^{\mathscr{P}}$,

$$\forall e = (u,v) \in E,\ j \in [N], \quad \mathbf{b} \perp \mathbf{h}_j^e \Leftrightarrow \sum_{i \in (\pi^{e,u})^{-1}(j)} \mathbf{b}(u,i) = \sum_{i' \in (\pi^{e,v})^{-1}(j)} \mathbf{b}(v,i') \tag{7}$$

We now define the subspace H of $\mathbb{R}^{\mathscr{P}}$ as: $H := \text{span}(\{\mathbf{h}_j^e \mid e \in E,\ j \in [N]\})$. Let $\mathbb{R}^{\mathscr{P}} = F \bigoplus H$ where $F \perp H$ is a subspace of $\mathbb{R}^{\mathscr{P}}$. The point set X constructed in Step 1 is folded over H, i.e. each point in X is replaced (with multiplicity) with its orthogonal projection on F. Let the resultant set of points be $\overline{X}$, which constitutes the final instance $\mathscr{A}_{final}$. The point set as well as the expected solution, say $\mathbf{b}$, are written in some orthonormal basis for F. Let $\overline{\mathbf{x}} \in \overline{X}$ be the orthogonal projection of a point $x \in X$ onto the subspace F, and let $\mathbf{b} \in F$ be the expected solution. Clearly we have $\langle \mathbf{b}, \overline{\mathbf{x}}\rangle = \langle \mathbf{b}, \mathbf{x}\rangle$. Also, since $\mathbf{b} \perp H$, we have from Equation

(7),

$$\forall e=(u,v)\in E,\ j\in[N],\quad \sum_{i\in(\pi^{e,u})^{-1}(j)} \mathbf{b}(u,i)=\sum_{i'\in(\pi^{e,v})^{-1}(j)} \mathbf{b}(v,i'). \tag{8}$$

We note that the objective value, which can be written as $\mathbb{E}_{\mathbf{x}\in X}[|\langle \mathbf{b},\mathbf{x}\rangle|^p]$, is unchanged under transformation of orthonormal basis since it is a function of inner product of vectors. The folding operation only ensures that the constraints given by Equation (8) are satisfied. Therefore, the instance $\mathscr{A}_{final}$ of Subspace$(\dim-1,p)$ is equivalent to the following optimization problem over solutions $\mathbf{b}\in\mathbb{R}^{\mathscr{P}}$:

$$\min\left(\mathbb{E}_{v\in V}\left[\mathbb{E}_{\mathbf{x}\in\{-1,1\}^{\mathscr{P}_v}}\left[|\langle \mathbf{b}_v,\mathbf{x}\rangle|^p\right]\right]\right)^{1/p} \tag{9}$$

subject to ,

$$\mathbb{E}_{v\in V}\left[\|\mathbf{b}_v\|_2^2\right]=1 \quad \text{and,} \tag{10}$$

$$\textstyle\sum_{i\in(\pi^{e,u})^{-1}(j)} \mathbf{b}(u,i)=\sum_{i'\in(\pi^{e,v})^{-1}(j)} \mathbf{b}(v,i'), \qquad \forall e=(u,v)\in E,\ j\in[N]. \tag{11}$$

Note that the last condition is equivalent to $\mathbf{b}\perp H$.

4.3 Completeness

If the instance $\mathscr{L}$ of Smooth Label Cover is a YES instance then there is a labeling L of the vertices of $\mathscr{L}$ that satisfies all the edges. Using this we construct a solution $\mathbf{b}^*$ to the instance $\mathscr{A}_{final}$ as follows: for any vertex $v\in V$ and element $i\in M$, $\mathbf{b}^*(v,i)$ is 1 if $L(v)=i$ and 0 otherwise.

Since L satisfies all edges, $\pi^{e,u}(L(u))=\pi^{e,v}(L(v))$ for all edges $e=(u,v)$. Therefore it is easy to see that $\mathbf{b}\perp H$. Moreover, since there is exactly one nonzero coordinate corresponding to each vertex on which $\mathbf{b}^*$ is 1, we have $\|\mathbf{b}_v^*\|_2=1$, for all $v\in V$. Therefore, $\mathbf{b}^*$ is a valid solution for $\mathscr{A}_{final}$ with objective value 1.

4.4 Soundness

We assume, towards contradiction, that $\mathbf{b}\in\mathbb{R}^{\mathscr{P}}$ is a solution to the instance $\mathscr{A}_{final}$ such that,

$$\mathbb{E}_{v\in V}\left[\mathbb{E}_{\mathbf{x}\in\{-1,1\}^{\mathscr{P}_v}}\left[|\langle \mathbf{b}_v,\mathbf{x}\rangle|^p\right]\right]\leqslant \gamma_p^p(1-\eta), \tag{12}$$

where $\eta>0$ is a positive constant. We begin with a lemma upper bounding the ℓ_2^2 mass of blocks of coordinates in $\mathbf{b}$ corresponding to small sets of vertices. This critically depends on the fact that $p>2$.

Lemma 4.2. *Let $S\subseteq V$ be a set of size $\theta|V|=\theta n$ for some $0<\theta<1$. Then, $\sum_{v\in S}\|\mathbf{b}_v\|_2^2\leqslant\gamma_p^2\theta^{1-2/p}n$.*

Proof. We need to upper bound β where,

$$\sum_{v\in S}\|\mathbf{b}_v\|_2^2=\beta n.$$

Note that the above implies that,

$$\mathbb{E}_{v\in S}\left[\|\mathbf{b}_v\|_2^2\right]=\frac{\beta}{\theta}. \tag{13}$$

We know from our assumption that $\mathbb{E}_{v\in V}\left[\mathbb{E}_{\mathbf{x}\in\{-1,1\}^{\mathscr{P}_v}}\left[|\langle \mathbf{b}_v,\mathbf{x}\rangle|^p\right]\right] \leqslant \gamma_p^p$. This implies,

$$\begin{array}{rcll} \gamma_p^p & \geqslant & \mathbb{E}_{v\in V}\left[\left[\left(\mathbb{E}_{\mathbf{x}\in\{-1,1\}^{\mathscr{P}_v}}\left[|\langle \mathbf{b}_v,\mathbf{x}\rangle|^p\right]\right)^{1/p}\right]^p\right] & \\ & \geqslant & \mathbb{E}_{v\in V}\left[\left[\left(\mathbb{E}_{\mathbf{x}\in\{-1,1\}^{\mathscr{P}_v}}\left[|\langle \mathbf{b}_v,\mathbf{x}\rangle|^2\right]\right)^{1/2}\right]^p\right] & (\text{Since } p>2) \\ & = & \mathbb{E}_{v\in V}\left[\|\mathbf{b}_v\|_2^p\right] & (\text{By Equation (6)}) \\ & \geqslant & \theta\mathbb{E}_{v\in S}\left[\|\mathbf{b}_v\|_2^p\right] & (\text{by averaging}) \\ & \geqslant & \theta\left(\mathbb{E}_{v\in S}\left[\|\mathbf{b}_v\|_2^2\right]\right)^{\frac{p}{2}} & (\text{by Jensen's Inequality}) \\ & = & \theta\left(\frac{\beta}{\theta}\right)^{\frac{p}{2}} & (\text{by Equation (13)}). \end{array}$$

Therefore, $\beta \leqslant \gamma_p^2\theta^{1-2/p}$ which completes the proof of the lemma. □

We next introduce the notion, of an *irregular* vertex : v is said to be irregular if there is a coordinate (v,i) for some $i\in[M]$ such that the value $|\mathbf{b}(v,i)|$ is large as compared to $\|\mathbf{b}_v\|_2$. Formally we have the following definition.

Definition 4.3. *(τ-irregular vertex) A vertex $v\in V$ is said to be τ-irregular if there exists $i\in[M]$ such that $|\mathbf{b}(v,i)|>\tau\|\mathbf{b}_v\|_2$. If not, the vertex is referred to as τ-regular.*

The following lemma follows from Lemma 2.5 of [26] in an analogous manner to Lemma 5.3. We therefore omit the proof.

Lemma 4.4. *For an appropriately small choice of $\tau>0$ depending on p the following holds. If $v\in V$ τ-regular then,*

$$\mathbb{E}_{\mathbf{x}\in\{-1,1\}^{\mathscr{P}_v}}\left[|\langle \mathbf{b}_v,\mathbf{x}\rangle|^p\right] \;\geqslant\; \gamma_p^p\|\mathbf{b}_v\|_2^p\left(1-\sqrt{\tau}\right). \tag{14}$$

The next lemma shows that for small enough τ, there is a significant fraction of vertices that are τ-irregular.

Lemma 4.5. *Let S_{irr} be the set of vertices that are τ-irregular and let $|S_{irr}|=\theta n$. Then for a small enough choice of $\tau=\tau(\eta,p)>0$ (in terms of η and p), $\theta=\theta(\eta,p)>0$ is a constant depending only on η and p.*

Proof. By Lemma 4.2 we have,

$$\begin{array}{rrcl} & \sum_{v\in S_{irr}}\|\mathbf{b}_v\|_2^2 & \leqslant & \gamma_p^2\theta^{1-2/p}n \\ \Leftrightarrow & \sum_{v\in V\setminus S_{irr}}\|\mathbf{b}_v\|_2^2 & \geqslant & n\left(1-\gamma_p^2\theta^{1-2/p}\right) \end{array} \tag{15}$$

Also, from our initial assumption on the objective value of $\mathbf{b}$ given by Equation (12) we have,

$$\begin{aligned}
\gamma_p^p(1-\eta)n &\geqslant \sum_{v\in V} \mathbb{E}_{\mathbf{x}\in\{-1,1\}^{\mathscr{P}_v}}\left[|\langle \mathbf{b}_v,\mathbf{x}\rangle|^p\right] \\
&\geqslant \sum_{v\in V\setminus S_{irr}} \mathbb{E}_{\mathbf{x}\in\{-1,1\}^{\mathscr{P}_v}}\left[|\langle \mathbf{b}_v,\mathbf{x}\rangle|^p\right] \\
&\geqslant \gamma_p^p(1-\sqrt{\tau})n(1-\theta)\left(\frac{1}{(1-\theta)n}\sum_{v\in V\setminus S_{irr}}\|\mathbf{b}_v\|_2^p\right) \quad \text{(by Lemma 4.4)} \\
&\geqslant \gamma_p^p(1-\sqrt{\tau})n(1-\theta)\mathbb{E}_{v\in V\setminus S_{irr}}\left[\|\mathbf{b}_v\|_2^p\right] \\
&\geqslant \gamma_p^p(1-\sqrt{\tau})n(1-\theta)\left(\mathbb{E}_{v\in V\setminus S_{irr}}\|\mathbf{b}_v\|_2^2\right)^{\frac{p}{2}} \quad \text{(by Jensen's Inequality)} \\
&\geqslant \gamma_p^p(1-\sqrt{\tau})n(1-\theta)\left(1-\gamma_p^2\theta^{1-2/p}\right)^{\frac{p}{2}} \quad \text{(by Equation (15))} \qquad (16)
\end{aligned}$$

From the above, choosing $0<\tau\leqslant\eta^6$, we obtain that $\theta>0$ is a constant depending only on η and p. Therefore, at least $\theta>0$ fraction of the vertices are τ-irregular where $0<\tau\leqslant\eta^6$ and $\theta=\theta(\eta,p)$. □

We choose an appropriately small value of $\tau=\tau(\eta,p)>0$ given by the above lemma. To complete the analysis of the soundness, in the rest of this section we show that the vector $\mathbf{b}$ can be decoded into a labeling for $\mathscr{L}$ that satisfies a significant fraction of the its edges, thereby contradicting the soundness property of Theorem 3.5.

Constructing a good labeling for $\mathscr{L}$

We now describe how to decode the vector $\mathbf{b}$ into a labeling for the set of τ-irregular vertices S_{irr}. Observe that since $|S_{irr}|\geqslant\theta n$, by the *Weak Expansion* property of Theorem 3.5,

$$|E(S_{irr})| \;\geqslant\; (\theta^2/2)|E|, \qquad (17)$$

where $E(S_{irr})$ is the set of edges induced by S_{irr}. For every vertex $v\in S_{irr}$, define,

$$\Gamma_0(v) := \left\{i\in[M] \;\middle|\; |\mathbf{b}(v,i)|\geqslant\frac{\tau}{2}\|\mathbf{b}_v\|_2\right\} \qquad \Gamma_1(v) := \left\{i\in[M] \;\middle|\; |\mathbf{b}(v,i)|\geqslant\frac{\tau}{10t}\|\mathbf{b}_v\|_2\right\},$$

where $t=4^R$ is the parameter from Theorem 3.5. Clearly, for every vertex $v\in S_{irr}$:

$$\emptyset\neq\Gamma_0(v)\subseteq\Gamma_1(v), \quad |\Gamma_0(v)|\leqslant\frac{4}{\tau^2}, \text{ and } |\Gamma_1(v)|\leqslant\frac{100t^2}{\tau^2}. \qquad (18)$$

Let v be any vertex in S_{irr}. Call an edge e incident on v to be "good" for v if $\pi^{e,v}$ maps the set $\Gamma_1(v)$ injectively (one to one) into $[N]$. Using the *smoothness* property of Theorem 3.5 yields the following bound on the probability that a random edge incident on v is "good":

$$\mathbf{Pr}_{e\ni v}[e \text{ is "good" for } v] \;\geqslant\; 1-\frac{|\Gamma_1(v)|^2}{J} \;\geqslant\; 1-\frac{10000t^4}{\tau^2 J} =: 1-\zeta. \qquad (19)$$

Since the graph of $\mathscr{L}$ is regular, this implies that the total number of edges induced by S_{irr} that are not "good" for at least one of the end points in S_{irr} is at most $2\zeta|E|$. Let $E'\subseteq E(S_{irr})$ be the set of edges induced by S_{irr} that are "good" for both endpoints. The above bounds combined with Equation (17) imply that

$|E'| \geqslant \left(\frac{\theta^2}{2} - 2\zeta\right)|E|$. The following lemma shows that the folding constraints enforce a structural property on the sets $\Gamma_0(v)$ with respect to the edges in E'.

Lemma 4.6. *Let $e = (u,v)$ be any edge in E'. Then $\pi^{e,u}(\Gamma_0(u)) \cap \pi^{e,v}(\Gamma_0(v)) \neq \emptyset$.*

Proof. Clearly, u and v are τ-irregular. Without loss of generality assume that $\|\mathbf{b}_u\|_2 \geqslant \|\mathbf{b}_v\|_2$. Since u is τ-irregular, there is a coordinate (u,i_u) ($i_u \in [M]$) such that $|\mathbf{b}(u,i_u)| \geqslant \tau\|\mathbf{b}_u\|_2$. By construction $i_u \in \Gamma_0(u)$.

Let $j_0 := \pi^{e,u}(i_u)$. Since $e \in E'$, $(\pi^{e,u})^{-1}(j_0) \cap \Gamma_1(u) = \{i_u\}$. This implies that for all $i \in (\pi^{e,u})^{-1}(j_0)$ and $i \neq i_u$, $|\mathbf{b}(u,i)| < \frac{\tau}{10t}\|\mathbf{b}_u\|_2$. Moreover, from Theorem 3.5 $|(\pi^{e,u})^{-1}(j_0)| \leqslant t$. Combining these observations yields,

$$\left|\sum_{i\in(\pi^{e,u})^{-1}(j_0)} \mathbf{b}(u,i)\right| \;\geqslant\; \left(\tau - t\left(\frac{\tau}{10t}\right)\right)\|\mathbf{b}_u\|_2 \;=\; \left(\frac{9\tau}{10}\right)\|\mathbf{b}_u\|_2. \tag{20}$$

We next show that $(\pi^{e,v})^{-1}(j_0) \cap \Gamma_0(v) \neq \emptyset$, which would imply that $j_0 \in \pi^{e,u}(\Gamma_0(u)) \cap \pi^{e,v}(\Gamma_0(v))$ thus completing the proof of the lemma.

For a contrapositive assume that $(\pi^{e,v})^{-1}(j_0) \cap \Gamma_0(v) = \emptyset$. Moreover, since $e \in E'$, $(\pi^{e,v})^{-1}(j_0) \cap \Gamma_1(v) \leqslant 1$. This yields the following bound,

$$\left|\sum_{i'\in(\pi^{e,v})^{-1}(j_0)} \mathbf{b}(v,i')\right| \;\leqslant\; \left(\frac{\tau}{2} + t\left(\frac{\tau}{10t}\right)\right)\|\mathbf{b}_v\|_2 \;=\; 0.6\tau\|\mathbf{b}_v\|_2. \tag{21}$$

However, the folding constraints (Equation (8)) imply that,

$$\sum_{i\in(\pi^{e,u})^{-1}(j_0)} \mathbf{b}(u,i) \;=\; \sum_{i'\in(\pi^{e,v})^{-1}(j_0)} \mathbf{b}(v,i'),$$

which is a contradiction to Equations (20) and (21) combined with $\|\mathbf{b}_u\|_2 \geqslant \|\mathbf{b}_v\|_2 > 0$ (by the definition of S_{irr}). This completes the proof of the lemma. ☐

Let L^* be a labeling to the vertices in S_{irr} constructed by independently and uniformly at random choosing a label from the set $\Gamma_0(v)$ for every vertex $v \in S_{irr}$. By Lemma 4.6, every edge $e = (u,v) \in E'$ is satisfied with probability at least $\frac{1}{|\Gamma_0(u)||\Gamma_0(v)|} \geqslant \frac{\tau^4}{16}$ (by Equation (18)). Therefore, in expectation the total fraction Δ of edges satisfied is bounded by, $\Delta \geqslant \left(\frac{\tau^4}{16}\right)\left(\frac{\theta^2}{2} - 2\zeta\right)$. Choosing $J > (4^R)^5$ and $R \gg 1$ large enough (depending on η) so that $\zeta \ll \theta$ one can ensure that $\Delta > 2^{-c_0 R}$ thereby yielding a contradiction to the soundness of Theorem 3.5.

5 Hardness Reduction for QM(p)

In this section we shall describe the NP-hardness reduction from Smooth Label Cover to the L_p Quadratic Grothendieck Maximization Problem QM(p). Specifically we will show the following:

Theorem 5.1. *For any fixed $p > 2$ and constant $\varepsilon > 0$, there is a polynomial time reduction from an instance $\mathscr{L}$ of* Smooth Label Cover *with appropriately chosen parameters J and R to an instance A of* QM(p) *such that,*

- *(Completeness) If $\mathscr{L}$ is a YES instance, then* $\mathrm{Val}_p(A) = 1$.
- *(Soundness) If $\mathscr{L}$ is a NO instance, then* $\mathrm{Val}_p(A) \leqslant \gamma_p^{-2}(1+\varepsilon)$.

The above implies that it is NP*-hard to approximate* QM(p) *within a factor of* $(1-\varepsilon)\gamma_p^2$ *for all* $\varepsilon > 0$.

The instance of QM(p) in our reduction is not explicitly given as a matrix. Instead we construct a set of coordinates and a quadratic form over any mapping from the set of coordinates to real numbers. A solution to this instance of QM(p) would be a mapping that maximizes the value of the quadratic form subject to the appropriate bound on the pth norm of the mapping. While the initial construction would not ensure that the diagonal terms of the quadratic form are all 0, our analysis shall prove that setting them to zero would not change the optimum of the instance significantly.

We start with an instance of Smooth Label Cover $\mathscr{L}(G(V,E),N,M,\{\pi^{e,v}|e\in E, v\in e\})$ as given in Theorem 3.5. The parameters J and R shall be chosen appropriately for the soundness analysis of the reduction in Section 5.2. Let $n := |V|$. Define the parameters B and D as follows:

$$D := d\cdot n^{12}\cdot |E|^2\cdot 2^M\cdot M \quad \text{and} \quad B := n^{10}\cdot |E|^2\cdot 2^M, \tag{22}$$

where d is the degree of the (regular) graph G of the instance $\mathscr{L}$ as given in Theorem 3.5. The first step of our construction is to define the coordinates.

Coordinates. For each vertex $v\in V$ there are D sets of coordinates C_v^j for $j = 1,\ldots,D$. Each set C_v^j consists of 2^M coordinates indexed by all elements of $\{-1,1\}^M$, and we denote the coordinate corresponding to $x\in\{-1,1\}^M$ by $C_v^j(x)$. Let,

$$\mathscr{C} := \bigcup_{v\in V}\bigcup_{j\in D}\bigcup_{x\in\{-1,1\}^M}\{C_v^j(x)\}$$

denote the set of all coordinates.

Let $F:\mathscr{C}\mapsto\mathbb{R}$ be a mapping from the set of coordinates to real numbers. The quadratic form we construct shall be defined over F. Before we do so, we need to define some additional quantities. Given F we define $f_v^j, f_v : \{-1,1\}^M\mapsto\mathbb{R}$ for all $v\in V$ and $j\in[D]$ by setting,

$$f_v^j(x) = F(C_v^j(x)) \quad \text{and,} \tag{23}$$

$$f_v(x) = \mathbb{E}_{j\in[D]}[f_v^j(x)] \quad \forall x\in\{-1,1\}^M. \tag{24}$$

In other words, f_v is a point-wise average of f_v^j over all $j\in[D]$. The L_q norm of F for $q\geqslant 1$ is given by:

$$\|F\|_q := \left(\mathbb{E}_{v\in V}\mathbb{E}_{j\in[D]}\left[\|f_v^j\|_q^q\right]\right)^{1/q}, \tag{25}$$

where $\|f_v^j\|_q = \left(\mathbb{E}_{x\in\{-1,1\}^M}|f_v^j(x)|^q\right)^{1/q}$. Now, since $\|f\|_q^q$ is a convex function for $q\geqslant 1$, we have by Jensen's inequality,

$$\|F\|_q \geqslant \left(\mathbb{E}_{v\in V}\|f_v\|_q^q\right)^{1/q}. \tag{26}$$

We note that the functions f_v $(v\in V)$ can be written in their Fourier expansion with the basis functions χ_S $(S\subseteq[M])$ with Fourier coefficients $\widehat{f_v}(S)$. Note that the Fourier coefficients are linear forms on the values of the function f_v. In our construction the quadratic form for F shall be defined as a quadratic form over the Fourier coefficients. For convenience, we shall abuse notation to denote the Fourier coefficients corresponding to singleton sets $\{i\}$ $(i\in[M])$ by $\widehat{f_v}(i)$.

The Quadratic Form. We define quadratic forms on F : $A_{cons}(F)$, $A_{dict}(F)$ and $A_{prel}(F)$ in terms of the

Fourier coefficients of the functions f_v ($v \in V$) as follows.

$$
\begin{aligned}
A_{cons}(F) \quad := \quad & -B\mathbb{E}_{e=(u,w)}\left[\sum_{j\in[N]}\left(\sum_{i\in(\pi^{e,u})^{-1}(j)}\widehat{f}_u(i) - \sum_{i'\in(\pi^{e,w})^{-1}(j)}\widehat{f}_w(i')\right)^2\right] \\
& -B\mathbb{E}_{v\in V}\left[\sum_{\substack{S\subseteq[M]\\|S|\neq 1}}\widehat{f}_v(S)^2\right], \qquad (27)
\end{aligned}
$$

$$
A_{dict}(F) = \mathbb{E}_{v\in V}\left[\sum_{i\in[M]}\widehat{f}_u(i)^2\right], \tag{28}
$$

and,

$$
A_{prel}(F) = A_{cons}(F) + A_{dict}(F). \tag{29}
$$

In our reduction F denotes a solution to our instance of QM(p) over which the quadratic form is defined. Therefore, F satisfies the bound on its p-norm: $\|F\|_p = 1$.

We shall prove the completeness and soundness claims for the quadratic form $A_{prel}(F)$. However, we note that the $A_{prel}(F)$ may have non-zero diagonal terms. We address this issue in Section 5.3 wherein we show the existence of another quadratic form $A_{fin}(F)$ such that $|A_{fin}(F) - A_{prel}(F)| \leqslant 1/n$ for all F such that $\|F\|_p \leqslant 1$, which then suffices to prove Theorem 5.1.

Before we proceed, we state the instance of QM(p) as an optimization problem. The objective is to compute:

$$
\max_{F:\mathscr{C}\mapsto\mathbb{R}} A_{prel}(F) \tag{30}
$$

$$
\text{s.t.} \qquad \|F\|_p = 1. \tag{31}
$$

5.1 Completeness

Suppose the Smooth Label Cover instance $\mathscr{L}$ has a labeling $\sigma : V \mapsto [M]$ that satisfies all edges. Then construct the vector F by defining f_v^j, for all $v \in V, j \in [D]$, as follows:

$$
f_v^j(x) = x(\sigma(v)), \quad \forall x \in \{-1,1\}^M.
$$

The above also implies that $f_v(x) = x(\sigma(v))$ for all $x \in \{-1,1\}^M$, i.e. f_v is the 'dictator' function given by the $\sigma(v)$-th coordinate. Therefore, we obtain $\widehat{f_v}(\sigma(v)) = 1$ and $\widehat{f_v}(S) = 0$ for all $S \neq \{\sigma(v)\}$. Clearly, $\|F\|_p = 1$ since F is either 1 or -1 at any coordinate.

To analyze the value of $A_{prel}(F)$ we observe that $\pi_{e,u}(\sigma(u)) = \pi_{e,w}(\sigma(w))$ for all edges $e = (u,w)$, which along with the fact that $f_u(x) = x(\sigma(u))$ and $f_w(x) = x(\sigma(w))$ implies that $A_{cons}(F) = 0$. Also, $A_{dict}(F) = 1$, which gives us that $A_{prel}(F) = 1$.

5.2 Soundness

For a contradiction we assume that there is a vector F such that $A_{prel}(F) \geqslant \gamma_p^{-2}(1+\eta)$ for some constant $\eta > 0$. We shall show that this implies the existence of a labeling to $\mathscr{L}$ that satisfies a significant fraction of edges depending only on η.

The following is a straightforward upper bound on $A_{dict}(F)$:

$$\begin{aligned}
A_{dict}(F) &= \mathbb{E}_{v\in V}\left[\sum_{i\in[M]} \widehat{f}_v(i)^2\right] \\
&\leqslant \mathbb{E}_{v\in V}\left[\|f_v\|_2^2\right] \\
&\leqslant \left(\mathbb{E}_{v\in V}\left[\|f_v\|_2^p\right]\right)^{2/p} \quad \text{(By Jensen's Inequality)} \\
&\leqslant \left(\mathbb{E}_{v\in V}\left[\|f_v\|_p^p\right]\right)^{2/p} \quad (\text{since } \|f\|_p \geqslant \|f\|_2) \\
&\leqslant \|F\|_p^2 \leqslant 1. \qquad (32)
\end{aligned}$$

For every $v \in V$ define: $a_1^v := \sum_{|S|\neq 1, S\subseteq[M]} |\widehat{f}_v(S)|$ and $a_2^v := \left(\sum_{|S|\neq 1, S\subseteq[M]} |\widehat{f}_v(S)|^2\right)^{1/2}$. Clearly, $a_1^v \leqslant 2^{M/2} a_2^v$. Moreover, since $A_{prel}(F) \geqslant 0$, Equations (27), (29) and (32) along with the Cauchy-Schwarz inequality yield the following:

$$\begin{aligned}
\frac{1}{B} &\geqslant \mathbb{E}_{v\in V}\left[(a_2^v)^2\right] \\
&\geqslant \left(\mathbb{E}_{v\in V}\left[a_2^v\right]\right)^2 \\
&\geqslant \left(\mathbb{E}_{v\in V}\left[2^{-M/2} a_1^v\right]\right)^2 \\
\Rightarrow \left(\frac{2^M}{B}\right)^{1/2} &\geqslant \mathbb{E}_{v\in V}\left[a_1^v\right] \\
\Rightarrow \left(\frac{n^2 2^M}{B}\right)^{1/2} &\geqslant \max_{v\in V} a_1^v =: a_{\max} \qquad (33)
\end{aligned}$$

For the remainder of the analysis our focus shall be on the degree one Fourier spectrum of the functions f_v. Define, for every vertex v: $f_v^{=1} := \sum_{i\in[M]} \widehat{f}_v(i)\chi_{\{i\}}$, i.e. the function obtained by taking only the degree one Fourier spectrum of f_v. Clearly, we have $|f_v(x) - f_v^{=1}(x)| \leqslant a_{\max}$ for all $x \in \{-1,1\}^M$ and $v \in V$. By the triangle inequality for L_q, L_q' norms (for $q, q' \geqslant 1$) this implies,

$$\left| \|f_v\|_q - \|f_v^{=1}\|_q \right| \leqslant a_{\max} \quad \forall v \in V \qquad (34)$$

$$\text{and,} \quad \left| \left(\mathbb{E}_{v\in V}\left[\|f_v\|_q^{q'}\right]\right)^{1/q'} - \left(\mathbb{E}_{v\in V}\left[\|f_v^{=1}\|_q^{q'}\right]\right)^{1/q'} \right| \leqslant a_{\max}. \qquad (35)$$

By our setting of B, $a_{\max}$ is at most $1/n^4$. Since $p \geqslant 2$ is a fixed constant, setting $q = q' = p$ in Equation (35) implies (for large enough n),

$$\mathbb{E}_{v\in V}\left[\|f_v^{=1}\|_p^p\right] \leqslant \left(\|F\|_p + 1/n^4\right)^p \leqslant 1 + 1/n^3. \quad \text{(By Equations (26), (31))} \qquad (36)$$

Using Jensen's inequality we also obtain that,

$$\mathbb{E}_{v\in V}\left[\|f_v^{=1}\|_p^2\right] \leqslant \left(\mathbb{E}_{v\in V}\left[\|f_v^{=1}\|_p^p\right]\right)^{2/p} \leqslant \left(\|F\|_p + 1/n^4\right)^2 \leqslant 1 + 1/n^3. \qquad (37)$$

Using the above we obtain the following upper bound on the sum of the values $\|f_v^{=1}\|_2^2$ for small sets of vertices.

Lemma 5.2. *Let $S \subseteq V$ be a set of size $\theta|V| = \theta n$ for some $0 < \theta < 1$. Then,*

$$\sum_{v \in S} \|f_v^{=1}\|_2^2 \leqslant \theta^{1-2/p} n(1 + 1/n^3). \tag{38}$$

Proof. We need to upper bound β where,

$$\sum_{v \in S} \|f_v^{=1}\|_2^2 = \beta n.$$

Note that the above implies that,

$$\mathbb{E}_{v \in S}\left[\|f_v^{=1}\|_2^2\right] = \frac{\beta}{\theta}. \tag{39}$$

Equation (36) yields $\mathbb{E}_{v \in V}\left[\|f_v^{=1}\|_p^p\right] \leqslant 1 + 1/n^3$. This implies,

$$\begin{aligned}
1 + 1/n^3 &\geqslant \mathbb{E}_{v \in V}\left[\|f_v^{=1}\|_2^p\right] && (\text{since } \|f\|_p \geqslant \|f\|_2) \\
&\geqslant \theta \mathbb{E}_{v \in S}\left[\|f_v^{=1}\|_2^p\right] && (\text{by averaging}) \\
&\geqslant \theta \left(\mathbb{E}_{v \in S}\left[\|f_v^{=1}\|_2^2\right]\right)^{\frac{p}{2}} && (\text{by Jensen's Inequality}) \\
&= \theta \left(\frac{\beta}{\theta}\right)^{\frac{p}{2}} && (\text{by Equation (39)}).
\end{aligned}$$

Therefore, $\beta \leqslant \theta^{1-2/p}(1 + 1/n^3)^{2/p} \leqslant \theta^{1-2/p}(1 + 1/n^3)$, (since $p \geqslant 2$) which completes the proof of the lemma. ☐

Before proceeding we choose a parameter $\tau > 0$ which we shall later fix appropriately to depend only on η and p. We now define the following set $V' \subset V$ of vertices which have significantly large "mass" as follows.

$$V' := \{v \in V \mid \|f_v^{=1}\|_2^2 \geqslant 1/n^3\} \tag{40}$$

Further, define a subset $S_{irr} \subseteq V'$ as:

$$S_{irr} := \{v \in V' \mid \exists i \in [M] \text{ s.t. } |\widehat{f_v^{=1}}(i)| > \tau \|f_v^{=1}\|_2\} \tag{41}$$

We shall refer to S_{irr} as the set of τ-*irregular* vertices. Reusing notation for convenience, we assume that $|S_{irr}| = \theta n$. Our goal is to show that θ is a significantly large constant depending on η and p (for an appropriate choice of τ, again depending on η and p). Lemma 5.2 applied to S_{irr} directly gives the following:

$$\sum_{v \in S_{irr}} \|f_v^{=1}\|_2^2 \leqslant \theta^{1-2/p} n(1 + 1/n^3). \tag{42}$$

We also note that since $A_{cons}(F) \leqslant 0$ by definition, $A_{prel}(F) \leqslant A_{dict}(F) = \mathbb{E}_{v\in V}\left[\|f_v^{=1}\|_2^2\right]$. Therefore,

$$\begin{aligned}
\gamma_p^{-2}(1+\eta)n &\leqslant \sum_{v\in V}\|f_v^{=1}\|_2^2 \\
&= \sum_{v\in V'\setminus S_{irr}}\|f_v^{=1}\|_2^2 + \sum_{v\in S_{irr}}\|f_v^{=1}\|_2^2 + \sum_{v\in V\setminus V'}\|f_v^{=1}\|_2^2 \\
&\leqslant \sum_{v\in V'\setminus S_{irr}}\|f_v^{=1}\|_2^2 + \theta^{1-2/p}n(1+1/n^3) + 1/n^2,
\end{aligned}$$

where we used Equation (42) along with the bound $\sum_{v\in V\setminus V'}\|f_v^{=1}\|_2^2 \leqslant n\cdot(1/n^3) = 1/n^2$. Rearranging the above we obtain,

$$\sum_{v\in V'\setminus S_{irr}}\|f_v^{=1}\|_2^2 \geqslant \gamma_p^{-2}(1+\eta)n - \theta^{1-2/p}n(1+1/n^3) - 1/n^2. \tag{43}$$

To show that θ is large, we need to upper bound the LHS of the above equation. For this we use the following lemma which follows from Lemma 3.3.

Lemma 5.3. *For an appropriately small choice of $\tau > 0$ depending on p the following holds. For all vertices $v \in V'\setminus S_{irr}$,*

$$\|f_v^{=1}\|_p^2 \geqslant \gamma_p^2\|f_v^{=1}\|_2^2\left(1-\sqrt{\tau}\right). \tag{44}$$

Proof. We consider probability space given by the uniform distribution over $x \in \{-1,1\}^M$. Defining $X_i := x_i\left(\frac{\widehat{f}_v(i)}{\|f_v^{=1}\|_2}\right)$ we observe X_i are Bernoulli variables with $\mathbb{E}[X_i] = 0$ for all $1 \leqslant i \leqslant n$ and $\sum_{j=1}^n \mathbb{E}[X_j^2] = 1$. Moreover, since $v \in V'\setminus S_{irr}$, $\sum_{j=1}^n \mathbb{E}[|X_j|^3] \leqslant \tau$. Applying Lemma 3.3 and observing that for small enough $\tau > 0$ depending on p, $(1-4\tau(\log(1/\tau))^{p/2})^2 \geqslant (1-\sqrt{\tau})$, we obtain the desired bound. □

The above analysis yields the following sequence of inequalities which gives us a lower bound on θ depending on η.

$$\begin{aligned}
(1+1/n^3)n &\geqslant \sum_{v\in V}\|f_v^{=1}\|_p^2 \quad \text{(By Equation (37))} \\
&\geqslant \sum_{v\in V'\setminus S_{irr}}\|f_v^{=1}\|_p^2 \\
&\geqslant \gamma_p^2\left(1-\sqrt{\tau}\right)\sum_{v\in V'\setminus S_{irr}}\|f_v^{=1}\|_2^2 \quad \text{(By Lemma 5.3)} \\
&\geqslant \left(\gamma_p^2\left(1-\sqrt{\tau}\right)\right)\left(\gamma_p^{-2}(1+\eta)n - \theta^{1-2/p}n(1+1/n^3) - 1/n^2\right) \\
&= n(1-\sqrt{\tau})\left(1+\eta-\gamma_p^2\theta^{1-2/p}(1+1/n^3) - \gamma_p^2/n\right)
\end{aligned}$$

Choosing $0 < \tau \leqslant \eta^6$ in the above inequality, and using the fact that $p > 2$ is a fixed constant, for large enough n, we obtain that $\theta > 0$ is (at least) a positive constant depending on η and p. Therefore, S_{irr} contains at least a constant fraction of vertices in V. To complete the soundness analysis we shall use this to obtain a substantially good labeling to the instance $\mathscr{L}$.

Recovering a good labeling to $\mathscr{L}$

F shall be decoded into a labeling for the set of τ-irregular vertices S_{irr}. Observe that since $|S_{irr}| \geqslant \theta n$, by the *Weak Expansion* property of Theorem 3.5,

$$|E(S_{irr})| \ \geqslant \ (\theta^2/2)|E|, \tag{45}$$

where $E(S_{irr})$ is the set of edges induced by S_{irr}. For every vertex $v \in S_{irr}$, define,

$$\Gamma_0(v) := \left\{ i \in [M] \ \mid \ |\widehat{f_v^{=1}}(i)| \geqslant \frac{\tau}{2}\|f_v^{=1}\|_2 \right\}$$

and,

$$\Gamma_1(v) := \left\{ i \in [M] \ \mid \ |\widehat{f_v^{=1}}(i)| \geqslant \frac{\tau}{10t}\|f_v^{=1}\|_2 \right\},$$

where $t = 4^R$ is the parameter from Theorem 3.5. Clearly, for every vertex $v \in S_{irr}$:

$$\emptyset \neq \Gamma_0(v) \subseteq \Gamma_1(v), \quad |\Gamma_0(v)| \leqslant \frac{4}{\tau^2}, \ \text{and} \ |\Gamma_1(v)| \leqslant \frac{100t^2}{\tau^2}. \tag{46}$$

Let v be any vertex in S_{irr}. Call an edge e incident on v to be "good" for v if $\pi^{e,v}$ maps the set $\Gamma_1(v)$ injectively (one to one) into $[N]$. Using the *smoothness* property of Theorem 3.5 yields the following bound on the probability that a random edge incident on v is "good":

$$\mathbf{Pr}_{e \ni v}[e \text{ is "good" for } v] \ \geqslant \ 1 - \frac{|\Gamma_1(v)|^2}{J} \ \geqslant \ 1 - \frac{10000t^4}{\tau^4 J} =: 1 - \zeta. \tag{47}$$

Since the graph of $\mathscr{L}$ is regular, this implies that the total number of edges induced by S_{irr} that are not "good" for at least one of the end points in S_{irr} is at most $2\zeta|E|$. Let $E' \subseteq E(S_{irr})$ be the set of edges induced by S_{irr} that are "good" for both endpoints. The above bounds combined with Equation (45) imply,

$$|E'| \geqslant \left(\frac{\theta^2}{2} - 2\zeta \right) |E|. \tag{48}$$

The following lemma shows that the constraints given by Equation (27) enforce a structural property on the sets $\Gamma_0(v)$ with respect to the edges in E'.

Lemma 5.4. *Let $e = (u,v)$ be any edge in E'. Then $\pi^{e,u}(\Gamma_0(u)) \cap \pi^{e,v}(\Gamma_0(v)) \neq \emptyset$.*

Proof. Clearly, u and v are in S_{irr}. Without loss of generality assume that $\|f_u^{=1}\|_2 \geqslant \|f_v^{=1}\|_2 \geqslant 1/n^3$ where the lower bound is because S_{irr} is a subset of V'. Since u is τ-irregular, there is exists $i_u \in [M]$ such that $|\widehat{f_u^{=1}}(i_u)| \geqslant \tau\|f_u^{=1}\|_2$. By construction $i_u \in \Gamma_0(u)$.

Let $j_0 := \pi^{e,u}(i_u)$. Since $e \in E'$, $(\pi^{e,u})^{-1}(j_0) \cap \Gamma_1(u) = \{i_u\}$. This implies that for all $i \in (\pi^{e,u})^{-1}(j_0)$ and $i \neq i_u$, $|\widehat{f_u^{=1}}(i)| < \frac{\tau}{10t}\|f_u^{=1}\|_2$. Moreover, from Theorem 3.5 $|(\pi^{e,u})^{-1}(j_0)| \leqslant t$. Combining these observations yields,

$$\left| \sum_{i \in (\pi^{e,u})^{-1}(j_0)} \widehat{f_u^{=1}}(i) \right| \ \geqslant \ \left(\tau - t\left(\frac{\tau}{10t} \right) \right) \|f_u^{=1}\|_2 \ = \ \left(\frac{9\tau}{10} \right) \|f_u^{=1}\|_2. \tag{49}$$

We shall now show that $(\pi^{e,v})^{-1}(j_0) \cap \Gamma_0(v) \neq \emptyset$, which would imply that $j_0 \in \pi^{e,u}(\Gamma_0(u)) \cap \pi^{e,v}(\Gamma_0(v))$ thus completing the proof of the lemma.

For a contrapositive assume that $(\pi^{e,v})^{-1}(j_0) \cap \Gamma_0(v) = \emptyset$. Moreover, since $e \in E'$, $(\pi^{e,v})^{-1}(j_0) \cap \Gamma_1(v) \leqslant 1$. This yields the following bound,

$$\left| \sum_{i' \in (\pi^{e,v})^{-1}(j_0)} \widehat{f_v^{=1}}(i') \right| \leqslant \left(\frac{\tau}{2} + t\left(\frac{\tau}{10t} \right) \right) \|f_v^{=1}\|_2 = 0.6\tau \|f_v^{=1}\|_2 \leqslant 0.6\tau \|f_u^{=1}\|_2. \tag{50}$$

Noting that for any vertex $w \in V$ and $i \in [M]$, $\widehat{f_w^{=1}}(i) = \widehat{f_w}(i)$, the constraints given by Equation (27) imply,

$$\left| \sum_{i \in (\pi^{e,u})^{-1}(j_0)} \widehat{f_u^{=1}}(i) - \sum_{i' \in (\pi^{e,v})^{-1}(j_0)} \widehat{f_v^{=1}}(i') \right| \leqslant \frac{|E|}{\sqrt{B}} \leqslant \frac{1}{n^5},$$

by our setting of B. Since $\tau > 0$ is a constant, this is a contradiction to Equations (49) and (50) combined with $\|f_u\|_2 \geqslant 1/n^3$, for large enough n. This completes the proof of the lemma. $\square$

Let L^* be a labeling to the vertices in S_{irr} constructed by independently and uniformly at random choosing a label from the set $\Gamma_0(v)$ for every vertex $v \in S_{irr}$. By Lemma 5.4, every edge $e = (u,v) \in E'$ is satisfied with probability at least $\frac{1}{|\Gamma_0(u)||\Gamma_0(v)|} \geqslant \frac{\tau^4}{16}$ (by Equation (46)). Therefore, in expectation the total fraction Δ of edges satisfied is lower bounded by,

$$\Delta \geqslant \left(\frac{\tau^4}{16} \right) \left(\frac{\theta^2}{2} - 2\zeta \right).$$

Choosing $J > (4^R)^5$ and $u \gg 1$ large enough (depending on η) so that $\zeta \ll \theta$ one can ensure that $\Delta > 2^{-c_0 R}$ thereby yielding a contradiction to the soundness of Theorem 3.5. This completes the analysis of the soundness case.

5.3 Removing the Diagonal Terms

In this section we show that there is a quadratic form $A_{fin}(F)$ with no diagonal entries, such $|A_{fin}(F) - A_{prel}(F)| \leqslant 1/n$ for F such that $\|F\|_p = 1$. Our reduction would output a slightly scaled version of $A_{fin}(F)$ as the quadratic form to ensure that in the completeness case of Theorem 5.1 the optimum is at least 1. Since the scaling is by at most a factor of $(1 + 1/n)$ the error induced in the soundness case can be absorbed in the constant $\varepsilon > 0$ of Theorem 5.1, thus completing its proof. The rest of this section is devoted to computing $A_{fin}(F)$. We first note that there are two sources of the diagonal terms in $A_{prel}(F)$:
(i) expressions of the form $\widehat{f_v}(S)^2$ for some $v \in V$ and $S \subseteq [M]$ and,
(ii) expressions of the form $\left(\sum_{s \in (\pi^{e,v})^{-1}(r)} \widehat{f_v}(s) \right)^2$ for some $v \in V$, $e \sim v$ and $r \in [N]$.
We shall analyze each of the above separately as follows.

For any vertex $v \in V$ and set $S \subseteq [M]$ we have,

$$\begin{aligned}
\widehat{f_v}(S)^2 &= \left(\mathbb{E}_{j \in [D]} \left[\widehat{f_v^j}(S) \right] \right)^2 \\
&= \frac{1}{D^2} \cdot \sum_{j_1, j_2 \in [D]} \widehat{f_v^{j_1}}(S) \widehat{f_v^{j_2}}(S) \\
&= \frac{1}{D^2} \cdot \sum_{\substack{j_1, j_2 \in [D] \\ j_1 \neq j_2}} \widehat{f_v^{j_1}}(S) \widehat{f_v^{j_2}}(S) + \frac{1}{D^2} \cdot \sum_{j \in [D]} \widehat{f_v^j}(S)^2.
\end{aligned} \tag{51}$$

Note that the diagonal terms are a consequence of only the second term in the expression in Equation (51). Therefore, the quadratic form $\widehat{f}_v(S)^2 - T_v(S)$ has no diagonal terms where,

$$T_v(S) := \frac{1}{D^2} \cdot \sum_{j \in [D]} \widehat{f_v^j}(S)^2, \tag{52}$$

for $v \in V$ and $S \subseteq [M]$. Note that $T_v(S) \geqslant 0$. In addition we have,

$$\begin{aligned}
\mathbb{E}_{v \in V}\left[\sum_{S \subseteq [M]} T_v(S)\right] &= \mathbb{E}_{v \in V}\left[\sum_{S \subseteq [M]} \frac{1}{D^2} \cdot \sum_{j \in [D]} \widehat{f_v^j}(S)^2\right] \\
&= \mathbb{E}_{v \in V}\left[\frac{1}{D^2} \sum_{j \in [D]} \sum_{S \subseteq [M]} \widehat{f_v^j}(S)^2\right] \\
&= \frac{1}{D}\left(\mathbb{E}_{v \in V} \mathbb{E}_{j \in [D]}\left[\sum_{S \subseteq [M]} \widehat{f_v^j}(S)^2\right]\right) \\
&= \frac{1}{D}\left(\mathbb{E}_{v \in V} \mathbb{E}_{j \in [D]}\left[\|f_v^j\|_2^2\right]\right) \\
&= \frac{1}{D}\|F\|_2^2 \;\leqslant\; \frac{1}{D}\|F\|_p^2 \;=\; \frac{1}{D}.
\end{aligned} \tag{53}$$

Now consider an expression of the form $\left(\sum_{s \in (\pi^{e,v})^{-1}(r)} \widehat{f}_v(s)\right)^2$. We have,

$$\begin{aligned}
\left(\sum_{s \in (\pi^{e,v})^{-1}(r)} \widehat{f}_v(s)\right)^2 &= \left(\mathbb{E}_{j \in [D]}\left[\sum_{s \in (\pi^{e,v})^{-1}(r)} \widehat{f_v^j}(s)\right]\right)^2 \\
&= \frac{1}{D^2} \cdot \sum_{\substack{j_1, j_2 \in [D] \\ j_1 \neq j_2}} \left[\left(\sum_{s \in (\pi^{e,v})^{-1}(r)} \widehat{f_v^{j_1}}(s)\right)\left(\sum_{s \in (\pi^{e,v})^{-1}(r)} \widehat{f_v^{j_2}}(s)\right)\right] \\
&\quad + \frac{1}{D^2} \cdot \sum_{j \in [D]} \left(\sum_{s \in (\pi^{e,v})^{-1}(r)} \widehat{f_v^j}(s)\right)^2
\end{aligned} \tag{54}$$

Clearly, the contribution to the diagonal entries comes from the last term in the above sum. Therefore, the quadratic form

$$\left(\sum_{s \in (\pi^{e,v})^{-1}(r)} \widehat{f}_v(s)\right)^2 - R_{e,v,r} \quad ,$$

does not have any diagonal entries where,

$$R_{e,v,r} = \frac{1}{D^2} \cdot \sum_{j \in [D]} \left(\sum_{s \in (\pi^{e,v})^{-1}(r)} \widehat{f_v^j}(s)\right)^2 \tag{55}$$

By Cauchy-Schwartz inequality we have that,

$$\begin{aligned} R_{e,v,r} &\leqslant \frac{1}{D^2} \cdot \sum_{j \in [D]} |(\pi^{e,v})^{-1}(r)| \sum_{s \in (\pi^{e,v})^{-1}(r)} \widehat{f_v^j}(s)^2 \\ &\leqslant \frac{1}{D^2} \cdot \sum_{j \in [D]} M \sum_{s \in (\pi^{e,v})^{-1}(r)} \widehat{f_v^j}(s)^2 \end{aligned} \tag{56}$$

Since the graph of $\mathscr{L}$ is regular, we have that,

$$\begin{aligned} \mathbb{E}_{e=(u,w)\in E}\left[\sum_{r\in[N]} (R_{e,u,r} + R_{e,w,r})\right] &= 2M\mathbb{E}_{v\in V}\left[\frac{1}{D^2}\sum_{j\in[D]}\sum_{s\in[M]} \widehat{f_v^j}(s)^2\right] \\ &\leqslant \frac{2M}{D}, \end{aligned} \tag{57}$$

where the last inequality follows from Equation (53).

Again observing that the graph of $\mathscr{L}$ is regular, it can be seen that $A_{cons}(F)$ can be written as,

$$\begin{aligned} A_{cons}(F) &= -B\cdot\mathbb{E}_{e=(u,w)\in E}\left[\sum_{r\in[N]}\left(\left(\sum_{s\in(\pi^{e,u})^{-1}(r)} \widehat{f_u}(s)\right)^2 + \left(\sum_{s'\in(\pi^{e,w})^{-1}(r)} \widehat{f_w}(s')\right)^2\right)\right] \\ &\quad -B\cdot\mathbb{E}_{v\in V}\left[\sum_{\substack{S\subseteq[M]\\|S|\neq 1}} \widehat{f_v}(S)^2\right] \\ &\quad +A_{cross}(F) \end{aligned}$$

where $A_{cross}(F)$ is a quadratic form involving terms of the type $\widehat{f_u}(i)\widehat{f_w}(j)$ where $u,w\in V$, $u\neq w$ and $i,j\in[M]$, and therefore does not contribute any diagonal terms. Defining $A^*_{cons}(F)$ as,

$$\begin{aligned} A^*_{cons}(F) &= A_{cons}(F) + B\mathbb{E}_{e=(u,w)\in E}\left[\sum_{r\in[N]}(R_{e,u,r}+R_{e,w,r})\right] \\ &\quad +B\mathbb{E}_{v\in V}\left[\sum_{\substack{S\subseteq[M]\\|S|\neq 1}} T_v(S)\right] \end{aligned} \tag{58}$$

we see that $A^*_{cons}(F)$ does not contain any diagonal terms. Similarly, defining $A^*_{dict}(F)$ as,

$$A^*_{dict}(F) = A_{dict}(F) - \mathbb{E}_{v\in V}\left[\sum_{\substack{S'\subseteq[M]\\|S'|=1}} T_v(S')\right] \tag{59}$$

it can be seen that $A^*_{dict}(F)$ does not contain any diagonal terms. Let $A_{fin}(F) = A^*_{cons}(F) + A^*_{dict}(F)$. Using

Equations (53), (57), (58) and (59) it is easy to see that,

$$|A^*_{cons}(F) - A_{cons}(F)| \leqslant \frac{2BM+B}{D} \quad \text{and} \quad |A^*_{dict}(F) - A_{dict}(F)| \leqslant \frac{1}{D}.$$

Therefore, we obtain the desired bound,

$$\left|A_{fin}(F) - A_{prel}(F)\right| \leqslant \frac{2BM+B+1}{D} \leqslant 1/n, \tag{60}$$

by our setting of B and D.

6 Approximation for $\mathrm{QM}(p)$

In this section we shall prove the following theorem.

Theorem 6.1. *For any fixed $p \geqslant 2$, $\mathrm{Vec}_p(A) \leqslant \gamma_p^2 \cdot \mathrm{Val}_p(A)$ for any instance A of $\mathrm{QM}(p)$. This implies a polynomial time γ_p^2 approximation for $\mathrm{QM}(p)$.*

Furthermore, for all constants $\varepsilon > 0$, there is a polynomial time (randomized) rounding procedure that rounds the solution to $\mathrm{Vec}_p(A)$ to obtain a $(1+\varepsilon)\gamma_p^2$ approximate solution to $\mathrm{Val}_p(A)$.

Let $A = (a_{ij})_{i,j=1}^n$ be an $n \times n$ symmetric matrix with diagonal entries all zero, given as an instance of $\mathrm{QM}(p)$ for a fixed $p \geqslant 2$. We have,

$$\mathrm{Val}_p(A) = \max\left\{ \sum_{i,j=1}^n a_{ij}x_ix_j \ : \ \{x_1,\ldots,x_n\} \subseteq \mathbb{R}, \ \sum_{i=1}^n |x_i|^p \leqslant 1 \right\}. \tag{61}$$

As shown in Kindler *et al.* [26] the above can be relaxed to the following convex program,

$$\mathrm{Vec}_p(A) = \max\left\{ \sum_{i,j=1}^n a_{ij}\langle u_i, u_j\rangle \ : \ \{u_1,\ldots,u_n\} \subseteq L_2, \ \sum_{i=1}^n \|u_i\|_2^p \leqslant 1 \right\}. \tag{62}$$

Let $v_1,\ldots,v_n$ denote an optimal solution to the above convex program. Let $h_1,\ldots,h_n$ be mean zero Gaussian random variables obtained by defining $h_i := \langle G, v_i\rangle$ $(1 \leqslant i \leqslant n)$, where G is a random Gaussian vector in the space spanned by $v_1,\ldots,v_n$. It is easy to see that the following properties are satisfied.

$$\sum_{i=1}^n \left(\mathbb{E}[h_i^2]\right)^{p/2} \leqslant 1 \qquad \text{and} \quad \mathbb{E}\left[\sum_{i,j=1}^n a_{ij}h_ih_j\right] = \mathrm{Vec}_p(A). \tag{63}$$

Now we simply note that,

$$
\begin{aligned}
\mathrm{Vec}_p(A) &= \mathbb{E}\left[\sum_{i,j=1}^{n} a_{ij}h_ih_j\right] \\
&= \mathbb{E}\left[\left(\sum_{k=1}^{n}|h_k|^p\right)^{2/p}\left[\sum_{i,j=1} a_{ij}\left(\frac{h_i}{(\sum_{k=1}^{n}|h_k|^p)^{1/p}}\right)\left(\frac{h_j}{(\sum_{k=1}^{n}|h_k|^p)^{1/p}}\right)\right]\right] \\
&\leqslant \mathbb{E}\left[\left(\sum_{k=1}^{n}|h_k|^p\right)^{2/p}\cdot \mathrm{Val}_p(A)\right] \quad (\text{By Definition of } \mathrm{Val}_p(A)) \\
&\leqslant \left(\sum_{k=1}^{n}\mathbb{E}\left[|h_k|^p\right]\right)^{2/p}\mathrm{Val}_p(A) \quad (\text{By Jensens Inequality and since } p \geqslant 2) \\
&= \left(\sum_{k=1}^{n}\gamma_p^p\left(\mathbb{E}\left[h_k^2\right]\right)^{p/2}\right)^{2/p}\mathrm{Val}_p(A) \quad (\text{By Definition of } \gamma_p \text{ and since } h_k \text{ is Gaussian}) \\
&\leqslant \gamma_p^2\mathrm{Val}_p(A) \quad (\text{By Equation (63)})
\end{aligned}
$$

which is the upper bound we wanted. Note that the upper bound is obtained directly without rounding the vectors. To complete the proof of Theorem 6.1 we need to demonstrate a polynomial time rounding algorithm that extracts a $\gamma_p^2(1+\delta)$ approximate solution $x_1^*,\ldots,x_n^*$ to $\mathrm{Val}_p(A)$ from the vectors $v_1,\ldots,v_n$ for any constant $\delta > 0$. This shall be our goal in the remainder of the section.

Before we do so we can first assume without the loss of generality that

$$|a_{12}| = 1 = \max_{1\leqslant i,j\leqslant n}|a_{ij}|, \tag{64}$$

by appropriately relabeling the entries of the matrix A and scaling them. Setting $x_1 = 1/2$ and $x_2 = a_{12}/|a_{12}|$ and $x_3,\ldots,x_n = 0$ we obtain that

$$\mathrm{Vec}_p(A) \geqslant \mathrm{Val}_p(A) \geqslant 1/4. \tag{65}$$

The following is the rounding algorithm that we shall analyze.

Algorithm Round$(A,\{v_1,\ldots,v_n\})$:

1. Let $T := n^{22}$. Sample T random Gaussian vectors $G_1,\ldots,G_T$ in the span of $v_1,\ldots,v_n$.

2. Define random variables $z_i^{(t)} := \langle G_t, v_i\rangle$ for all $1 \leqslant t \leqslant T$ and $1 \leqslant i \leqslant n$. In addition define

$$x_i^{(t)} := \frac{z_i^{(t)}}{\left(\sum_{k=1}^{n}|z_j^{(t)}|^p\right)^{1/p}},$$

   and,

$$\Delta_t := \sum_{i,j=1}^{n} a_{ij}x_i^{(t)}x_j^{(t)}.$$

3. Let $t^* \in \{1,\ldots,T\}$ be such that $\Delta_{t^*} = \max_{1\leqslant t\leqslant T}\Delta_t$. Output $x_1^*,\ldots,x_n^*$ as the solution where $x_i^* = x_i^{(t^*)}$

for $1 \leqslant i \leqslant n$.

Let $\mathbb{E}_t$ denote the expectation over the uniformly at random choice of t from $1,\ldots,T$. We begin with the following lemma.

Lemma 6.2. *Given the random variables constructed in the procedure* $\mathrm{Round}(A,\{v_1,\ldots,v_n\})$*, with probability at least* $1-1/n^8$ *over the choice of* $G_1,\ldots,G_T$ *the following inequality holds,*

$$\sum_{i,j=1}^{n} a_{ij}\mathbb{E}_t[z_i^{(t)}z_j^{(t)}] \geqslant (1-8/n^2)\mathrm{Vec}_p(A). \tag{66}$$

Proof. We begin by ignoring the terms corresponding to pairs i,j ($1 \leqslant i,j \leqslant n$) such that $\langle v_i,v_j\rangle$ is very small. Formally, Let $R := \{(i,j)\in[n]\times[n] \mid \left|\langle v_i,v_j\rangle\right| \geqslant 1/n^4\}$. We have,

$$\begin{aligned}
\left|\sum_{(i,j)\in R} a_{ij}\langle v_i,v_j\rangle - \sum_{i,j=1}^{n} a_{ij}\langle v_i,v_j\rangle\right| &\leqslant \sum_{(i,j)\notin R}\left|a_{ij}\langle v_i,v_j\rangle\right| \\
&\leqslant \sum_{(i,j)\notin R}|a_{ij}|\left(\frac{1}{n^4}\right) \\
&\leqslant n^2\left(\frac{1}{n^4}\right) = 1/n^2. \quad \text{(By Equation (64))}
\end{aligned} \tag{67}$$

Now consider any $(i,j)\in R$. As before, we have Gaussian random variables h_i and h_j such that $\mathbb{E}[h_ih_j] = \langle v_i,v_j\rangle$. Moreover, since $(i,j)\in R$, $|\mathbb{E}[h_ih_j]| \geqslant 1/n^4$. We also need a bound on the variance of h_ih_j. Clearly, $\mathbf{Var}[h_ih_j] \leqslant \mathbb{E}[h_i^2h_j^2]$. Also, from Equation (63) we have that $\mathbb{E}[h_i^2],\mathbb{E}[h_j^2] \leqslant 1$. Therefore, $\mathbb{E}[h_i^2h_j^2]$ is upper bounded by $\mathbb{E}[g^4]=3$ where g is a standard Gaussian variable with variance 1. We note that over the choice of $G_1,\ldots,G_T$, the random variables $z_i^{(t)}z_j^{(t)}$ are identically distributed as h_ih_j for all $1\leqslant t\leqslant T$. Moreover, since $G_1,\ldots,G_T$ are independent Gaussian vectors, the random variables $z_i^{(t)}z_j^{(t)}$ are also independent for $1\leqslant t\leqslant T$. Therefore,

$$\mathbf{Var}\left[\mathbb{E}_t[z_i^{(t)}z_j^{(t)}]\right] \leqslant \frac{\mathbf{Var}[h_ih_j]}{T} \leqslant \frac{1}{n^{20}}, \tag{68}$$

by our choice of T and where the variance is over the choice of $G_1,\ldots,G_T$. Moreover, since $\mathbb{E}\left[\mathbb{E}_t[z_i^{(t)}z_j^{(t)}]\right] = \mathbb{E}[h_ih_j]$, we have the following bound using Chebyshev's inequality.

$$\mathbf{Pr}\left[\left|\mathbb{E}_t[z_i^{(t)}z_j^{(t)}] - \mathbb{E}[h_ih_j]\right| \geqslant 1/n^5\right] \leqslant 1/n^{10}. \tag{69}$$

Since the above analysis holds for all pairs $(i,j)\in R$, using a union bound over all pairs the above implies that with probability at least $1-1/n^8$, the following holds,

$$\left|\sum_{(i,j)\in R} a_{ij}\mathbb{E}_t[z_i^{(t)}z_j^{(t)}] - \sum_{(i,j)\in R} a_{ij}\mathbb{E}[h_ih_j]\right| \leqslant (1/n^5)\sum_{(i,j)\in R} a_{ij} \leqslant 1/n^3, \tag{70}$$

where the final inequality is obtained using Equation (64). Combining the above with Equation (67) implies that the following holds with probability at least $1-1/n^8$,

$$\left|\sum_{i,j=1}^{n} a_{ij}\mathbb{E}_t[z_i^{(t)} z_j^{(t)}] - \sum_{i,j=1}^{n} a_{ij}\langle v_i, v_j\rangle\right| \leqslant 1/n^2 + 1/n^3 \leqslant 2/n^2$$

This implies that with probability at least $1-1/n^8$,

$$\sum_{i,j=1}^{n} a_{ij}\mathbb{E}_t[z_i^{(t)} z_j^{(t)}] \geqslant \mathrm{Vec}_p(A) - 2/n^2 \geqslant (1-8/n^2)\mathrm{Vec}_p(A),$$

where the last inequality follows from Equation (65). This completes the proof of the lemma. □

The next lemma also proves a similar bound for the pth moments of the variables Gaussian variables h_i.

Lemma 6.3. *With probability at least $1-1/n^8$ over the choice of $G_1,\ldots,G_T$ the following holds for every $i=1,\ldots,n$.*

$$\left|\mathbb{E}_t[|z_i^{(t)}|^p] - \mathbb{E}[|h_i|^p]\right| \leqslant 1/n^4. \tag{71}$$

Proof. Let us fix $i \in \{1,\ldots,n\}$ for the moment. As noted before, over the choice of $G_1,\ldots,G_T$ the random variables $z_i^{(t)}$, $1 \leqslant t \leqslant T$, are independent random variables distributed identically to h_i. Now we have,

$$\mathbf{Var}[|h_i|^p] \leqslant \mathbb{E}[|h_i|^p] \leqslant \gamma_p^p(\mathbb{E}[|h_i^2|])^{2/p} \leqslant \gamma_p^p,$$

where the second last inequality is by the definition of γ_p and the last inequality uses Equation (63). Since $z_i^{(t)}$ are independent for $1 \leqslant t \leqslant T$, this implies,

$$\mathbf{Var}\left[\mathbb{E}_t[|z_i^{(t)}|^p]\right] \leqslant \frac{\mathbf{Var}[|h_i|^p]}{T} \leqslant \frac{\gamma_p^p}{n^{22}} \leqslant \frac{1}{n^{20}}, \tag{72}$$

for large enough n. Therefore, by Chebyshev's inequality we obtain,

$$\mathbf{Pr}\left[\left|\mathbb{E}_t[|z_i^{(t)}|^p] - \mathbb{E}[|h_i|^p]\right| \geqslant 1/n^5\right] \leqslant 1/n^{10}. \tag{73}$$

Taking a union bound over all $i=1,\ldots,n$ and rearranging Equation (73) proves the lemma. □

We are now ready to prove the desired bounds on the performance of the rounding algorithm Round$(A,\{v_1,\ldots,v_n\})$. For this we need to prove a upper bound on $\mathrm{Vec}_p(A)$ in terms of Δ^*. This is shown through the following series of inequalities implied by the two previous lemmas whose conditions hold with probability at least $1-2/n^8$.

$$
\begin{aligned}
(1-8/n^2)\mathrm{Vec}_p(A) &\leqslant \mathbb{E}_t\left[\sum_{i,j=1}^n a_{ij} z_i^{(t)} z_j^{(t)}\right] \qquad \text{(By Lemma 6.2)} \\
&\leqslant \mathbb{E}_t\left[(\sum_{k=1}^n |z_k^{(t)}|^p)^{2/p} \sum_{i,j=1}^n a_{ij}\left(\frac{z_i^{(t)}}{(\sum_{k=1}^n |z_k^{(t)}|^p)^{1/p}}\right)\left(\frac{z_j^{(t)}}{(\sum_{k=1}^n |z_k^{(t)}|^p)^{1/p}}\right)\right] \\
&\leqslant \mathbb{E}_t\left[(\sum_{k=1}^n |z_k^{(t)}|^p)^{2/p} \Delta^*\right] \qquad \text{(By the definition of } \Delta^*) \\
&\leqslant \left(\sum_{k=1}^n \mathbb{E}_t[|z_k^{(t)}|^p]\right)^{2/p} \Delta^* \qquad \text{(By Jensen's inequality since } p>2) \\
&\leqslant \left(\sum_{k=1}^n \left[\mathbb{E}[|h_k|^p] + 1/n^4\right]\right)^{2/p} \Delta^* \qquad \text{(By Lemma 6.3)} \\
&= \left(\sum_{k=1}^n \left[\gamma_p^p (\mathbb{E}[|h_k|^2])^{p/2} + 1/n^4\right]\right)^{2/p} \Delta^* \qquad \text{(By Definition of } \gamma_p) \\
&\leqslant \gamma_p^2 \left(1+1/n^3\right)^{2/p} \Delta^*. \qquad \text{(By Equation (63))}
\end{aligned}
$$

Since the parameter n is large enough the above analysis proves the approximation achieved by the rounding algorithm. This completes the proof of Theorem 6.1 and concludes this section.

References

[1] N. Alon, K. Makarychev, Y. Makarychev, and A. Naor. Quadratic forms on graphs. In *37th ACM STOC*, pages 486–493, 2005. 4

[2] N. Alon and A. Naor. Approximating the cut-norm via Grothendieck's inequality. *SIAM J. Comput.*, 35(4):787–803, 2006. 4, 5

[3] S. Arora, B. Barak, and D. Steurer. Subexponential algorithms for Unique Games and related problems. In *51st IEEE FOCS*, pages 563–572, 2010. 3

[4] S. Arora, E. Berger, E. Hazan, G. Kindler, and M. Safra. On non-approximability for quadratic programs. In *46th IEEE FOCS*, pages 206–215, 2005. 5

[5] S. Arora, S. Khot, A. Kolla, D. Steurer, M. Tulsiani, and N. K. Vishnoi. Unique games on expanding constraint graphs are easy: extended abstract. In *Proc. 40th ACM STOC*, pages 21–28, 2008. 3

[6] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998. 30

[7] S. Arora and S. Safra. Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998. 30

[8] B. Barak, P. Raghavendra, and D. Steurer. Rounding semidefinite programming hierarchies via global correlation. In *Proc. 52nd IEEE FOCS*, pages 472–481, 2011. 3

[9] A. Brieden, P. Gritzman, and V. Klee. Inapproximability of some geometric and quadratic optimization problems. In *Approximation and Complexity in Numerical Optimization: Continuous and Discrete Problems*, 2000. 5

[10] M. Charikar and A. Wirth. Maximizing quadratic programs: Extending Grothendieck's inequality. In *45th IEEE FOCS*, pages 54–60, 2004. 4, 5

[11] A. Deshpande, M. Tulsiani, and N. Vishnoi. Algorithms and hardness for subspace approximation. In *Proc. SODA*, 2011. 1, 3, 5, 6

[12] A. Deshpande and K. R. Varadarajan. Sampling-based dimension reduction for subspace approximation. In *Proc. 39th STOC*, pages 641–650, 2007. 5

[13] D. Feldman, M. Monemizadeh, C. Sohler, and D. P. Woodruff. Coresets and sketches for high dimensional subspace approximation problems. In *Proc. SODA*, pages 630–649, 2010. 5

[14] V. Feldman, V. Guruswami, P. Raghavendra, and Y. Wu. Agnostic learning of monomials by halfspaces is hard. In *Proc.* 50th *IEEE FOCS*, 2009. 5, 6

[15] G. Golub and C. van Loan. Matrix Computations. *Johns Hopkins University Press*, 1996. 5

[16] P. Gopalan, S. Khot, and R. Saket. Hardness of reconstructing multivariate polynomials over finite fields. In *Proc.* 48th *IEEE FOCS*, 2007. 6

[17] V. Guruswami, J. Håstad, R. Manokaran, P. Raghavendra, and M. Charikar. Beating the random ordering is hard: Every ordering CSP is approximation resistant. *SIAM J. Comput.*, 40(3):878–914, 2011. 3

[18] V. Guruswami and A. K. Sinop. Lasserre hierarchy, higher eigenvalues, and approximation schemes for graph partitioning and quadratic integer programming with PSD objectives. In *Proc. 52nd IEEE FOCS*, pages 482–491, 2011. 3

[19] S. Har-Peled and K. R. Varadarajan. Projective clustering in high dimensions using core-sets. In *Symposium on Computational Geometry*, pages 312–318, 2002. 5

[20] S. Khot. Hardness results for coloring 3-colorable 3-uniform hypergraphs. In *43rd IEEE FOCS*, pages 23–32, 2002. 6, 31

[21] S. Khot. On the power of unique 2-prover 1-round games. In *Proc.* 34th *ACM STOC*, pages 767–775, 2002. 3, 5

[22] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput.*, 37(1):319–357, 2007. 3, 4

[23] S. Khot and O. Regev. Vertex cover might be hard to approximate to within 2-epsilon. *J. Comput. Syst. Sci.*, 74(3):335–349, 2008. 3

[24] S. Khot and R. Saket. A 3-query non-adaptive PCP with perfect completeness. In *Proc. IEEE CCC*, pages 159–169, 2006. 6

[25] S. Khot and R. Saket. On hardness of learning intersection of two halfspaces. In *Proc.* 40^{th} *ACM STOC*, pages 345–354, 2008. 5, 6

[26] G. Kindler, A. Naor, and G. Schechtman. The UGC hardness threshold of the l_p Grothendieck problem. *Math. Oper. Res.*, 35(2):267–283, 2010. 1, 4, 5, 6, 7, 8, 12, 24

[27] A. Megretski. Relaxations of quadratic programs in operator theory and system analysis. In *Systems, approximation, singular integral operators, and related topics (Bordeaux)*, number 3, pages 365–392, 2001. 5

[28] E. Mossel. Gaussian bounds for noise correlation of functions and tight analysis of long codes. In *49^{th} IEEE FOCS*, pages 156–165, 2008. 4

[29] A. Naor and G. Schechtman. An approximation scheme for quadratic form maximization on convex bodies. *Manuscript*. 1, 4, 7

[30] A. Nemirovski, C. Roos, and T. Terlaky. On maximization of quadratic form over intersection of ellipsoids with common center. *Math. Program.*, 86(3, Ser. A):463–473, 1999. 4, 5

[31] Y. Nesetrov. Global quadratic optimization via conic relaxation. *Working Paper CORE*, 1998. 5

[32] P. Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proc.* 40^{th} *ACM STOC*, pages 245–254, 2008. 3

[33] R. Raz. A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998. 31

[34] N. D. Shyamalkumar and K. R. Varadarajan. Efficient subspace approximation algorithms. In *Proc. SODA*, pages 532–540, 2007. 5

[35] K. R. Varadarajan, S. Venkatesh, Y. Ye, and J. Zhang. Approximating the radii of point sets. *SIAM J. Comput.*, 36(6):1764–1776, 2007. 5

[36] U. Zwick. Approximation algorithms for constraint satisfaction problems involving at most three variables per constraint. In *Proc.* 9^{th} *ACM-SIAM SODA*, pages 201–210, 1998. 4

A Construction of Smooth Label Cover and Proof of Theorem 3.5

An instance ϕ of Max-3-SAT(5) is a 3-CNF formula in which each variable occurs in exactly 5 clauses. The PCP Theorem [7, 6] states that it is NP-hard to decide whether ϕ is satisfiable (YES instance), or at most $1-\varepsilon_0$ fraction of its clauses can be satisfied by any assignment, for some universal constant $\varepsilon_0 > 0$ (NO instance).

Consider the following reduction from Max-3-SAT(5) to a 2-Prover 1-Round Game. Given an instance ϕ of Max-3-SAT(5), the verifier chooses a subset $\mathscr{C}$ of $(J+1)R$ clauses at random, and a subset $\mathscr{C}'$ of R clauses from $\mathscr{C}$. Independently for each clause in $\mathscr{C}'$, the verifier chooses a variable uniformly at random from those in the clause. Let the collection of these variables be $\mathscr{X}$. The verifier sends the set $\mathscr{C}$ of clauses to the first prover, and the sets $\mathscr{C} \setminus \mathscr{C}'$ of clauses and $\mathscr{X}$ of variables to the second prover. The verifier expects back from each prover an assignment to all the variables received by it, and accepts if the answers of the provers are consistent on the common variables and satisfy all the clauses in $\mathscr{C}$.

The above construction is due to Khot [20], and is equivalent to a bipartite Label Cover instance $\mathscr{B}$ on vertex sets U and V, with projections $\pi_{vu} : [M] \mapsto [N]$ for edges between $u \in U$ and $v \in V$, where $M = 7^{(J+1)R}$ and $N = 2^R 7^{JR}$. The graph is bi-regular with the degree of vertices in U being 5^R, and of vertices in V being $\binom{(J+1)R}{R} 3^R$. For each $i \in [N]$ and edge (u,v), $\left|\pi_{vu}^{-1}(i)\right| \leqslant t := 4^R$. The size of the construction is $n^{O(JR)}$, where n is the number of clauses in ϕ. A pair of labelings $\sigma_U : U \mapsto [N]$ and $\sigma_V : V \mapsto [M]$ satisfy an edge (u,v) if $\sigma_U(u) = \pi_{vu}(\sigma_V(v))$.

Consider an R-round parallel repetition of the standard clause-variable 2-Prover 1-Round game derived from ϕ. It is shown formally by Khot [20] that this game can be *embedded* into $\mathscr{L}$. Therefore, the PCP Theorem and the Parallel Repetition Theorem [33] yield a universal constant $c_0 > 0$ such that,

- (YES Case) If ϕ is a YES instance then there is a labeling to U and V that satisfies all the edges of $\mathscr{B}$.
- (NO Case) If ϕ is a NO instance then every labeling to U and V satisfies less than a fraction $2^{-c_0 R}$ fraction of edges of $\mathscr{B}$.

Furthermore, as Khot [20] shows, the instance $\mathscr{B}$ satisfies the following smoothness property: for any vertex $w \in V$ and $i, j \in [M]$ s.t. $i \neq j$,

$$\mathbf{Pr}_{u \sim w}[\pi_{wu}(i) = \pi_{wu}(j)] \leqslant 1/J,$$

where the probability is taken over a random neighbor u of w. This is because i and j correspond to distinct assignments to a fixed set of clauses $\mathscr{C}$ (received by first prover) differing on at least one clause, say $C \in \mathscr{C}$. The LHS of the above equation is upper bounded by the probability (over the choice of $\mathscr{C}'$) that $C \in \mathscr{C}'$, which is at most $1/J$.

The above instance is converted into an instance $\mathscr{L}$ of Smooth Label Cover with vertex set V and label sets $[M]$ and $[N]$ as follows: for every vertex $u \in U$ and its neighbors v and w, add an edge $e = \{v, w\}$ in $\mathscr{L}$ with projections $\pi^{e,v} = \pi_{vu}$ and $\pi^{e,w} = \pi_{wu}$. The bi-regularity and smoothness of $\mathscr{B}$ along with this construction directly imply that $\mathscr{L}$ is regular with degree depending only on R and J, and satisfies the smoothness property in Theorem 3.5.

Given labelings σ_U and σ_V that satisfy all edges in $\mathscr{B}$, σ_V satisfies all edges in $\mathscr{L}$. Thus, a YES instance $\mathscr{B}$ is transformed into a YES instance $\mathscr{L}$. On the other hand, assume that there is a labeling σ_V to V that satisfies ζ fraction of edges of $\mathscr{L}$. Consider the following randomized labeling σ_U to U: independently for each $u \in U$, choose a neighbor $v \in V$ u.a.r, and assign u the label $\pi^{e,v}(\sigma_V(v))$. It is easy to see that the expected fraction edges of $\mathscr{B}$ satisfied by σ_V and σ_U is the probability over a uniformly random $u \in U$, and two of its neighbors v and w chosen independently and u.a.r, that $\pi_{vu}(\sigma_V(v)) = \pi_{wu}(\sigma_V(w))$. This is exactly the probability over the choice of a random edge $e = \{v, w\}$ of $\mathscr{L}$, that $\pi^{e,v}(\sigma_V(v)) = \pi^{e,w}(\sigma_V(w))$. By our assumption this is at least ζ. Therefore, if $\mathscr{B}$ is a NO instance, then $\mathscr{L}$ is also a NO instance.

Finally, to see the weak expansion property, let $V' \subseteq V$ s.t. $|V'| = \delta|V|$. For each $u \in U$, let p_u be the fraction of the neighbors of u which are in V'. Thus, $\mathbb{E}_u[p_u] = \delta$, and the fraction of edges [1] in $\mathscr{L}$ induced by V' is $\mathbb{E}_u[p_u^2] \geqslant \mathbb{E}_u[p_u]^2 = \delta^2$.

[1] We are also counting the (negligible) fraction self-loops produced in $\mathscr{L}$. These are satisfied by any labeling, and are ignored in the hardness reductions.

ECCC ISSN 1433-8092
http://eccc.hpi-web.de