The UGC hardness threshold of the ℓ_p Grothendieck problem

Guy Kindler* Assaf Naor† Gideon Schechtman‡

Abstract

For $p \geq 2$ we consider the problem of, given an $n \times n$ matrix $A = (a_{ij})$ whose diagonal entries vanish, approximating in polynomial time the number

$$\mathrm{Opt}_p(A) \colon = \max \left\{ \sum_{i,j=1}^n a_{ij} x_i x_j : \ (x_1, \ldots, x_n) \in \mathbb{R}^n \ \wedge \ \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \leq 1 \right\}.$$

When $p = 2$ this is simply the problem of computing the maximum eigenvalue of A, while for $p = \infty$ (actually it suffices to take $p \approx \log n$) it is the Grothendieck problem on the complete graph, which was shown to have a $O(\log n)$ approximation algorithm in [27, 26, 15], and was used in [15] to design the best known algorithm for the problem of computing the maximum correlation in Correlation Clustering. Thus the problem of approximating $\mathrm{Opt}_p(A)$ interpolates between the spectral ($p = 2$) case and the Correlation Clustering ($p = \infty$) case. From a physics point of view this problem corresponds to computing the ground states of spin glasses in a hard-wall potential well.

We design a polynomial time algorithm which, given $p \geq 2$ and an $n \times n$ matrix $A = (a_{ij})$ with zeros on the diagonal, computes $\mathrm{Opt}_p(A)$ up to a factor $\frac{p}{e} + 30 \log p$. On the other hand, assuming the unique games conjecture (UGC) we show that it is NP-hard to approximate (2) up to a factor smaller than $\frac{p}{e} + \frac{1}{4}$. Hence as $p \to \infty$ the UGC-hardness threshold for computing $\mathrm{Opt}_p(A)$ is exactly $\frac{p}{e}(1 + o(1))$.

1 Introduction

In this paper we consider the problem of maximizing a multilinear quadratic polynomial over a convex set $K \subseteq \mathbb{R}^n$. Namely, given a symmetric $n \times n$ matrix $A = (a_{ij})$ whose diagonal entries vanish, the goal is to approximate in polynomial time the number

$$\max \left\{ \sum_{i,j=1}^n a_{ij} x_i x_j : \ (x_1, \ldots, x_n) \in K \right\}. \tag{1}$$

In recent years there has been a lot of work on approximating such polynomials in the special case where K is the hypercube (by convexity, it makes no difference if one considers the solid hypercube $[-1, 1]^n$ or the discrete hypercube $\{-1, 1\}^n$). The case of optimizing over the hypercube has a wide

*Weizmann Institute. Email: gkindler@weizmann.ac.il.

†Courant Institute. Email: naor@cims.nyu.edu. Supported by in part by NSF grants CCF-0635078 and DMS-0528387.

‡Weizmann Institute. Email: gideon@weizmann.ac.il. Supported in part by the Israel Science Foundation.

range of applications to combinatorial optimization, and also has connections to topics in classical Banach space theory such as Grothendieck's inequality. We refer to [27, 26, 4, 15, 3, 5, 23, 2] and the references therein for both positive and negative results in this case, as well as for their applications.

Optimization over other bodies K is interesting as well. The case where K is a simplex has been investigated in [20, 16], partly in connection to problems in computational biology. The case when K is a polytope with polynomially many facets is classical, and is among the most important non-linear optimization problems, with a wide range of applications in operations research, computational biology and economics (see [17, 10, 13] for more information on the computational complexity of such problems).

In this work our focus is on cases where K is the unit ball in ℓ_p^n for some parameter p. The specific case mentioned above where K is the hypercube is obtained by setting $p = \infty$, and it is computationally hard: an $O(\log n)$ approximation algorithm is known (it was discovered independently in [29], [28], [25], and [15]), but it was shown to be NP-hard in [31], it was shown NP-hard to approximate within some constant factor in [4], and in [5] it was shown to be NP-hard for any constant factor approximation. The latter paper also showed that getting an approximation factor better than $(\log n)^\gamma$ is quasi-NP hard for some universal constant $\gamma > 0$, and that even improving on the $O(\log n)$ approximation can be ruled out under some plausible complexity assumption.

Setting $p = 2$, on the other hand, corresponds to the case where K is the Euclidean unit ball, and is much easier computationally. In this case the value in (1) is the maximum eigenvalue of the coefficient matrix $A = (a_{ij})$, and can be computed efficiently with arbitrarily good precision.

1.1 Our results

It is natural to ask what happens for values of p that lie between 2 and ∞. Roughly speaking, this set of problems can be viewed as a smooth interpolation between Spectral Partitioning and Correlation Clustering (the connection between (1) when K is the hypercube and the Correlation Clustering problem was discovered in [15]). In this paper we investigate these problems and give both new algorithms and complexity lower-bounds (the lower bounds being based on the Unique Games Conjecture). We note that the proofs of our results use the assumption $p \geq 2$, and therefore do not apply to the case $1 \leq p < 2$, which we did not investigate. The case $p = 1$ was studied in [20, 16], and apart from that nothing seems to be known for $1 < p < 2$.

The following theorem contains the approximation factors that we can achieve for $2 < p < \infty$, as well as the hardness factors that we can prove.

Theorem 1.1. *There is a polynomial time algorithm which, given $p \geq 2$ and an $n \times n$ matrix $A = (a_{ij})$ whose diagonal entries vanish, computes the number*[1]

$$\max \left\{ \sum_{i,j=1}^{n} a_{ij} x_i x_j : \ (x_1, \ldots, x_n) \in \mathbb{R}^n \ \wedge \ \sum_{j=1}^{n} |x_j|^p \leq 1 \right\} \tag{2}$$

up to a factor $\frac{p}{e} + 30 \log p$.

On the other hand, assuming the unique games conjecture, it is NP-hard to approximate (2) *for any constant $p > 2$ up to a factor smaller than $\frac{p}{e} + \frac{1}{4}$.*

[1] As we remark in Section 3, we can actually realize an approximate solution for eqrefeq:QP(p).

Hence, assuming the unique games conjecture, as $p \to \infty$ the NP-hardness threshold of the convex program (2) is $\frac{p}{e}(1 + o(1))$.

The unique games conjecture (UGC), which has been put forth by Khot [22], is a commonly used assumption in complexity theory. We describe it formally in Section 2. The UGC has been used in the context of hardness results for quadratic programs such as (1) in [23]. For readers that are not familiar with the UGC let us say at this point that the hardness result in Theorem 1.1 should be viewed as evidence that efficiently computing (2) up to a factor smaller than $\frac{p}{e} + \frac{1}{4}$ is hard.

When p is close to 2. Theorem 1.1 is stated asymptotically as $p \to \infty$, but our hardness result actually shows that for every $\delta \in (0, 1)$ and $p > 2$ it is UGC hard to approximate (2) up to a factor smaller than $(1 - \delta)\gamma_p^2$, where γ_p is the p'th norm of a standard Gaussian random variable. Since

$$\gamma_p = \left(\frac{2^{p/2}\Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}} \right)^{1/p} \geq \sqrt{1 + c(p-2)},$$

where c is universal constant, we also obtain a non-trivial hardness of approximation result for every $p > 2$.

1.2 The relation to spin glass models.

The problems described above are natural from the point of view of solid state physics, since they are intimately related to the problem of efficient evaluation of ground states of spin glasses. In the spin glass model we are given n particles, denoted by $\{1, \ldots, n\}$, each of which has a spin, or magnetization, $x_i \in \mathbb{R}$. The energy corresponding to each pair i, j in the system is proportional to $x_i x_j$: we are given an $n \times n$ matrix of pairwise interactions $A = (a_{ij}) \in M_n(\mathbb{R})$ (the diagonal entries are zeros), and the total energy of the system is given by $\sum_{i,j=1}^{n} a_{ij} x_i x_j$. The system is constrained to be in a potential well, or equivalently, there is an external magnetic field acting on the particles. A "hard wall" potential well[2] would simply correspond to imposing the constraint $(x_1, \ldots, x_n) \in K$ for some $K \subseteq \mathbb{R}^n$.

Since nature seems to seek the state where the energy is minimized, many physicists are interested in the computational complexity of computing the ground state, i.e. the configuration which minimizes the total energy of the constrained system. This is precisely the problem described above (with the matrix A replaced by $-A$). One reason why physicists would be interested in the computation complexity of a ground state is that if this state is hard to find computationally, it may explain (or even predict) why certain systems cannot settle in their ground state. On the other hand, if a system does settle in a computationally-hard ground state, this would imply that it can somehow perform a hard computation.

There is a vast amount of literature on the computational aspects of the evaluation of ground states of physical systems, as this corresponds to understanding efficient mechanisms for pattern

[2]Physicists often also consider "soft constraints", in which we are given a potential $V : \mathbb{R}^n \to \mathbb{R}$ and we wish to minimize $\sum_{i,j=1}^{n} a_{ij} x_i x_j + V(x_1, \ldots, x_n)$ (note that this has a similar effect to requiring that $(x_1, \ldots, x_n)$ is in the set where $V(x_1, \ldots, x_n)$ is not very large). The case when $V(x_1, \ldots, x_n) = \sum_{j=1}^{n} v(x_j)$ is called the case of one body potentials, and it is studied explicitly in the physics literature, including the especially important case $v(x) = |x|^p$, which corresponds to the problems studied here (for example, the ℓ_p case is studied in [30, 11]).

formation (see for example [33, 24]). The Ising case of the spin glass model corresponds to the assumption $x_i = \pm 1$. Computing the ground state in this famous simplified version clearly corresponds to (1) when K is the hypercube. Rigorous algorithmic results on the Ising case were obtained in [12, 9, 7, 3, 8]. The Ising model was introduced as a more tractable simplification of the original spin glass model, and in physically realistic scenarios the magnetization of the particles should be allowed to take real values (actually the most interesting case is when the spins are elements of the 2-dimensional sphere S^2).

Our results give evidence of a threshold behavior of the computational tractability of the ground state, when the hard wall constraint corresponds to the ℓ_p^n ball, $p > 2$. We believe that this phenomenon holds true for more general potentials (see the "Discussion and open problems" section). It would be interesting to find a physical explanation of this computational phase transition[3].

1.3 About the techniques

The algorithm that we design in Theorem 1.1 departs from the standard semidefinite programming approach that was used in [26, 4, 15, 3] by considering the convex program (which is not an SDP for finite p) that computes the quantity

$$\mathrm{Vec}_p(A) \colon = \max \left\{ \sum_{i,j=1}^{n} a_{ij} \langle v_i, v_j \rangle : \ \{v_1, \ldots, v_n\} \subseteq \mathbb{R}^n \ \wedge \ \left(\sum_{i=1}^{n} \|v_i\|_2^p \right)^{1/p} \leq 1 \right\}.$$

This program can be solved with arbitrarily small error in polynomial time using Grötschel-Lovász-Schrijver theory [18].

Let $v_1, \ldots, v_n \in \mathbb{R}^n$ be the the output of this program. In the case $p = \infty$, which corresponds to the constraint $\max_{1 \leq j \leq n} \|v_j\|_2 \leq 1$, one can proceed as follows. Since the value of $\mathrm{Vec}_\infty(A)$ depends only on the scalar products $\{\langle v_i, v_j \rangle\}_{i,j=1}^n$, we may replace $v_1, \ldots, v_n$ by $Uv_1, \ldots, Uv_n$ for any orthogonal matrix U. A standard computation shows that if we choose U uniformly at random among all $n \times n$ orthogonal matrices then for each $1 \leq j \leq n$ and $t > 0$ the probability that $(Uv_j)_k$ (the k^{th} coordinate of Uv_j) is greater than $\frac{t}{\sqrt{n}}$ is at most $e^{-\Omega(t^2)}$. Hence we can use the union bound to select with constant positive probability an orthogonal matrix U such that $x_{jk} \colon = \sqrt{cn/\log n} \cdot (Uv_j)_k \leq 1$ for all $1 \leq i, j \leq n$, where c is a universal constant. Since by definition

$$\mathrm{Vec}_\infty(A) = \sum_{i,j=1}^{n} a_{ij} \langle v_i, v_j \rangle = \sum_{i,j=1}^{n} a_{ij} \langle Uv_i, Uv_j \rangle = \frac{\log n}{c} \cdot \frac{1}{n} \sum_{k=1}^{n} \sum_{i,j=1}^{n} a_{ij} x_{ik} x_{jk},$$

there exists some $1 \leq k \leq n$ for which $\sum_{i,j=1}^n a_{ij} x_{ik} x_{jk} \geq \frac{c}{\log n} \mathrm{Vec}_\infty(A)$. We therefore get the required $O(\log n)$ approximation algorithm (originally obtained in [29, 28, 25, 15]) by rounding v_j to $x_{jk} \in [-1, 1]$.

For bounded p we wish to obtain a constant factor approximation, so using a similar union bound approach would not work here. We overcome this problem via a bootstrapping argument which is a non-trivial adaptation of the Gaussian Hilbert space approach to Grothendieck's inequality [21] (this approach was previously used for algorithmic purposes in [4, 3]). Namely, we take a standard

[3] It should be pointed out here that this type of "static" complexity phase transition is different from another popular research direction in statistical physics—the relation between statistical phase transitions and average case hardness in random models. We refer to [1] and the references therein for more information on this topic

Gaussian vector $G \in \mathbb{R}^n$ and consider the scalars $\langle v_1, G\rangle, \dots, \langle v_n, G\rangle$. Rather than truncating each of the numbers $\langle v_i, G\rangle$ separately, we consider the event $\mathcal{E} = \{\sum_{i=1}^n |\langle v_i, G\rangle|^p \le M\}$ for some appropriately chosen $M > 0$. Our rounding algorithm rounds the vector v_j to the number $x_j\colon = \frac{1}{M^{1/p}}\langle v_j, G\rangle \mathbf{1}_{\mathcal{E}}$. We show that this rounding procedure works by bounding the expectation of the error term $\sum_{i,j=1}^n a_{ij}\langle v_i, v_j\rangle - \sum_{i,j=1}^n a_{ij}x_ix_j$ using Hölder's inequality. This error term is shown to be a small proportion of $\mathrm{Vec}_p(A)$ for an appropriate choice of M and the exponent used in Hölder's inequality. A careful optimization of these two parameters yields the optimal $\frac{p}{e}\big(1+o(1)\big)$ integrality gap. The details are presented in Section 3.

The hardness result in Theorem 1.1 is achieved via a reduction from the Unique Label Cover problem. While the reduction is similar to the one used in [5], its analysis is considerably more involved, and we need to reduce from the Unique Label Cover problem (and hence use the UGC) rather than from the Label Cover problem. This complication stems, in essence, from the following technical fact: if $\sum_{i=1}^n \varepsilon_i a_i$, where the ε_i's are i.i.d. symmetric Bernoulli random variables, is bounded in L_∞, then $\sum_{i=1}^n |a_i|$ is bounded, while a bound on $\mathbb{E}\,|\sum_{i=1}^n a_i\varepsilon_i|^p$ does not imply any bound on $\sum_{i=1}^n |a_i|$ which is independent on n. This issue raises significant technical difficulties in the case of bounded p, and analytic ideas that we use to overcome them might be of independent interest in the context of computational hardness results. See Section 4 for more details (specifically, the remark on the difference from the L_∞ case on page 15).

2 Preliminaries and notation

In this section we describe some definitions, notation, and basic facts that will be used throughout this paper. We start with a formal definition of the ℓ_p Grothendieck problem, which we also call the ℓ_p Quadratic Maximization problem.

Definition 2.1 (ℓ_p Quadratic Maximization problem)**.** *The Quadratic Maximization problem over ℓ_p, denoted* QM*(p) for short, is defined as follows. For a given parameter $p \ge 1$ (which is possibly a function of n), an instance of the* QM*(p) problem is a square matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ with zero diagonal entries. The goal is to compute*

$$\mathrm{Opt}_p(A)\colon = \max\left\{\sum_{i,j=1}^n a_{ij}x_ix_j \,:\, (x_1,\dots,x_n)\in\mathbb{R}^n \ \wedge\ \sum_{i=1}^n |x_i|^p \le 1\right\}.$$

Our hardness results will be based on the Unique Games Conjecture (see [22]). We shall now briefly present the necessary background on this topic.

In what follows, for an integer r, we denote the set $\{1,2,\dots,r\}$ by $[r]$.

Definition 2.2 (Unique Label Cover problem)**.** *An instance $\mathscr{L}$ of the* Unique Label Cover problem, which we denote ULC for short, is a tuple $\mathscr{L} = \big(V, E, r, \{\pi_e\}_{e\in E}\big)$, where V is a set of nodes, and $E \subseteq V\times V$ is a symmetric set of directed edges, namely if $(u,w)\in E$ then also $(w,u)\in E$. For a node $v \in V$, we define $d(v)$ to be the number of edges of the form (v,w) in E. We assume that E contains no loops, and that $d(v)\ge 1$ for every $v\in V$.

For every edge $e=(u,w)\in E$ a permutation $\pi_e : [r]\to[r]$ is given such that $\pi_{(u,w)} = \pi_{(w,u)}^{-1}$ for every edge $(u,w)\in E$. A function $\mathscr{A} : V \to [r]$ is called an assignment for $\mathscr{L}$. We say that an edge $e=(u,w)$ is satisfied by $\mathscr{A}$ if $\mathscr{A}(w) = \pi_e(\mathscr{A}(u))$. The goal in the Unique Label Cover problem is to find an assignment which maximizes the fraction of satisfied edges. We denote the maximum fraction of satisfied edges in an instance $\mathscr{L}$ by $\mathrm{Opt}(\mathscr{L})$. The number r is called the number of labels in $\mathscr{L}$.

Conjecture 2.3 (unique games conjecture)**.** *For any constants $\delta, \varepsilon > 0$ satisfying $\delta + \varepsilon < 1$ there is an integer $r = r(\delta, \varepsilon)$ such that it is NP-hard to distinguish between* ULC *instances $\mathscr{L}$ with r labels for which $\mathrm{Opt}(\mathscr{L}) \le \delta$, and instances with r labels for which $\mathrm{Opt}(\mathscr{L}) \ge 1 - \varepsilon$.*

As noted in the introduction, the unique games conjecture was put forth by Khot in [22], and it is a commonly used complexity assumption. Despite several recent attempts to get better polynomial time approximation algorithms for the Unique Label Cover problem (see the table in [14] for a description of the known results), the unique games conjecture still stands.

We shall now record some moment bounds for Gaussian random variables and sums of independent random variables. These bounds will be used extensively in what follows.

Let g be a standard Gaussian random variable. Recall that we denote its p'th moment by γ_p, i.e. $\gamma_p \colon = (\mathbb{E}|g|^p)^{1/p}$. Then

$$\gamma_p^p = \frac{2}{\sqrt{2\pi}} \int_0^\infty x^p e^{-x^2/2} dx \overset{(x=\sqrt{2t})}{=} \frac{2^{p/2}}{\sqrt{\pi}} \int_0^\infty t^{\frac{p+1}{2}-1} e^{-t} dt = \frac{2^{p/2}\Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}}.$$

The following version of Stirling's formula [6],

$$\sqrt{\frac{2\pi}{x}} \cdot \left(\frac{x}{e}\right)^x \le \Gamma(x) \le \sqrt{\frac{2\pi}{x}} \cdot \left(\frac{x}{e}\right)^x e^{\frac{1}{12x}},$$

implies that

$$\gamma_p^p \in \left[\sqrt{\frac{2}{e}} \left(\frac{p+1}{e}\right)^{p/2}, \sqrt{\frac{2}{e}} \left(\frac{p+1}{e}\right)^{p/2} e^{\frac{1}{6(p+1)}}\right]. \tag{3}$$

In fact the following stronger lower bound holds true for $p > 2$ (This bound is optimal in terms of the second order terms as $p \to 2$ and $p \to \infty$. Since we will not use it here, we shall not include its elementary, though tedious, proof):

$$\gamma_p^2 \ge \max\left\{1 + \left(1 - \frac{\gamma + \ln 2}{2}\right)(p-2), \frac{p}{e} + \frac{\ln 2}{e}\right\}, \tag{4}$$

where $\gamma = 0.5772...$ is Euler's constant.

Lemma 2.4. *Let $X_1, \dots, X_n$ be independent random variables such that $\mathbb{E}X_j = 0$ and $\sum_{j=1}^n \mathbb{E}X_j^2 = 1$. Assume also that for some $\delta \in \left(0, e^{-4}\right)$ we have $\sum_{j=1}^n \mathbb{E}|X_j|^3 < \delta$. Then for every $p \ge 1$,*

$$\left(\mathbb{E}\left|\sum_{j=1}^n X_j\right|^p\right)^{1/p} \ge \gamma_p \cdot \left(1 - 4\left(\log(1/\delta)\right)^{p/2} \delta\right).$$

Proof. By the Berry-Esseen theorem (see [19]. The constant we use below follows from [32]), for every $u > 0$ we have

$$\Pr\left(\left|\sum_{j=1}^n X_j\right| \ge u\right) \ge \Pr(|g| \ge u) - 2\sum_{j=1}^n \mathbb{E}|X_j|^3 \ge \Pr(|g| \ge u) - 2\delta.$$

Therefore, for every $a > 0$ we have

$$\begin{aligned}\mathbb{E}\left|\sum_{j=1}^n X_j\right|^p &= \int_0^\infty pu^{p-1}\Pr\left(\left|\sum_{j=1}^n X_j\right| \ge u\right)du \ge \int_0^a pu^{p-1}\Pr(|g| > u)du - 2\delta a^p \\ &\ge \sqrt{\frac{2}{\pi}}\int_0^a u^p e^{-u^2/2}du - 2\delta a^p = \gamma_p^p - \sqrt{\frac{2}{\pi}}\int_a^\infty u^p e^{-u^2/2}du - 2\delta a^p. \quad (5)\end{aligned}$$

Choosing $a = \gamma_p\sqrt{\log(1/\delta)}$ yields the required result, where we use the bound $\int_a^\infty u^p e^{-u^2/2}du \le 2a^{p-1}e^{-a^2/2}$, which holds whenever $a^2 > 2p$—this estimate follows from the inequality

$$\int_a^\infty u^p e^{-u^2/2}du = a^{p-1}e^{-a^2/2} + (p-1)\int_a^\infty u^{p-2}e^{-u^2/2}du \le a^{p-1}e^{-a^2/2} + \frac{p-1}{a^2}\int_a^\infty u^p e^{-u^2/2}du.$$

The proof of Lemma 2.4 is complete. □

For large p we will need a better bound, which is contained in the following lemma. Recall that a random variable X is symmetric if X and $-X$ have the same distribution.

Lemma 2.5. *Let $X_1, \ldots, X_n$ be independent symmetric random variables such that $\sum_{j=1}^n \mathbb{E}X_j^2 = 1$. Fix $p \ge 2$ and assume that $\max_{1\le j\le n}\sqrt{\mathbb{E}X_j^2} \le \frac{2}{p}$. Then*

$$\left(\mathbb{E}\left|\sum_{j=1}^n X_j\right|^p\right)^{1/p} \ge \sqrt{\frac{p}{e}} - \frac{2}{\sqrt{p}}.$$

Proof. Let $k \in \mathbb{N}$ be the largest integer such that $2k \le p$. Denote $a_j := \sqrt{\mathbb{E}X_j^2}$, so that $\sum_{j=1}^n a_j^2 = 1$, and write $\delta := \max_{1\le j\le n}|a_j| \le \frac{2}{p} \le \frac{1}{k}$. Note that

$$\mathbb{E}\left(\sum_{j=1}^n X_j\right)^{2k} = \sum_{\substack{\ell_1,\ldots,\ell_n\in\mathbb{N}\cup\{0\}\\ \ell_1+\cdots+\ell_n=k}} \frac{(2k)!}{\prod_{j=1}^n(2\ell_j)!}\prod_{j=1}^n \mathbb{E}X_j^{2\ell_j} \ge \frac{(2k)!}{2^k}\sum_{\substack{S\subseteq\{1,\ldots,n\}\\ |S|=k}}\prod_{j\in S}a_j^2. \quad (6)$$

On the other hand

$$
\begin{aligned}
1 &= \left(\sum_{j=1}^n a_j^2\right)^k \\
&= \sum_{\substack{\ell_1,\ldots,\ell_n\in\mathbb{N}\cup\{0\}\\ \ell_1+\cdots+\ell_n=k}} \frac{k!}{\prod_{j=1}^n \ell_j!}\prod_{j=1}^n a_j^{2\ell_j} \\
&\le k!\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\prod_{j\in S}a_j^2+\sum_{j=1}^n a_j^2 \sum_{\substack{r_1,\ldots,r_n\in\mathbb{N}\cup\{0\}\\ r_1+\cdots+r_n=k-1\\ r_j\ge 1}} \frac{k}{r_j+1}\cdot\frac{(k-1)!}{\prod_{i=1}^n r_i!}\prod_{i=1}^n a_i^{2r_i} \\
&\le k!\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\prod_{j\in S}a_j^2+\frac{k\delta^2}{2}\sum_{\substack{r_1,\ldots,r_n\in\mathbb{N}\cup\{0\}\\ r_1+\cdots+r_n=k-1}} |\{1\le j\le n:\ r_j\ge 1\}|\,\frac{(k-1)!}{\prod_{i=1}^n r_i!}\prod_{i=1}^n a_i^{2r_i} \\
&\le k!\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\prod_{j\in S}a_j^2+\frac{k(k-1)\delta^2}{2}\sum_{\substack{r_1,\ldots,r_n\in\mathbb{N}\cup\{0\}\\ r_1+\cdots+r_n=k-1}} \frac{(k-1)!}{\prod_{i=1}^n r_i!}\prod_{i=1}^n a_i^{2r_i} \\
&= k!\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\prod_{j\in S}a_j^2+\frac{k(k-1)\delta^2}{2}\left(\sum_{j=1}^n a_j^2\right)^{k-1} \\
&= k!\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\prod_{j\in S}a_j^2+\frac{k(k-1)\delta^2}{2}.
\end{aligned}
$$

Hence,

$$
\sum_{\substack{S\subseteq\{1,\ldots,n\}\\|S|=k}}\prod_{j\in S}a_j^2 \ge \frac{1-\frac{k(k-1)\delta^2}{2}}{k!}\ge\frac{1}{2k!}. \tag{7}
$$

Combining (7) with (6) we get that

$$
\left(\mathbb{E}\left|\sum_{j=1}^n X_j\right|^p\right)^{\frac1p} \ge \left(\mathbb{E}\left(\sum_{j=1}^n X_j\right)^{2k}\right)^{\frac{1}{2k}} \ge \left(\frac{(2k)!}{2^{k+1}k!}\right)^{\frac{1}{2k}} \ge \sqrt{\frac{2k}{e}}-\frac{1}{4\sqrt{k}} \ge \sqrt{\frac{p}{e}}-\frac{2}{\sqrt{p}},
$$

where we used Stirling's formula and the fact that $p\ge 2k\ge p-2$. □

3 A $\frac{p}{e}(1+o(1))$ approximation algorithm for $\mathrm{QM}(p)$

Let $A=(a_{ij})\in M_n(\mathbb{R})$ be an $n\times n$ matrix with zeros on the diagonal and fix $p\ge 2$. Recall that

$$
\mathrm{Opt}_p(A)=\max\left\{\sum_{i,j=1}^n a_{ij}x_ix_j:\ \{x_1,\ldots,x_n\}\subseteq\mathbb{R}\ \wedge\ \sum_{i=1}^n |x_i|^p\le 1\right\}, \tag{8}
$$

We define the following parameter:

$$\mathrm{Vec}_p(A) :\,= \max\left\{\sum_{i,j=1}^{n} a_{ij}\langle v_i, v_j\rangle :\ \{v_1,\ldots,v_n\}\subseteq L_2\ \wedge\ \sum_{i=1}^{n}\|v_i\|_2^p \le 1\right\} \tag{9}$$

(the maximum is indeed attained, from compactness).

In this section we show that $\mathrm{Vec}_p(A)$ can be efficiently computed, and that it approximates $\mathrm{Opt}_p(A)$ as stated in the first part of Theorem 1.1. We also note that the approximation actually works even if the diagonal entries of A are non-zero.

Claim 3.1. *There is a PTAS for computing* $\mathrm{Vec}_p(A)$*.*

Proof. Let P_n denote the space of all $n\times n$ matrices $P\in M_n(\mathbb{R})$ which are positive semidefinite. We also write $K :\,= \left\{(m_{ij})\in M_n(\mathbb{R}) :\ \sum_{i=1}^{n}|m_{ii}|^{p/2}\le 1\right\}$. Since $p\ge 2$ the set K is convex. Moreover, K has a polynomial time membership oracle (more generally, any body given by $\sum_{i=1}^{n} f(|m_{ii}|)\le 1$, where f is convex and computable, has polynomial time membership oracle). It follows from the Grötschel-Lovász-Schrijver convex optimization theory [18] that there is a PTAS for computing the maximum of the linear functional $\sum_{i,j=1}^{n} a_{ij}m_{ij}$ on $P_n\cap K$. This maximum is precisely $\mathrm{Vec}_p(A)$. $\square$

Theorem 3.2. *We have the following inequality:*

$$\mathrm{Opt}_p(A)\le \mathrm{Vec}_p(A)\le\left(\frac{p}{e}+30\log p\right)\cdot\mathrm{Opt}_p(A).$$

Therefore there is a polynomial time algorithm which computes $\mathrm{Opt}_p(A)$ *up to a factor of* $\left(\frac{p}{e}+30\log p\right)$*.*

Proof. The left hand inequality in Theorem 3.2 is obvious. As is often the case with Grothendieck-type inequalities, we will work with the Gaussian Hilbert Space—this approach to Grothendieck's inequality first appeared in print in a paper of Johnson and Lindenstrauss [21], and was used extensively in [4, 3]. Let $g_1, g_2, \ldots$ be i.i.d. standard Gaussian random variables, and assume that they are defined on some probability space $(\Omega, \Pr)$. The Gaussian Hilbert space H is the closure in $L_2(\Omega)$ of the linear span of $\{g_1, g_2, \ldots\}$. By the definition of $\mathrm{Vec}_p(A)$ there are $h_1,\ldots,h_n\in H$ such that

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

Fix $M>1$ which will be determined later, and consider the event

$$S :\,= \left\{\sum_{i=1}^{n}|h_i|^p > M\cdot\gamma_p^p\right\}. \tag{10}$$

The crucial point to note here is that since each h_i is a Gaussian random variable we have the identity $\mathbb{E}|h_i|^p=\gamma_p^p\left(\mathbb{E}h_i^2\right)^{p/2}$. Markov's inequality implies that

$$\Pr[S]\le\frac{\sum_{i=1}^{n}\mathbb{E}|h_i|^p}{M\cdot\gamma_p^p}\le\frac{1}{M}\sum_{i=1}^{n}\left(\mathbb{E}h_i^2\right)^{p/2}\le\frac{1}{M}. \tag{11}$$

Now (11), and an application of Hölder's inequality, implies that for every $q > 1$,

$$B\colon = \sum_{i=1}^n \left(\mathbb{E}\left(h_i \mathbf{1}_S\right)^2\right)^{p/2} \le \sum_{i=1}^n \left(\mathbb{E}h_i^{2q}\right)^{\frac{p}{2q}} \left(\Pr[S]\right)^{\frac{p(q-1)}{2q}} \le \sum_{i=1}^n \left(\gamma_{2q}^{2q} \cdot \left(\mathbb{E}h_i^2\right)^q\right)^{\frac{p}{2q}} \frac{1}{M^{\frac{p(q-1)}{2q}}}$$
$$= \left(\frac{\gamma_{2q}^{2q}}{M^{q-1}}\right)^{\frac{p}{2q}} \sum_{i=1}^n \left(\mathbb{E}h_i^2\right)^{p/2} \le \left(\frac{\gamma_{2q}^{2q}}{M^{q-1}}\right)^{\frac{p}{2q}}.$$

Hence, an application of the definition of $\mathrm{Vec}_p(A)$ to the vectors $\frac{h_i \mathbf{1}_S}{B^{1/p}} \in L_2(\Omega)$ implies that

$$\mathbb{E} \sum_{i,j=1}^n a_{ij} h_i h_j \mathbf{1}_S \le B^{2/p} \mathrm{Vec}_p(A) \le \left(\frac{\gamma_{2q}^{2q}}{M^{q-1}}\right)^{1/q} \cdot \mathrm{Vec}_p(A). \tag{12}$$

On the other hand, the definition of $\mathrm{Opt}_p(A)$ implies that

$$\mathbb{E} \sum_{i,j=1}^n a_{ij} h_i h_j \mathbf{1}_{\Omega \setminus S} = M^{2/p} \gamma_p^2 \cdot \mathbb{E} \sum_{i,j=1}^n a_{ij} \frac{h_i}{M^{1/p}\gamma_p} \cdot \frac{h_j}{M^{1/p}\gamma_p} \mathbf{1}_{\Omega \setminus S} \le M^{2/p} \cdot \gamma_p^2 \cdot \mathrm{Opt}_p(A). \tag{13}$$

Combining (12) and (13) we get that

$$\mathrm{Vec}_p(A) = \mathbb{E} \sum_{i,j=1}^n a_{ij} h_i h_j = \mathbb{E} \sum_{i,j=1}^n a_{ij} h_i h_j \mathbf{1}_S + \mathbb{E} \sum_{i,j=1}^n a_{ij} h_i h_j \mathbf{1}_{\Omega \setminus S}$$
$$\le \left(\frac{\gamma_{2q}^{2q}}{M^{q-1}}\right)^{1/q} \cdot \mathrm{Vec}_p(A) + M^{2/p} \cdot \gamma_p^2 \cdot \mathrm{Opt}_p(A). \tag{14}$$

The right-hand side of (14) is mimimized for

$$M = \left(\frac{p(q-1)}{2q} \cdot \left(\frac{\gamma_{2q}}{\gamma_p}\right)^2 \cdot \frac{\mathrm{Vec}_p(A)}{\mathrm{Opt}_p(A)}\right)^{\frac{pq}{pq+2q-p}}. \tag{15}$$

Plugging this value of M into (14) we get the inequality

$$\mathrm{Vec}_p(A) \le \left[\mathrm{Vec}_p(A)\right]^{\frac{2q}{pq+2q-p}} \cdot \left[\mathrm{Opt}_p(A)\right]^{\frac{p(q-1)}{pq+2q-p}}$$
$$\cdot \left(\gamma_{2q}\right)^{\frac{4q}{pq+2q-p}} \cdot \left(\gamma_p\right)^{\frac{2p(q-1)}{pq+2q-p}} \cdot \left[\left(\frac{2q}{p(q-1)}\right)^{\frac{p(q-1)}{pq+2q-p}} + \left(\frac{p(q-1)}{2q}\right)^{\frac{2q}{pq+2q-p}}\right],$$

which simplifies to give the bound,

$$\frac{\mathrm{Vec}_p(A)}{\mathrm{Opt}_p(A)} \le \left(\gamma_p\right)^2 \cdot \left(\gamma_{2q}\right)^{\frac{4q}{p(q-1)}} \cdot \left[\left(\frac{2q}{p(q-1)}\right)^{\frac{p(q-1)}{pq+2q-p}} + \left(\frac{p(q-1)}{2q}\right)^{\frac{2q}{pq+2q-p}}\right]^{1+\frac{2q}{p(q-1)}}. \tag{16}$$

Choosing $q = 2$ in (16), and using the bounds in (3), yields the required result. But by optimizing over q one can get better bounds – in particular one gets much better approximation factors for small p. For large p the optimal choice of q is $q = \Theta(\log p)$, and the bound becomes

$$\mathrm{Vec}_p(A) \le \left(\frac{p}{e} + \frac{2}{e}\log p + \frac{2}{e}\log\log p + O(1)\right)\mathrm{Opt}_p(A).$$

In any case, the proof of Theorem 3.2 is complete. □

Remark 3.3. We can actually round the vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ for which $\mathrm{Vec}_p(A)$ is approximately attained to scalars $x_1, \ldots, x_n$ which realize the approximation to $\mathrm{Opt}_p(A)$. Specifically, we can find random numbers $x_1, \ldots, x_n$ which satisfy $\sum_{i=1}^n |x_i|^p \le 1$ and $\mathbb{E}\sum_{i,j=1}^n a_{ij}x_ix_j \ge \mathrm{Opt}_p(A)/\left(\frac{p}{e} + O(\log p)\right)$. We can actually further show that the x_i's satisfy the above inequality with positive probability which only depends on p, but we omit the argument.

To choose the x_i's, we can concretely realize the Gaussians h_i that appeared in the proof of Theorem 3.2 as $h_i = \langle v_i, G\rangle$, where G is a standard Gaussian vector in $\mathbb{R}^n$. We then define $x_i \colon= \frac{1}{\widetilde{M}^{1/p}}\langle v_i, G\rangle \mathbf{1}_{\Omega\setminus S}$, where analogously to (15),

$$\widetilde{M} = \left(\frac{p(q-1)}{2q}\cdot\left(\frac{\gamma_{2q}}{\gamma_p}\right)^2\cdot\left(\frac{p}{e} + 30\log p\right)\right)^{\frac{pq}{pq+2q-p}},$$

$q \approx \log p$ and S is as in (10). By repeating the argument in the proof of Theorem 3.2 with the value $\widetilde{M}$ instead of the value M in (15), i.e. by replacing the term $\frac{\mathrm{Vec}_p(A)}{\mathrm{Opt}_p(A)}$ in the definition of M with its a priori bound which we already proved, shows that this rounding procedure yields the desired approximation factor.

4 UGC hardness

In this section we prove the UGC hardness of $\mathrm{QM}(p)$. We will make use of the notation and definitions in Section 2, and of the following definition.

Definition 4.1. *For $p \ge 2$ and $\varepsilon > 0$ let $\phi(p,\varepsilon)$ be the largest $\phi > 0$ such that for all $n \in \mathbb{N}$ if $(a_1, \ldots, a_n) \in \mathbb{R}^n$ satisfy $\sum_{i=1}^n a_i^2 = 1$ and $\max_{i\in[n]} a_i^2 \le \phi$ then*

$$\left(\mathbb{E}\left|\sum_{i=1}^n a_i\varepsilon_i\right|^p\right)^{1/p} \ge (1-\varepsilon)\gamma_p,$$

where $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. symmetric ± 1 Bernoulli random variables. The existence of $\phi(p,\varepsilon)$ follows from Lemma 2.4. Moreover, Lemma 2.5, combined with (3) *implies that*

$$\phi\left(p, \frac{10}{p}\right) \ge \frac{4}{p^2}. \tag{17}$$

It will be convenient to introduce the following notation. Given $p \ge 1$ and a function $f : X \to \mathbb{R}$ defined on a finite set X, we write $\mathbb{E}_{x\in X}f(x) \colon= \frac{1}{|X|}\sum_{x\in X} f(x)$ and $\|f\|_p \colon= (\mathbb{E}_{x\in X}|f(x)|^p)^{1/p}$ (thus in what follows ℓ_p norms will correspond exclusively to uniform distributions on finite sets). If μ is a probability distribution μ on X then we use the notation $\mathbb{E}_{x\sim\mu}f(x) \colon= \int_X f(x)d\mu(x)$.

The main result of this section is the following reduction.

Theorem 4.2. *There exists a reduction algorithm from* ULC *to* QM*(p) which, given a* ULC *instance* $\mathscr{L} = (V, E, r, \{\pi_e\}_{e\in E})$*, outputs an instance* A *of* QM*(p). The reduction algorithm runs in time polynomial in* $|V|$ *and* 2^r*, and has the following properties:*

- Completeness: $\mathrm{Opt}_p(A) \geq \mathrm{Opt}(\mathscr{L}) - \frac{1}{|V|}$
- Soundness: *For every* $\varepsilon > 0$ *it holds that*

$$\mathrm{Opt}_p(A) \leq \left(\frac{4\gamma_p^{12}}{\varepsilon^4 \phi(p,\varepsilon)^2} \cdot \mathrm{Opt}(\mathscr{L})\right)^{(p-2)/p} + \frac{1 + 2\varepsilon + 3/|V|}{(1-\varepsilon)\gamma_p^2} + \frac{1}{|V|}. \tag{18}$$

The proof of Theorem 4.2 spans the rest of this section, but before we commence with it, let us show how it implies the UGC hardness for approximating QM(p) (and thus the hardness part of Theorem 1.1). We note that while the following corollary shows QM(p) to be hard for constant p, it can be extended for some non-constant p's if one is willing to assume a hypothesis which is somewhat stronger than the Unique Games Conjecture.

Corollary 4.3. *Let* $p > 2$ *be any constant, and let* $\delta > 0$ *be a constant such that* $(1-\delta)\gamma_p^2 > 1$*. Then assuming the unique games conjecture, it is NP-hard to approximate* QM*(p) within a factor* $(1-\delta)\gamma_p^2$*. Using* (4) *it follows that assuming the unique games conjecture it is NP-hard to approximate* QM*(p) within a factor* $\frac{p}{e} + \frac{1}{4}$ *(note that we are using a crude, sub-optimal in terms of the additive constant term, version of* (4) *here).*

Proof. Set $\varepsilon' = \varepsilon = \delta/8$, and pick $\delta' > 0$ small enough so that $d' < 1 - \varepsilon'$ and

$$\left(\frac{4\gamma_p^{12}}{\varepsilon^4 \phi(p,\varepsilon)^2} \cdot \delta'\right)^{(p-2)/p} \leq \frac{\delta}{4\gamma_p^2}.$$

From the unique games conjecture we have that there exists an integer r, such that it is NP-hard to distinguish between instances $\mathscr{L}$ of ULC with r labels for which $\mathrm{Opt}(\mathscr{L}) \geq 1 - \varepsilon'$, and instances with r labels for which $\mathrm{Opt}(\mathscr{L}) \leq \delta'$. The reduction stated in Theorem 4.2 maps instances $\mathscr{L}$ of the first kind into QM(p) instances A with $\mathrm{Opt}_p(A) \geq 1 - \varepsilon' - 1/|V| = 1 - \delta/8 - 1/|V|$ (note that the reduction runs in polynomial time). We call such instances 'yes' instances.

By the soundness property, it maps instances $\mathscr{L}$ of the latter kind into instances A satisfying

$$\mathrm{Opt}_p(A) \leq \frac{\delta}{4\gamma_p^2} + \frac{1 + 2\varepsilon + 3/|V|}{(1-\varepsilon)\gamma_p^2} + \frac{1}{|V|} \leq \frac{\delta}{4\gamma_p^2} + \frac{1 + \delta/2 + 3/|V|}{\gamma_p^2} + \frac{1}{|V|}.$$

These are called 'no' instances.

Thus the ratio between the values of $\mathrm{Opt}_p(A)$ for 'yes' and for 'no' instances tends, as $|V|$ goes to infinity, to a number which is greater than $(1-\delta)\gamma_p^2$. This completes the proof of Corollary 4.3. □

4.1 The reduction

Let us begin by describing the reduction algorithm. Let $\mathscr{L} = (V, E, r, \{\pi_e\}_{e\in E})$ be an instance of ULC. The reduction will make use of the following parameters:

$$D \colon = |V|^3 \cdot |E| \cdot 2^{2r} \qquad \text{and} \qquad B \colon = |V|^2 \cdot |E| \cdot 2^{2r}. \tag{19}$$

Some notation. Given $\mathscr{L}$, the reduction should output a square matrix A with zero diagonal entries. Equivalently, we prefer to think of the output of the reduction as a *multilinear form* defined on vectors F, and to write $A(F)$ instead of $\langle AF, F\rangle$ (multilinearity, i.e. linearity in each of the coordinates of the vector, is equivalent in matrix notation to A having zero entries on the diagonal).

The coordinates. It will be convenient to use meaningful indices for the coordinates of F. For every node $v \in V$ there will be $D \cdot d(v)$ *sets* of coordinates $\left\{C_v^j\right\}_{j\in[D\cdot d(v)]}$. Each such set C_v^j will contain 2^r coordinates, labeled $\left\{C_v^j(x)\right\}_{x\in\{-1,1\}^r}$. These coordinates serve as encodings of assignments for v: For F to encode the assignment i for v, it should satisfy

$$\forall\ j \in [D \cdot d(v)],\ \forall x \in \{-1, 1\}^r \qquad F_{C_v^j(x)} = x_i.$$

Given a vector F with coordinates indexed as above, we define a function $f_v^j : \{-1, 1\}^r \to \mathbb{R}$ for every $v \in V$ and $j \in [D \cdot d(v)]$, by setting

$$f_v^j(x) = F_{C_v^j(x)}$$

(for simplicity of notation, we keep the dependency of f_v^j on F implicit). We also define for every $v \in V$ a function $f_v : \{-1, 1\}^r \to \mathbb{R}$ by taking

$$f_v(x) := \mathbb{E}_{j\in[D\cdot d(v)]}\left[f_v^j(x)\right]. \tag{20}$$

The distribution μ. Before we proceed to define the value of the multilinear form A on F, let us note that for every $q \geq 1$ we can write the L_q norm of a vector F as follows. Let μ be the probability distribution on V defined by $\mu(v) := \frac{d(v)}{|E|}$ (recall that we defined $d(v)$ to be just the outgoing degree of v, and therefore this really is a probability distribution). Then for every $q \geq 1$,

$$\|F\|_q = \left(\mathbb{E}_{v\sim\mu}\mathbb{E}_{j\in[D\cdot d(v)]}\left\|f_v^j\right\|_q^q\right)^{1/q}. \tag{21}$$

Equation (21) follows directly from the definition of the functions f_v^j. Since $q \geq 1$, it now follows from the triangle inequality in L_q and from (20) that

$$\left(\mathbb{E}_{v\sim\mu}\|f_v\|_q^q\right)^{1/q} \leq \|F\|_q. \tag{22}$$

The inner and outer forms. We are ready to define the quadratic form A. Although A is defined over the coordinates of the vector F, it will be easier to describe it as a form over the Fourier coefficients of the functions $\{f_v\}_{v\in V}$. Such a quadratic form is also a quadratic form over F since each of these coefficients can be written as a linear combination of the coordinates of F. We shall deal with the multilinearity of the form later.

The form A will be a sum of two terms, one called the *inner* term and the other, the *outer* term. The inner term will serve to prevent A from taking large values unless the functions $\{f_v\}_{v\in V}$ associated with F are very close to being linear homogeneous. The role of the outer term is, roughly, to allow A to obtain large values on vectors F which encode assignments for $\mathscr{L}$ that satisfy a large fraction of the edges.

For each $v \in V$, we define a quadratic form A_v on F by

$$A_v(F) \colon = -B \cdot \sum_{\substack{S \subseteq [r] \\ |S| \neq 1}} \widehat{f_v}(S)^2.$$

We take the inner term to be

$$A_{inner}(F) \colon = \mathbb{E}_{v \sim \mu}\left[A_v(F)\right].$$

For every edge $e = (u, w) \in E$ we let

$$A_e(F) \colon = \sum_{i \in [r]} \widehat{f_u}(i) \widehat{f_w}(\pi_e(i)),$$

and define the outer term to be

$$A_{outer} \colon = \mathbb{E}_{e \in E}\left[A_e(F)\right].$$

The quadratic form. We define $A'(F) = A_{inner}(F) + A_{outer}(F)$. We would have liked to take A' as the output of the reduction algorithm, but it turns out to have some non-multilinear terms (i.e. some non-zero diagonal entries in matrix language). For simplicity of analysis, we therefore first establish the completeness and soundness properties for A' as it is defined here. In Subsection 4.4 we slightly change A' to remove the diagonal entries, obtaining the final quadratic form A, and show that for every vector F with $\|F\|_p \leq 1$,

$$|A'(F) - A(F)| \leq \frac{1}{|V|}. \tag{23}$$

Running time. It is obvious from the construction that the form A' can be constructed in time polynomial in $|V|$ and in 2^r as required. This will also hold for the actual form A, that is defined in Subsection 4.4.

Our next step is to establish, in the next two subsections, the completeness and soundness properties of A', and therefore, assuming (23), also of A.

4.2 Completeness

To get the completeness property we start with an assignment $\mathscr{A} : V \to [r]$ for $\mathscr{L}$ which satisfies an $\mathrm{Opt}(\mathscr{L})$ fraction of the edges, and use is to get a vector F with $\|F\|_p \leq 1$ such that $A'(F) = \mathrm{Opt}(\mathscr{L})$. Assuming (23), this gives the completeness property.

The vector F. We define F by

$$\forall\, v \in V,\ \forall\, j \in [D \cdot d(v)],\ \forall\, x \in \{-1,1\}^r, \qquad F_{C_v^j(x)} = x_{\mathscr{A}(v)}.$$

It is clear that $\|F\|_p = 1$, since all of its coordinates are either 1 or -1, and that the functions $\{f_v\}_{v\in V}$ associated with F satisfy

$$\forall\, v \in V,\ x \in \{-1,1\}^r, \qquad\qquad f_v(x) = x_{\mathscr{A}(v)}. \tag{24}$$

The functions f_v are therefore linear, which implies that $A_{inner}(F) = 0$. It also follows from (24) that for every edge $e \in E$, the value of $A_e(F)$ is 1 if $\mathscr{A}$ satisfies e and 0 otherwise, and therefore

$$A_{outer}(F) = \mathbb{E}e \in E\,[A_e(F)] = \Pr_{e\in E}[\mathscr{A} \text{ satisfies } e] = \mathrm{Opt}(\mathscr{L}).$$

Overall, we thus have $A'(F) = A_{inner}(F) + A_{outer}(F) = \mathrm{Opt}(\mathscr{L})$, as desired.

4.3 Soundness

To establish the soundness property, let F be any vector with $\|F\|_p \le 1$. We shall prove that

$$A'(F) \le \left(\frac{4\gamma_p^{12}}{\varepsilon^4 \phi(p,\varepsilon)^2} \cdot \mathrm{Opt}(\mathscr{L})\right)^{(p-2)/p} + \frac{1+2\varepsilon+3/|V|}{(1-\varepsilon)\gamma_p^2}, \tag{25}$$

and then the soundness property will follow from (25) together with (23). We assume w.l.o.g. that $A'(F) \ge 0$.

Our aim is to use the vector F to define an assignment for $\mathscr{L}$, which will prove that $\mathrm{Opt}(\mathscr{L})$ is large enough to make (25) hold.

On the difference between the ℓ_p case and the L_∞ case. Our definition of the assignment for $\mathscr{L}$ is different than that in [5]. In that paper every $v \in V$ was given a multi-assignment by taking all the elements $i \in [r]$ for which $|\widehat{f_v}(i)|$ was greater than a certain threshold. The assignment for v was then chosen randomly from among the elements of its multi-assignment. Since in [5] the L_∞ norm of F was bounded by 1, it was possible there to get a bound on $\sum_{i=1}^n |\widehat{f_v}(i)|$, and therefore on the size of the multi-assignment (using the fact that the L_∞ norm of a linear function f equals $\sum_{i=1}^n |\widehat{f}(i)|$). Since here we only have a bound on the ℓ_p norm of F, that approach does not work and the definition of the multi-assignment, as well as its analysis, are more involved.

Before we define the assignment for $\mathscr{L}$ we make the following definitions and observations.

Claim 4.4. *Suppose that a real number t_v is associated with every node $v \in V$. Then*

$$\mathbb{E}_{e=(u,w)\in E}\,[t_u t_w] \le \mathbb{E}_{v\sim\mu}\left[|t_v|^2\right]. \tag{26}$$

Proof. Consider the set of pairs $\{(v,j)\}_{v\in V,\, j\in[d(v)]}$ as a set of indices, and define a vector X by $X_{(v,j)} = t_v$. Then the r.h.s. of (26) equals $\|X\|_2^2$. The l.h.s. of (26) is the inner product of X with a vector whose coordinates are simply a permutation of the coordinates of X (this follows from the fact that E is symmetric). Hence (26) follows from the Cauchy-Schwarz inequality. □

The following claim follows by the same argument as in Claim 4.4.

Claim 4.5. *Let $S \subseteq E$ be a symmetric subset (that is, if $(u,w) \in S$ then also $(w,u) \in S$), and let $\{t_v\}_{v \in V}$ be any real numbers. Then*

$$\sum_{(u,w)\in S} t_u t_w \leq \sum_{(u,w)\in S} |t_u|^2.$$

The linear part of the functions f_v play an important role later on. Let us set a notation for them.

Definition 4.6. *For every $v \in V$, let f_v^l denote the linear homogeneous part of f_v, namely $f_v^l(x) = \sum_{i\in[r]} \widehat{f_v}(i) x_i$. Let $f_v^h = f_v - f_v^l$.*

We next establish some properties of the functions f_v^l and f_v^h.

Claim 4.7. $\max_{v\in V,\, x\in\{-1,1\}^r} \left(|f_v^h(x)|\right) \leq 2^r \cdot \sqrt{\frac{|E|}{B}} = \frac{1}{|V|}$.

Proof. First, note that

$$\begin{aligned}
A_{outer}(F) &= \mathbb{E}_{e=(u,w)\in E}\left[\sum_{i\in[r]} \widehat{f}_u(i)\widehat{f}_w(\pi_e(i))\right] \\
&\leq \mathbb{E}_{e=(u,w)\in E}\left[\|f_u\|_2 \|f_w\|_2\right] && \text{(Cauchy-Schwarz)} \\
&\leq \mathbb{E}_{v\sim\mu}\left[\|f_v\|_2^2\right] \leq \|F\|_2^2 && \text{by Claim 4.4 and by (22)} \\
&\leq \|F\|_p^2 \leq 1.
\end{aligned}$$

Hence, since we assumed that $A'(F) \geq 0$, we have from the definition of A_{inner}, μ, and the functions f_v^h that

$$-1 \leq A_{inner}(F) = -B \cdot \mathbb{E}_{v\sim\mu}\left[\left\|f_v^h\right\|_2^2\right] = -B \cdot \frac{\sum_{v\in V} d(v) \left\|f_v^h\right\|_2^2}{|E|}.$$

It follows that

$$\max_{v\in V}\left(\left\|f_v^h\right\|_1^2\right) \leq \max_{v\in V}\left(\left\|f_v^h\right\|_2^2\right) \leq \frac{|E|}{B},$$

and therefore that

$$\max_{v\in V,\, x\in\{-1,1\}^r}\left(|f_v^h(x)|\right) \leq 2^r \max_{v\in V}\left(\left\|f_v^h\right\|_1\right) \leq 2^r \cdot \sqrt{\frac{|E|}{B}} = \frac{1}{|V|},$$

where the last equality follows from the definition of B (19). This completes the proof of the claim. $\square$

Claim 4.8. $\left[\mathbb{E}_{v\sim\mu}\left\|f_v^l\right\|_p^p\right]^{1/p} \leq 1 + \frac{1}{|V|}$.

Proof. Using the triangle inequality in ℓ_p we have

$$\left[\mathbb{E}_{v\sim\mu}\left\|f_v^l\right\|_p^p\right]^{1/p} \leq \left[\mathbb{E}_{v\sim\mu}\|f_v\|_p^p\right]^{1/p} + \left[\mathbb{E}_{v\sim\mu}\left\|f_v^h\right\|_p^p\right]^{1/p} \overset{(22)}{\leq} \|F\|_p + \max_{v\in V,\, x\in\{-1,1\}^r}\left(|f_v^h(x)|\right) \leq 1 + \frac{1}{|V|},$$

where the last inequality follows from Claim 4.7 and the assumption that $\|F\|_p \leq 1$. $\square$

Influential coordinates. We now identify the influential coordinates in the functions f_v^l, which will serve as the basis for getting an assignment for $\mathscr{L}$. For this purpose, we pick the parameters

$$a\colon = \frac{\gamma_p^2}{\phi(p,\varepsilon)} \qquad \text{and} \qquad b\colon = \frac{\gamma_p^4 \cdot a}{\varepsilon^2},$$

where $\phi(p,\varepsilon)$ is as in Definition 4.1. We define for every $v \in V$ subsets $\alpha(v), \beta(v), \gamma(v) \subseteq [r]$ as follows: $\alpha(v)$ is taken to be the set of size a of coordinates $i \in [r]$ for which $|\widehat{f_v}(i)|$ is largest (ties are broken arbitrarily). $\beta(v)$ is taken to be the set of b coordinates[4] with the largest values of $|\widehat{f_v}(i)|$ among $[r] \setminus \alpha(v)$, and $\gamma(v)$ is defined by $\gamma(v) = [r] \setminus (\alpha(v) \cup \beta(v))$.

For every $v \in V$, define $f_v^\alpha(x)\colon = \sum_{i\in\alpha(v)} \widehat{f_v}(i)x_i$, $f_v^\beta(x)\colon = \sum_{i\in\beta(v)} \widehat{f_v}(i)x_i$, $f_v^\gamma(x)\colon = \sum_{i\in\gamma(v)} \widehat{f_v}(i)x_i$. So that $f_v^\alpha + f_v^\beta + f_v^\gamma = f_v^l$.

The assignment for $\mathscr{L}$. We view the sets $\{\alpha(v) \cup \beta(v)\}_{v\in V}$ as a multi-assignment for $\mathscr{L}$, namely an assignment of several values for each $v \in V$. We say that an edge $e = (u,w) \in E$ is satisfied by the multi-assignment if there exists an $i \in \alpha(u) \cup \beta(u)$ such that $\pi_e(i) \in \alpha(w) \cup \beta(w)$. We define $S \subseteq E$ to be the set of edges that are satisfied by the multi-assignment. Note that S is a symmetric set.

We use the multi-assignment to choose a *random* assignment $\mathscr{A}$ for $\mathscr{L}$, choosing $\mathscr{A}(v)$ to be a uniformly distributed element in $\alpha(v) \cup \beta(v)$. In this case, each edge $e \in S$ is satisfied by $\mathscr{A}$ with probability at least $\left(\frac{1}{a+b}\right)^2$, and therefore

$$\frac{|S|}{|E|} \le (a+b)^2 \cdot \mathrm{Opt}(\mathscr{L}). \tag{27}$$

Bounding A'(F). We are now ready to bound $A'(F)$. We first bound it by a sum of several terms, and then bound each term separately.

$$\begin{aligned}
A'(F) \;&\le\; A_{outer}(F) = \mathbb{E}_{e=(u,w)\in E}\left[\sum_{i\in[r]} \widehat{f_u}(i)\widehat{f_w}(\pi_e(i))\right] \\
&=\; \mathbb{E}_{e=(u,w)\in E}\left[\sum_{\substack{i\in\alpha(u)\cup\beta(u)\\ \pi_e(i)\in\alpha(w)\cup\beta(w)}} \widehat{f_u}(i)\widehat{f_w}(\pi_e(i))\right] && (28)\\
&+\; \mathbb{E}_{e=(u,w)\in E}\left[\sum_{\substack{i\in\alpha(u)\\ \pi_e(i)\in\gamma(w)}} \widehat{f_u}(i)\widehat{f_w}(\pi_e(i))\right] + \mathbb{E}_{e=(u,w)\in E}\left[\sum_{\substack{i\in\gamma(u)\\ \pi_e(i)\in\alpha(w)}} \widehat{f_u}(i)\widehat{f_w}(\pi_e(i))\right] && (29)\\
&+\; \mathbb{E}_{e=(u,w)\in E}\left[\sum_{\substack{i\in\beta(u)\\ \pi_e(i)\in\gamma(w)}} \widehat{f_u}(i)\widehat{f_w}(\pi_e(i))\right] + \mathbb{E}_{e=(u,w)\in E}\left[\sum_{\substack{i\in\gamma(u)\\ \pi_e(i)\in\alpha(w)}} \widehat{f_u}(i)\widehat{f_w}(\pi_e(i))\right]. && (30)
\end{aligned}$$

[4]To be precise, we should have take the size of $\alpha(v)$ and $\beta(v)$ to be $\lceil a\rceil$ and $\lceil b\rceil$ respectively. However for simplicity we disregard this minor issue.

We use the following lemmas to bound the above terms.

Lemma 4.9. $(28) \leq \big((a+b)^2 \cdot \mathrm{Opt}(\mathscr{L})\big)^{(p-2)/p}$.

Lemma 4.10. $(29) \leq 2 \cdot \sqrt{\dfrac{a}{b}} = \dfrac{2\varepsilon}{\gamma_p^2}$.

Lemma 4.11. $(30) \leq \dfrac{1+3/|V|}{(1-\varepsilon)\gamma_p^2}$.

Using the bounds from Lemma 4.9, Lemma 4.10, and Lemma 4.11, and substituting the values of a and b, we get (25) and establish the soundness property of A' (and assuming (23), also the soundness of A). Let us now prove these three lemmas.

Proof of Lemma 4.9. To prove this lemma we need the following claim.

Claim 4.12. *Let $S \subseteq E$ be a symmetric subset, and let $\{t_v\}_{v\in V}$ be real numbers. Then*

$$\mathbb{E}_{e=(u,w)\in E}\left[t_u t_w \cdot \mathbf{1}_S(e)\right] \leq \left(\frac{|S|}{|E|}\right)^{(p-2)/p} \cdot \left[\mathbb{E}_{v\sim\mu}|t_v|^p\right]^{2/p}, \tag{31}$$

where $\mathbf{1}_S$ denotes the indicator of the set S.

Proof. For every $v \in V$, let $s(v) = \frac{|(\{v\}\times V)\cap S|}{d(v)}$. Then by the definition of μ,

$$\mathbb{E}_{v\sim\mu}\left[s(v)\right] = \frac{1}{|E|} \cdot \sum_{v\in V} |(\{v\}\times V)\cap S| = \frac{|S|}{|E|}. \tag{32}$$

Now by Claim 4.5,

$$\begin{aligned}
\mathbb{E}_{e=(u,w)\in E}\left[t_u t_w \cdot \mathbf{1}_S(e)\right] &= \frac{1}{|E|} \cdot \sum_{(u,w)\in S} t_u t_w \\
&\leq \frac{1}{|E|} \cdot \sum_{(u,w)\in S} |t_u|^2 \\
&= \frac{1}{|E|} \cdot \sum_{v\in V} d(v) \cdot s(v) \cdot |t_v|^2 \\
&= \mathbb{E}_{v\sim\mu}\left[s(v) \cdot |t_v|^2\right] \\
&\leq \left[\mathbb{E}_{v\sim\mu} s(v)^{p/(p-2)}\right]^{(p-2)/p} \cdot \left[\mathbb{E}_{v\sim\mu}|t_v|^p\right]^{2/p} \\
&\leq \left[\mathbb{E}_{v\sim\mu} s(v)\right]^{(p-2)/p} \cdot \left[\mathbb{E}_{v\sim\mu}|t_v|^p\right]^{2/p} \\
&\overset{(32)}{=} \left(\frac{|S|}{|E|}\right)^{(p-2)/p} \cdot \left[\mathbb{E}_{v\sim\mu}|t_v|^p\right]^{2/p},
\end{aligned}$$

and the proof of Claim 4.12 is complete. ∎

We are now ready to bound (28). By our definition of S and by the Cauchy-Schwarz inequality we have that

$$\begin{aligned}
(28) &\le \mathbb{E}_{e=(u,w)\in E}\left[\mathbf{1}_S(e)\cdot \|f_u\|_2 \cdot \|f_w\|_2\right] \\
&\le \left(\frac{|S|}{|E|}\right)^{(p-2)/p} \cdot \left[\mathbb{E}_{v\sim\mu}\|f_v\|_2^p\right]^{2/p} && \text{(by Claim 4.12)} \\
&\le \left((a+b)^2 \cdot \mathrm{Opt}(\mathscr{L})\right)^{(p-2)/p} \cdot \left[\mathbb{E}_{v\sim\mu}\|f_v\|_p^p\right]^{2/p} && \text{(by (27))} \\
&\le \left((a+b)^2 \cdot \mathrm{Opt}(\mathscr{L})\right)^{(p-2)/p}, && \text{(by (22))}
\end{aligned}$$

as claimed. □

Proof of Lemma 4.10. By the symmetry of E, the two terms in (29) are equal. Using the Cauchy-Schwarz inequality we thus have that

$$(29) = 2\mathbb{E}_{e=(u,w)\in E}\left[\sum_{\substack{i\in\alpha(u)\\ \pi_e(i)\in\gamma(w)}} \widehat{f_u}(i)\widehat{f_w}(\pi_e(i))\right] \le 2\mathbb{E}_{e=(u,w)\in E}\left[\|f_u\|_2 \cdot \left(\sum_{\substack{i\in\alpha(u)\\ \pi_e(i)\in\gamma(w)}} \widehat{f_w}(\pi_e(i))^2\right)^{1/2}\right]. \quad (33)$$

We now consider the sum inside the square-root in (33). From the definition of $\gamma(w)$, we know that whenever $\pi_e(i) \in \gamma(w)$,

$$\widehat{f_w}(\pi_e(i))^2 \le \frac{\|f_w^l\|_2^2}{a+b} \le \frac{\|f_w\|_2^2}{a+b}$$

(this is because larger coefficients would have gone into $\alpha(w)$ or $\beta(w)$). Also, since $|\alpha(u)| = a$, the number of summands in the sum in (33) is at most a, and so for every w,

$$\left(\sum_{\substack{i\in\alpha(u)\\ \pi_e(i)\in\gamma(w)}} \widehat{f_w}(\pi_e(i))^2\right)^{1/2} \le \sqrt{\frac{a}{a+b}} \cdot \|f_w\|_2 \le \sqrt{\frac{a}{b}} \cdot \|f_w\|_2. \quad (34)$$

Substituting (34) in (33), we get

$$\begin{aligned}
(29) &\le 2 \cdot \sqrt{\frac{a}{b}} \cdot \mathbb{E}_{e=(u,w)\in E}\left[\|f_u\|_2 \cdot \|f_w\|_2\right] \\
&\le 2 \cdot \sqrt{\frac{a}{b}} \cdot \mathbb{E}_{v\sim\mu}\left[\|f_v\|_2^2\right] && \text{(using Claim 4.4)} \\
&\le 2 \cdot \sqrt{\frac{a}{b}} \cdot \|F\|_2^2 && \text{(using (22))} \\
&\le 2 \cdot \sqrt{\frac{a}{b}} = \frac{2\varepsilon}{\gamma_p^2}. && \text{(by the definition of } a \text{ and } b\text{)}
\end{aligned}$$

This completes the proof of Lemma 4.10 □

Proof of Lemma 4.11. Using Cauchy-Schwarz and (22) we have that

$$(30) \leq \mathbb{E}_{e=(u,w)\in E}\left[\sum_{\substack{i\in\beta(u)\cup\gamma(u)\\ \pi_e(i)\in\beta(w)\cup\gamma(w)}} \left|\widehat{f}_u(i)\right|\left|\widehat{f}_w(\pi_e(i))\right|\right]$$

$$\leq \mathbb{E}_{e=(u,w)\in E}\left[\left\|f_u^\beta + f_u^\gamma\right\|_2 \cdot \left\|f_w^\beta + f_w^\gamma\right\|_2\right] \leq \mathbb{E}_{v\sim\mu}\left[\left\|f_v^\beta + f_v^\gamma\right\|_2^2\right]. \tag{35}$$

For every $v \in V$ we would now like to bound $\left\|f_v^\beta + f_v^\gamma\right\|_2^2$ by considering two cases. If

$$\left\|f_v^\beta + f_v^\gamma\right\|_2^2 \geq \frac{1}{\gamma_p^2} \cdot \left\|f_v^l\right\|_2^2 \tag{36}$$

then by the choice of the parameter a and the definition of $\beta(v)$ and $\gamma(v)$, we have that all of the squares of the Fourier coefficients of $f_v^\beta + f_v^\gamma$ are smaller than $\frac{1}{a} \cdot \left\|f_v^l\right\|_2^2 = \frac{\phi(p,\varepsilon)}{\gamma_p^2} \cdot \left\|f_v^l\right\|_2^2 \leq \phi(p,\varepsilon) \cdot \left\|f_v^\beta + f_v^\gamma\right\|_2^2$. So the definition of $\phi(p,\varepsilon)$, together with the fact that f_v^α is a symmetric random variable which is independent of the random variable $f_v^\beta + f_v^\gamma$, we deduce that

$$\left\|f_v^l\right\|_p^2 = \left\|f_v^\alpha + f_v^\beta + f_v^\gamma\right\|_p^2 \geq \left\|f_v^\beta + f_v^\gamma\right\|_p^2 \geq (1-\varepsilon)\gamma_p^2 \cdot \left\|f_v^\beta + f_v^\gamma\right\|_2^2. \tag{37}$$

On the other hand, if (36) does not hold, then

$$\left\|f_v^\beta + f_v^\gamma\right\|_2^2 \leq \frac{1}{\gamma_p^2} \cdot \left\|f_v^l\right\|_p^2 \leq \frac{1}{(1-\varepsilon)\gamma_p^2} \cdot \left\|f_v^l\right\|_p^2, \tag{38}$$

and thus (38) hold in both cases. Continuing (35), we therefore have that

$$(30) \leq \frac{1}{(1-\varepsilon)\gamma_p^2} \cdot \mathbb{E}_{v\sim\mu}\left[\left\|f_v^l\right\|_p^2\right] \leq \frac{1}{(1-\varepsilon)\gamma_p^2} \cdot \left[\mathbb{E}_{v\sim\mu}\left\|f_v^l\right\|_p^p\right]^{2/p} \leq \frac{1+3/|V|}{(1-\varepsilon)\gamma_p^2},$$

where in the last inequality we used Claim 4.8. This completes the proof of Lemma 4.11. □

4.4 Removing diagonal entries

To complete the proof of the Theorem 4.2, it remain to define the actual form A, which will have no diagonal entries (that is, no non-multilinear terms), and to prove that it satisfies (23).

Identifying the source of diagonal entries. Let us begin by identifying the source of diagonal entries in A'. First, note that all the terms in A_{outer} are multilinear. This is true since it contains sums of terms of the form $\widehat{f}_u(i_1) \cdot \widehat{f}_w(i_2)$, where $(u,w) \in E$. Since by definition there are no loops in E, $\widehat{f}_u(i_1)$ and $\widehat{f}_w(i_2)$ are linear combinations of *disjoint* coordinates in F, and thus their product is multilinear in the coordinates of F.

As for A_{inner}, it is a weighted sum of terms of the form

$$A_v(F) = -B \cdot \sum_{\substack{S\subseteq[r]\\ |S|\neq 1}} \widehat{f}_v(S)^2,$$

where from the definition of f_v we have that

$$\widehat{f}_v(S)^2 = \left(\mathbb{E}_{j\in[D\cdot d(v)]}\left[\widehat{f}_v^j(S)\right]\right)^2 = \frac{1}{D^2\cdot d(v)^2}\cdot\sum_{j_1,j_2\in[D\cdot d(v)]}\widehat{f}_v^{j_1}(S)\widehat{f}_v^{j_2}(S).$$

Here the only terms that may contribute diagonal entries are those for which $j_1 = j_2$, which are of the form $\widehat{f}_v^j(S)^2$.

Defining A. Let us therefore define for every $v \in V$

$$A_v^*(F)\colon = -B\cdot\sum_{\substack{S\subseteq[r]\\|S|\neq 1}}\widehat{f}_v(S)^2 + \frac{B}{D^2\cdot d(v)^2}\cdot\sum_{j\in[D\cdot d(v)]}\sum_{\substack{S\subseteq[r]\\|S|\neq 1}}\widehat{f}_v^j(S)^2,$$

and set

$$A_{inner}^*(F)\colon = \mathbb{E}_{v\sim\mu}\left[A_v^*(F)\right] \qquad \text{and} \qquad A\colon = A_{inner}^* + A_{outer}.$$

Now A does not have nonzero diagonal entries. Let us verify that it satisfies (23). For every F such that $\|F\|_p \le 1$,

$$\begin{aligned}
\left|A'(F) - A(F)\right| &= |A_{inner}(F) - A_{inner}^*(F)| \\
&= \mathbb{E}_{v\sim\mu}\left[\frac{B}{D^2\cdot d(v)^2}\cdot\sum_{j\in[D\cdot d(v)]}\sum_{\substack{S\subseteq[r]\\|S|\neq 1}}\widehat{f}_v^j(S)^2\right] \\
&\le \frac{B}{D}\cdot\mathbb{E}_{v\sim\mu}\mathbb{E}_{j\in[D\cdot d(v)]}\left[\sum_{\substack{S\subseteq[r]\\|S|\neq 1}}\widehat{f}_v^j(S)^2\right]
\end{aligned}$$

(dropping a factor of $1/d(v)$ inside the expectation)

$$\le \frac{B}{D}\cdot\mathbb{E}_{v\sim\mu}\mathbb{E}_{j\in[D\cdot d(v)]}\left\|f_v^j\right\|_2^2 \le \frac{B}{D} = \frac{1}{|V|},$$

where the last inequality follows from (21), and the last equality is from the choice of the parameters B and D. This gives (23), and concludes the proof of Theorem 4.2.

5 Discussion and open problems

Several open problems arise naturally from our work. We list some of them below.

- Both our algorithm, and our hardness result, do not yield anything non-trivial when $1 \le p < 2$. It would be interesting to understand QM(p) in this case as well. When $p = 1$ it is easy to see that $\mathrm{Opt}_1(A)$ is up to a factor of 2 the same as $\max_{i,j\in[n]}|a_{ij}|$. Thus there is a trivial factor 2 approximation algorithm for QM(1). We have recently learned from Elad Hazan and Nimrod Megiddo (personal communication) that they obtained a PTAS for QM(1).

- We did not try to prove an integrality gap lower bound for the convex program that we used in Section 3, though we believe that a matching integrality gap should follow from an adaptation of the method used in [3].

- Is γ_p^2 the true hardness threshold for $\mathrm{QM}(p)$ for every fixed $p > 2$ (i.e. not only asymptotically as $p \to \infty$)?

- It would be interesting to study the complexity of correlated quadratic programs on more general bodies $K \subseteq \mathbb{R}^n$. Although our methods give non-trivial results in other cases, we did not pursue this research direction. In particular, if $K = \{x \in \mathbb{R}^n : \sum_{i=1}^n v(|x_i|) \leq 1\}$, where $v : \mathbb{R}_+ \to \mathbb{R}_+$ is convex (and computable), i.e. in the case of general one body potentials, a tempting conjecture would be that if $v\left(\sqrt{t}\right)$ is convex then the hardness threshold is $(\mathbb{E}v(g))^2$, where g is a standard Gaussian random variable.

- Is it possible to remove the UGC assumption in Theorem 1.1? Is it possible to prove a standard NP-hardness result in this context? Perhaps the complexity assumptions that appear in [5] can be used here as well.

- Is there a sharp hardness threshold in the Ising case (i.e. the L_∞ problem in (1))? Currently the gap between the known upper and lower bounds in this case is large, and it would be interesting to get sharp results as in the case of finite p.

Acknowledgments

We thank Gérard Ben-Arous for helpful discussions on spin glasses.

References

[1] D. Achlioptas, A. Naor, and Y. Peres. Rigorous location of phase transitions in hard optimization problems. *Nature*, 435:759–764, 2005.

[2] A. Alon and E. Berger. The Grothendieck constant of random and pseudo-random graphs. To appear in Discrete Optimization.

[3] N. Alon, K. Makarychev, Y. Makarychev, and A. Naor. Quadratic forms on graphs. *Invent. Math.*, 163(3):499–522, 2006.

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

[5] S. Arora, E. Berger, G. Kindler, E. Hazan, and S. Safra. On non-approximability for quadratic programs. In *46th Annual Symposium on Foundations of Computer Science*, pages 206–215. IEEE Computer Society, 2005.

[6] E. Artin. *The gamma function.* Translated by Michael Butler. Athena Series: Selected Topics in Mathematics. Holt, Rinehart and Winston, New York, 1964.

[7] C. P. Bachas. Computer-intractability of the frustration model of a spin glass. *J. Phys. A: Math. Gen.*, 17:709–712, 1984.

[8] N. Bansal, S. Bravyi, and B. M. Terhal. A classical approximation scheme for the ground-state energy of Ising spin Hamiltonians on planar graphs. Preprint, 2007. Available at `http://arxiv.org/abs/0705.1115`.

[9] F. Barahona. On the computational complexity of Ising spin glass models. *J. Phys. A: Math. Gen.*, 15:3241–3253, 1981.

[10] M. Bellare and P. Rogaway. The complexity of approximating a nonlinear program. *Math. Programming*, 69(3, Ser. A):429–441, 1995.

[11] G. Ben Arous, A. Dembo, and A. Guionnet. Aging of spherical spin glasses. *Probab. Theory Related Fields*, 120(1):1–67, 2001.

[12] L. Bieche, R. Maynard, R. Rammal, and J. P. Uhry. On the ground states of the frustration model of a spin glass by a matching method of graph theory. *J. Phys. A: Math. Gen.*, 13:2553–2576, 1980.

[13] A. Brieden. Geometric optimization problems likely not contained in $\mathbb{APX}$. *Discrete Comput. Geom.*, 28(2):201–209, 2002.

[14] M. Charikar, K. Makarychev, and Y. Makarychev. Near-optimal algorithms for unique games (extended abstract). In *STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pages 205–214, New York, 2006. ACM.

[15] M. Charikar and A. Wirth. Maximizing quadratic programs: extending Grothendieck's inequality. In *45th Annual Symposium on Foundations of Computer Science*, pages 54–60. IEEE Computer Society, 2004.

[16] E. de Klerk, M. Laurent, and P. A. Parrilo. A PTAS for the minimization of polynomials of fixed degree over the simplex. *Theoret. Comput. Sci.*, 361(2-3):210–225, 2006.

[17] U. Feige and L. Lovász. Two-prover one-round proof systems, their power and their problems. In *STOC'92: Proceedings of the 24th Annual ACM Symposium on Theory of Computing*, pages 733–744, New York, 1992. ACM.

[18] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, second edition, 1993.

[19] P. Hall. *Rates of convergence in the central limit theorem*, volume 62 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, Mass., 1982.

[20] E. Halperin and E. Hazan. HAPLOFREQ—estimating haplotype frequencies efficiently. *J. Comput. Biol.*, 13(2):481–500 (electronic), 2006.

[21] W. B. Johnson and J. Lindenstrauss. Basic concepts in the geometry of Banach spaces. In *Handbook of the geometry of Banach spaces, Vol. I*, pages 1–84. North-Holland, Amsterdam, 2001.

[22] S. Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, pages 767–775 (electronic), New York, 2002. ACM.

[23] S. Khot and R. O'Donnell. SDP gaps and UGC-hardness for MAXCUTGAIN. In *47th Annual Symposium on Foundations of Computer Science*, pages 217–226. IEEE Computer Society, 2006.

[24] J. Machta. The computational complexity of pattern formation. *J. Stat. Phys.*, 70:949–966, 1993.

[25] A. Megretski. Relaxation of quadratic programs in operator theory and system analysis. *Systems, Approximation, Singular Integral Operators, and Related Topics (Bordeaux, 2000)*, (3):365–392, 2001.

[26] A. Megretski. Relaxations of quadratic programs in operator theory and system analysis. In *Systems, approximation, singular integral operators, and related topics (Bordeaux, 2000)*, volume 129 of *Oper. Theory Adv. Appl.*, pages 365–392. Birkhäuser, Basel, 2001.

[27] 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.

[28] A. Nemirovski, C. Roos, and T. Terlaky. On maximization of quadratic form over intersection of ellipsoids with common center. *Mathematical Programming*, 86(3):463–473, 1999.

[29] Y. Nesterov. Global quadratic optimization via conic relaxation. *Working paper CORE*, 1998.

[30] H. Sompolinsky and A. Zippelius. Dynamic theory of the spin-glass phase. *Phys. Rev. Lett.*, 47(5):359–362, 1981.

[31] T.S.Motzkin and E. Straus. Maxima for graphs and a new proof of a theorem of turan. *Canadian Journal of Mathematics*, 17:533–540, 1965.

[32] P. van Beek. An application of Fourier methods to the problem of sharpening the Berry-Esseen inequality. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 23:187–196, 1972.

[33] S. Wolfram. Undecidability and intractability in theoretical physics. *Phys. Rev. Lett.*, 54:735–738, 1985.