Linear equations modulo 2 and the L_1 diameter of convex bodies

Subhash Khot*
Georgia Institute of Technology
khot@cc.gatech.edu

Assaf Naor†
Courant Institute of Mathematical Sciences
naor@cims.nyu.edu

Abstract

We design a randomized polynomial time algorithm which, given a 3-tensor of real numbers $A = \{a_{ijk}\}_{i,j,k=1}^n$ such that for all $i, j, k \in \{1, \ldots, n\}$ we have $a_{ijk} = a_{ikj} = a_{kji} = a_{jik} = a_{kij} = a_{jki}$ and $a_{iik} = a_{ijj} = a_{iji} = 0$, computes a number $\mathrm{Alg}(A)$ which satisfies with probability at least $\frac{1}{2}$,

$$\Omega\left(\sqrt{\frac{\log n}{n}}\right) \cdot \max_{x\in\{-1,1\}^n} \sum_{i,j,k=1}^n a_{ijk}x_ix_jx_k \le \mathrm{Alg}(A) \le \max_{x\in\{-1,1\}^n} \sum_{i,j,k=1}^n a_{ijk}x_ix_jx_k.$$

On the other hand, we show via a simple reduction from a result of Håstad and Venkatesh [22] that under the assumption $NP \nsubseteq DTIME\left(n^{(\log n)^{O(1)}}\right)$, for every $\varepsilon > 0$ there is no algorithm that approximates $\max_{x\in\{-1,1\}^n} \sum_{i,j,k=1}^n a_{ijk}x_ix_jx_k$ within a factor of $2^{(\log n)^{1-\varepsilon}}$ in time $2^{(\log n)^{O(1)}}$.

Our algorithm is based on a reduction to the problem of computing the diameter of a convex body in $\mathbb{R}^n$ with respect to the L_1 norm. We show that it is possible to do so up to a multiplicative error of $O\left(\sqrt{\frac{n}{\log n}}\right)$, while no randomized polynomial time algorithm can achieve accuracy $o\left(\sqrt{\frac{n}{\log n}}\right)$. This resolves a question posed by Brieden, Gritzmann, Kannan, Klee, Lovász and Simonovits in [10].

We apply our new algorithm to improve the algorithm of Håstad and Venkatesh [22] for the Max-E3-Lin-2 problem. Given an over-determined system $\mathcal{E}$ of N linear equations modulo 2 in $n \le N$ Boolean variables, such that in each equation appear only three distinct variables, the goal is to approximate in polynomial time the maximum number of satisfiable equations in $\mathcal{E}$ minus $\frac{N}{2}$ (i.e. we subtract the expected number of satisfied equations in a random assignment). Håstad and Venkatesh [22] obtained an algorithm which approximates this value up to a factor of $O\left(\sqrt{N}\right)$. We obtain a $O\left(\sqrt{\frac{n}{\log n}}\right)$ approximation algorithm. By relating this problem to the refutation problem for random $3-CNF$ formulas we give evidence that obtaining a significant improvement over this approximation factor is likely to be difficult.

*Research supported in part by NSF CARREER award CCF-0643626, and a Microsoft New Faculty Fellowship.
†Research supported by NSF grants CCF-0635078 and DMS-0528387.

1 Introduction

A function $f : \{-1,1\}^n \to \mathbb{R}$ has Fourier expansion $f(x) = \sum_{S\subseteq\{1,\dots,n\}} \widehat{f}(S)\prod_{i\in S} x_i$. Assume that f has a succinct representation in phase space, i.e. only polynomially many of the Fourier coefficients $\widehat{f}(S)$ are non-zero. Can we then compute in polynomial time a good approximation of the maximum of f over the discrete cube $\{-1,1\}^n$? In other words, if we are given polynomially many Fourier coefficients, is there a way to approximate $\max_{x\in\{-1,1\}^n} f(x)$ while only looking at the values of f on a tiny part of the cube? As we shall see below, under widely believed complexity assumptions the answer to this question is generally negative. But, under some additional structural information on the support of he Fourier transform it is possible to achieve this goal, and when this occurs such phenomena have powerful algorithmic applications. Currently our understanding of this fundamental problem is far from satisfactory, and the purpose of the present paper is to investigate cases which have previously eluded researchers. As a result, we uncover new connections to problems in algorithmic convex geometry and combinatorial optimization.

The Fourier maximization problem described above has been investigated extensively in the quadratic case, partly due to its connections to various graph partitioning problems. In [3] it has been shown that a classical inequality of Grothendieck can be used to give a constant factor approximation algorithm for computing the maximum of functions $f : \{-1,1\}^n \times \{-1,1\}^m \to \mathbb{R}$ which have the form $f(x,y) = \sum_{i=1}^n \sum_{j=1}^m a_{ij}x_iy_j$. This algorithm has various applications, including algorithmic versions of Szemerédi's regularity lemma. In the non-bipartite case several researchers [27, 25, 12] have discovered an algorithm which computes up to a factor $O(\log n)$ the maximum of functions $f : \{-1,1\}^n \to \mathbb{R}$ which have the form $f(x) = \sum_{i,j=1}^n a_{ij}x_ix_j$, where the matrix (a_{ij}) is assumed to be symmetric and vanish on the diagonal. This result was shown in [12] to imply the best-known approximation algorithms to graph partitioning problems such as MAXCUTGAIN and Correlation Clustering. In [2] the structure of the "Fourier support graph", i.e. the pairs $\{i,j\} \in \{1,\dots,n\}$ for which $a_{ij} \neq 0$, was taken into account. It was shown there that there exists an approximation algorithm which approximates the maximum of f up to a factor $O(\log\vartheta) = O(\log\chi)$, where ϑ is the Lovász Theta Function of the complement of the Fourier support graph, and χ is the chromatic number of this graph. We refer to [2] for more information on this topic, as well as its connection to the evaluation of ground states of spin glasses.

Negative results on the performance of the above mentioned algorithms, as well as complexity lower bounds were obtained in [3, 2, 5, 24, 1]. In particular, in was shown in [2] that the semidefinite relaxation that was used in the $O(\log n)$ algorithm discussed above had integrality gap $\Omega(\log n)$. Moreover in [5] it was shown that unless $NP \subseteq DTIME\left(n^{O((\log n)^3)}\right)$ there is no polynomial time algorithm which approximates the maximum of $\sum_{i,j=1}^n a_{ij}x_ix_j$ on $\{-1,1\}^n$ up to a factor smaller that $(\log n)^\gamma$, where γ is a universal constant. It was also shown in [5] that under the assumption of the existence of sufficiently strong PCPs it is also NP-hard to approximate this problem to within a factor of $O(\log n)$.

The motivation for the present paper is to study the case of functions whose Fourier expansion is supported on the third level. Specifically, given a 3-tensor of real numbers $A = \{a_{ijk}\}_{i,j,k=1}^n$ such that for all $i,j,k \in \{1,\dots,n\}$ we have $a_{ijk} = a_{ikj} = a_{kji} = a_{jik} = a_{kij} = a_{jki}$ and $a_{iik} = a_{ijj} = a_{iji} = 0$, we wish to approximate the maximum of the function $f : \{-1,1\}^n \to \mathbb{R}$ given by $f(x) = \sum_{i,j,k=1}^n a_{ijk}x_ix_jx_k$. Despite being a modest goal, this problem has eluded researchers for some time, as the "obvious" semidefinite programming approach that was previously applied to the quadratic case does not generalize to the degree-3 case. As we shall see below, this issue reflects a major difference from the quadratic case: Unless $NP \subseteq DTIME\left(n^{(\log n)^{O(1)}}\right)$, for every $\varepsilon > 0$ there is no algorithm that approximates the maximum of f within a factor of $2^{(\log n)^{1-\varepsilon}}$ in time $2^{(\log n)^{O(1)}}$. On the other hand, we will derive here a polynomial time algorithm which approximates the maximum of f to within a factor of $O\left(\sqrt{\frac{n}{\log n}}\right)$. This algorithm is based on a novel connection between this problem and the problem of efficient computation of the diameter of convex bodies under the ℓ_1^n norm, which is the main new insight of the present paper. We shall now describe our new approach, and its application to a fundamental problem in combinatorial optimization: The Max-E3-Lin-2 problem.

We associate to every 3-tensor A as above a convex body $K_A \subseteq \mathbb{R}^n$. The body K_A admits a polynomial time solution to the weak optimization problem for linear functionals (see [19, 18] for the relevant background on convex optimization). Moreover, we show that the ℓ_1^n diameter of K_A, i.e. $\mathrm{diam}_1(K_A) := \max_{a,b\in K_A} \|a-b\|_1$, is within a constant factor of $\max_{x_i\in\{-1,1\}} \sum_{i,j,k=1}^n a_{ijk}x_ix_jx_k$. This step is crucially based on an application of Grothendieck's inequality. We therefore reduce the level-3 Fourier maximization problem to the following question: Given a convex body $K \subseteq \mathbb{R}^n$ with a weak optimization oracle, approximate in oracle-polynomial time its ℓ_1^n diameter $\mathrm{diam}_1(K)$.

Such problems have been studied extensively in the literature, though mostly in the context of the Euclidean ℓ_2^n diameter (see for example [6, 8, 18, 10, 9, 31] and the references therein). In particular, a famous result of Bárány and Füredi states that no deterministic polynomial time algorithm can approximate the ℓ_2^n diameter of convex bodies up to a factor of $o\left(\sqrt{\frac{n}{\log n}}\right)$. In the paper [10] of Brieden, Gritzmann, Kannan, Klee, Lovász and Simonovits it is shown that unlike the case of volume computation, randomization does not help when it comes to approximating the Euclidean diameter of convex bodies: The same lower bound holds also for the accuracy of randomized oracle-polynomial time algorithms. The paper [10] also studies the case of the ℓ_1^n diameter, or more generally the ℓ_p^n diameter, i.e. $\mathrm{diam}_p(K) := \max_{a,b\in K_A} \|a-b\|_p$. It is shown there that there is an oracle polynomial time algorithm which approximates $\mathrm{diam}_1(K)$ to within a factor of $O\left(\sqrt{n}\right)$, and no polynomial time algorithm can achieve accuracy better than $O\left(\frac{\sqrt{n}}{\log n}\right)$. When $1 < p \le 2$ it is shown in [10] that $\mathrm{diam}_p(K)$ can be approximated within a factor $O\left(\frac{\sqrt{n}}{(\log n)^{(p-1)/p}}\right)$, and no polynomial time algorithm can achieve accuracy better than $O_p\left(\sqrt{\frac{n}{\log n}}\right)$. These bounds coincide only when $p = 2$, and the question of closing the gap in the remaining cases was raised in [10] (see also [9]). Here we resolve this problem by showing that the accuracy threshold for randomized oracle-polynomial algorithms that compute $\mathrm{diam}_p(K)$ is $\Theta\left(\sqrt{\frac{n}{\log n}}\right)$ for all $1 \le p \le 2$. Our improved accuracy lower bound when $p = 1$ is a slight variant of the argument in [10]. The main issue is obtaining an improved approximation algorithm—our approach is different from the polyhedral approximation of the ℓ_p^n ball that was used in [10] (though we believe that the construction of [10] is of independent interest).

We apply the results described above to obtain a significant improvement to the Max-E3-Lin-2 algorithm of Håstad and Venkatesh [22]. This fundamental problem is described as follows. Consider a system $\mathcal{E}$ of N linear equations modulo 2 in n Boolean variables $z_1, \ldots, z_n$, such that in each equation appear only three distinct variables. We assume throughout that $N \ge n$ (thus avoiding degenerate cases). Let $\mathrm{MAXSAT}(\mathcal{E})$ be the maximum number of equations in $\mathcal{E}$ that can be satisfied simultaneously. A random assignment of these variables satisfies in expectation $\frac{N}{2}$ equations, so in the Max-E3-Lin-2 problem it is natural to ask for an approximation algorithm to $\mathrm{MAXSAT}(\mathcal{E}) - \frac{N}{2}$. This problem was studied extensively by Håstad and Venkatesh in [22], where the best known upper and lower bounds were obtained. In particular, using the powerful methods of Håstad [21] they show that unless $NP \subseteq DTIME\left(n^{(\log n)^{O(1)}}\right)$, for every $\varepsilon > 0$ there is no algorithm that approximates $\mathrm{MAXSAT}(\mathcal{E}) - \frac{N}{2}$ within a factor of $2^{(\log n)^{1-\varepsilon}}$ in time $2^{(\log n)^{O(1)}}$. Moreover, they design a randomized polynomial time algorithm which approximates $\mathrm{MAXSAT}(\mathcal{E}) - \frac{N}{2}$ to within a factor of $O\left(\sqrt{N}\right)$.

Let $\mathcal{E}$ be a system of linear equations as above. Write $a_{ijk}(\mathcal{E}) = 1$ if the equation $z_i + z_j + z_k = 0$ is in the system $\mathcal{E}$. Similarly write $a_{ijk}(\mathcal{E}) = -1$ if the equation $z_i + z_j + z_k = 1$ is in $\mathcal{E}$. Finally, write $a_{ijk}(\mathcal{E}) = 0$ if no equation in $\mathcal{E}$ corresponds to $z_i + z_j + z_k$. Assume that the assignment $(z_1, \ldots, z_k)$ satisfies m of the equations in $\mathcal{E}$. Then $\sum_{i,j,k=1}^n a_{ijk}(\mathcal{E})(-1)^{z_i+z_j+z_k} = m - (N - m) = 2\left(m - \frac{N}{2}\right)$. It follows that

$$\max_{x_i \in \{-1,1\}} \sum_{i,j,k=1}^n a_{ijk}(\mathcal{E}) x_i x_j x_k = \max_{z_i \in \{0,1\}} \sum_{i,j,k=1}^n a_{ijk}(\mathcal{E})(-1)^{z_i+z_j+z_k} = 2\left(\mathrm{MAXSAT}(\mathcal{E}) - \frac{N}{2}\right).$$

Thus our algorithm yields a $O\left(\sqrt{\frac{n}{\log n}}\right)$ approximation to the Max-E3-Lin-2 problem. Note that when $N = \Theta(n)$ our improvement over the Håstad-Venkatesh algorithm is only logarithmic, but typically N can be as large as $\Theta\left(N^3\right)$. The above reasoning also allows us to apply the Håstad-Venkatesh hardness result for Max-E3-Lin-2 that was described above to the level-3 Fourier maximization problem. In particular it follows that this problem is computationally much harder than the quadratic case, in which a $O(\log n)$ approximation is possible. Finally, our reasoning comes full circle to shed light on the problem of approximating the ℓ_1^n diameter $\mathrm{diam}_1(K)$. While the proof in [10] is essentially an "entropy argument" showing that there are simply too many convex bodies to allow an approximation factor better than $O\left(\sqrt{n}\right)$, our reduction produces a concrete family of convex bodies for which computing the ℓ_1^n diameter within a factor of is $2^{(\log n)^{1-\varepsilon}}$ is hard.

2 A new algorithm for Max-E3-Lin-2

As described in the reduction that was presented in the introduction, our new algorithm for Max-E3-Lin-2 will follow from the more general algorithm for approximating the maximum of functions whose Fourier transform is supported on subsets of size 3. So, from now on let $\{a_{ijk}\}_{i,j,k=1}^n$ be real numbers such that for all $i, j, k \in \{1, \ldots, n\}$ we have $a_{ijk} = a_{ikj} = a_{kji} = a_{jik} = a_{kij} = a_{jki}$ and $a_{iik} = a_{ijj} = a_{iji} = 0$. Our first lemma reduces the problem of maximizing $\sum_{i,j,k=1}^n a_{ijk} x_i x_j x_k$ to the analogous tripartite case. Note that such a result *is false* in the quadratic case. Indeed, Theorem 3.5 in [2] implies that the gap between $\max_{x_i \in \{-1,1\}} \sum_{i,j=1}^n a_{ij} x_i x_j$ and $\max_{x_i, y_j \in \{-1,1\}} \sum_{i,j=1}^n a_{ij} x_i y_j$ can be as large as $\Omega\left(\frac{n}{\log n}\right)$. The key "trick" which allows us to prove that this cannot happen in the level-3 case is identity (1) below.

Lemma 2.1. *The following inequalities hold true:*

$$\frac{1}{10} \max_{x_i, y_j, z_k \in \{-1,1\}} \sum_{i,j,k=1}^n a_{ijk} x_i y_j z_k \le \max_{x_i \in \{-1,1\}} \sum_{i,j,k=1}^n a_{ijk} x_i x_j x_k \le \max_{x_i, y_j, z_k \in \{-1,1\}} \sum_{i,j,k=1}^n a_{ijk} x_i y_j z_k.$$

Proof. Define

$$M = \max_{x_i, y_j, z_k \in \{-1,1\}} \sum_{i,j,k=1}^n a_{ijk} x_i y_j z_k,$$

and

$$m = \max_{x_i \in \{-1,1\}} \sum_{i,j,k=1}^n a_{ijk} x_i x_j x_k.$$

Clearly $m \le M$, so we need to show that that $M \le 10m$. To see this observe first of all that $\sum_{i,j,k=1}^n a_{ijk} x_i x_j x_k$ is linear in x_i for each i. This implies that

$$m = \max_{|x_i| \le 1} \sum_{i,j,k=1}^n a_{ijk} x_i x_j x_k.$$

Moreover, since $\sum_{i,j,k=1}^n a_{ijk}x_ix_jx_k$ changes sign if we replace x_i by $-x_i$ for each i, we see that

$$m = \max_{|x_i|\le 1}\left|\sum_{i,j,k=1}^n a_{ijk}x_ix_jx_k\right|.$$

Now, for each $i, j, k \in \{1,\dots,n\}$ we have the identity

$$2x_iy_jy_k + 2x_jy_iy_k + 2x_ky_iy_j = (x_i+y_i)(x_j+y_j)(x_k+y_k) + (x_i-y_i)(x_j-y_j)(x_k-y_k) - 2x_ix_jx_k - 2y_iy_jy_k. \quad (1)$$

Multiplying this identity by a_{ijk}, summing over all $i, j, k \in \{1,\dots,n\}$, and using the symmetries of the coefficients a_{ijk}, we see that

$$6\sum_{i,j,k=1}^n a_{ijk}x_iy_jy_k = 8\sum_{i,j,k=1}^n a_{ijk}\frac{x_i+y_i}{2}\cdot\frac{x_j+y_j}{2}\cdot\frac{x_k+y_k}{2} + 8\sum_{i,j,k=1}^n a_{ijk}\frac{x_i-y_i}{2}\cdot\frac{x_j-y_j}{2}\cdot\frac{x_k-y_k}{2} - 2\sum_{i,j,k=1}^n a_{ijk}x_ix_jx_k - 2\sum_{i,j,k=1}^n a_{ijk}y_iy_jy_k.$$

It follows that

$$M' := \max_{x_i,y_j\in\{-1,1\}}\left|\sum_{i,j,k=1}^n a_{ijk}x_iy_jy_k\right| \le \frac{20}{6}m = \frac{10}{3}m.$$

As before, because $\sum_{i,j,k=1}^n a_{ijk}x_iy_jy_k$ is linear in each of the variables x_i and y_j, we have the identity

$$M' = \max_{|x_i|,|y_j|\le 1}\left|\sum_{i,j,k=1}^n a_{ijk}x_iy_jy_k\right|.$$

Now, consider the identity

$$y_jz_k + y_kz_j = (y_j+z_j)(y_k+z_k) - y_jy_k - z_jz_k.$$

Multiplying by $a_{ijk}x_i$, and summing up, we get the identity

$$2\sum_{i,j,k=1}^n a_{ijk}x_iy_jz_k = 4\sum_{i,j,k=1}^n a_{ijk}x_i\cdot\frac{y_j+z_j}{2}\cdot\frac{y_k+z_k}{2} - \sum_{i,j,k=1}^n a_{ijk}x_iy_jy_k - \sum_{i,j,k=1}^n a_{ijk}x_iz_jz_k \le 6M' \le 20m.$$

It follows that $M \le 10m$, as required. □

Let $(\ell_2^n)_\infty^n$ denote the space of $\vec{v} = (v_1,\dots,v_n) \in (\mathbb{R}^n)^n$, equipped with the norm

$$\left\|\vec{v}\right\|_{(\ell_2^n)_\infty^n} := \max_{1\le j\le n}\|v_j\|_2.$$

Similarly we let $(\ell_2^n)_1^n$ denote the space of $\vec{v} = (v_1,\dots,v_n) \in (\mathbb{R}^n)^n$, equipped with the norm

$$\left\|\vec{v}\right\|_{(\ell_2^n)_1^n} := \sum_{j=1}^n\|v_j\|_2.$$

Any $n \times n$ matrix $B = (b_{ij}) \in M_n(\mathbb{R})$ can be tensorized with the identity to yield an operator $B \otimes I : (\ell_2^n)_\infty^n \to (\ell_2^n)_1^n$ given by

$$((B \otimes I)\vec{v})_i := \sum_{j=1}^n b_{ij} v_j.$$

The operator norm of $B \otimes I$ is given by

$$\begin{aligned}
\|B \otimes I\|_{(\ell_2^n)_\infty^n \to (\ell_2^n)_1^n} &= \max\left\{ \sum_{i=1}^n \left\| \sum_{j=1}^n b_{ij} v_j \right\|_2 : \max_{1 \le j \le n} \|v_j\|_2 \le 1 \right\} \\
&= \max\left\{ \sum_{i=1}^n \left\langle \sum_{j=1}^n b_{ij} v_j, u_i \right\rangle : \max_{1 \le j \le n} \|v_j\|_2 \le 1 \ \wedge \ \max_{1 \le i \le n} \|u_i\|_2 \le 1 \right\} \\
&= \max_{\|u_i\|_2, \|v_j\|_2 \le 1} \sum_{i,j=1}^n b_{ij} \left\langle u_i, v_j \right\rangle. \qquad (2)
\end{aligned}$$

By Lemma 2.1 our goal is to approximate the value

$$\mathrm{Opt}(A) := \max_{x_i, y_j, z_k \in \{-1,1\}} \sum_{i,j,k=1}^n a_{ijk} x_i y_j z_k.$$

For every $x \in \{-1, 1\}^n$ define an $n \times n$ matrix $A(x) \in M_n(\mathbb{R})$ by

$$A(x)_{jk} = \sum_{i=1}^n a_{ijk} x_i.$$

Since for each $x \in \mathbb{R}^n$ we have $\|A(x) \otimes I\|_{(\ell_2^n)_\infty^n \to (\ell_2^n)_1^n} \ge \max_{|y_j|, |z_k| \le 1} \sum_{j,k=1}^n A(x)_{ij} y_j z_k$ it follows that $\mathrm{Opt}(A) \le \max_{x \in \{-1,1\}^n} \|A(x) \otimes I\|_{(\ell_2^n)_\infty^n \to (\ell_2^n)_1^n}$. On the other hand, using (2), Grothendieck's inequality (see the discussion in [3]) says that

$$\max_{x \in \{-1,1\}^n} \|A(x) \otimes I\|_{(\ell_2^n)_\infty^n \to (\ell_2^n)_1^n} \le \max_{x \in \{-1,1\}^n} K_G \max_{y, x \in \{-1,1\}^n} \sum_{j,k=1}^n A(x)_{jk} y_j z_k = K_G \cdot \mathrm{Opt}(A),$$

where $K_G \le 2$ is Grothendieck's constant. It therefore suffices to approximate the value of

$$\max_{x \in \{-1,1\}^n} \|A(x) \otimes I\|_{(\ell_2^n)_\infty^n \to (\ell_2^n)_1^n}.$$

Define a norm $\|\cdot\|_A$ on $\mathbb{R}^n$ by $\|x\|_A := \|A(x) \otimes I\|_{(\ell_2^n)_\infty^n \to (\ell_2^n)_1^n}$. Then the unit ball $B_A := \{x \in \mathbb{R}^n : \|x\|_A \le 1\}$ is a centrally symmetric convex body. Denote by $K_A = B_A^\circ$ the polar of B_A, i.e.

$$K_A = \{y \in \mathbb{R}^n : \forall x \in B_A, \langle x, y \rangle \le 1\}.$$

Then

$$\max_{x \in \{-1,1\}^n} \|A(x) \otimes I\|_{(\ell_2^n)_\infty^n \to (\ell_2^n)_1^n} = \max_{\|x\|_\infty \le 1} \|x\|_A = \max_{\|x\|_\infty \le 1} \max_{y \in K_A} \langle x, y \rangle = \max_{y \in K_A} \max_{\|x\|_\infty \le 1} \langle x, y \rangle = \max_{y \in K_A} \|y\|_1 = \frac{1}{2} \operatorname{diam}_1(K_A).$$

We have thus reduced our original problem to approximating $\operatorname{diam}_1(K_A)$ in polynomial time. Note that (2) implies that the computation of $\|x\|_A$ is a semidefinite program. Therefore by the theory of Grötschel, Lovász and Schrijver [19] it follows that linear functionals can be optimized on B_A in polynomial time. As shown in [19], this property is preserved under polarity, i.e. linear functionals can be optimized on $B_A^\circ = K_A$ in polynomial time. We have therefore reduced the problem of approximating $\max_{x_i\in\{-1,1\}} \sum_{i,j,k=1}^n a_{ijk}x_ix_jx_k$ to the problem of approximating $\mathrm{diam}_1(K)$ in oracle polynomial time, where K is a centrally symmetric convex body with a weak optimization oracle. This problem is resolved in Section 3, thus proving the following theorem, which is our main result.

Theorem 2.2. *There is a randomized polynomial time algorithm which, given a 3-tensor $A = \{a_{ijk}\}_{i,j,k=1}^n$ such that for all $i, j, k \in \{1, \dots, n\}$ we have $a_{ijk} = a_{ikj} = a_{kji} = a_{jik} = a_{kij} = a_{jki}$ and $a_{iik} = a_{ijj} = a_{iji} = 0$, computes a number* $\mathrm{Alg}(A)$ *which satisfies with probability at least $\frac{1}{2}$,*

$$\max_{x\in\{-1,1\}^n} \sum_{i,j,k=1}^n a_{ijk}x_ix_jx_k \le \mathrm{Alg}(A) \le O\left(\sqrt{\frac{n}{\log n}}\right) \max_{x\in\{-1,1\}^n} \sum_{i,j,k=1}^n a_{ijk}x_ix_jx_k.$$

3 An approximation algorithm for the L_1 diameter

The main result of this section is the following theorem, which settles a problem posed by Brieden, Gritzmann, Kannan, Klee, Lovász and Simonovits in [10].

Theorem 3.1. *Let $K \subseteq \mathbb{R}^n$ be a convex body with a weak optimization oracle. Then there exists a randomized algorithm which computes in oracle-polynomial time a number* $\mathrm{Alg}(K)$ *such that with probability at least $\frac{1}{2}$,*

$$\frac{1}{2}\sqrt{\frac{\log n}{n}} \cdot \mathrm{diam}_1(K) \le \mathrm{Alg}(K) \le \mathrm{diam}_1(K).$$

On the other hand, no randomized oracle-polynomial time algorithm can compute $\mathrm{diam}_1(K)$ *with accuracy $o\left(\sqrt{\frac{n}{\log n}}\right)$.*

Since we will be only using Theorem 3.1 when K is 0-symmetric, i.e. $K = -K$, we will prove it under this assumption. This is only for the sake of simplifying the notation—identical arguments work in the general case. Our starting point is the following distributional inequality.

Lemma 3.2. *For every $\delta \in \left(0, \frac{1}{2}\right)$ there is a constant $c(\delta) > 0$ with the following property. Fix $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ and let $\varepsilon_1, \dots, \varepsilon_n$ be i.i.d. symmetric Bernoulli random variables. Then*

$$\Pr\left(\sum_{j=1}^n a_j\varepsilon_j \ge \sqrt{\frac{\delta \log n}{n}} \cdot \|a\|_1\right) \ge \frac{c(\delta)}{n^\delta}.$$

Proof. Write $X = \sum_{j=1}^n a_j\varepsilon_j$. Assume first of all that

$$\frac{4\|a\|_2^2\sqrt{n}}{\|a\|_1^2} > \frac{1}{12n^\delta}. \tag{3}$$

The classical Paley-Zygmund inequality [28, 23, 4] states that for every $\theta \in (0, 1)$ we have

$$\Pr\left(X^2 \ge \theta\mathbb{E}X^2\right) \ge (1-\theta)^2 \cdot \frac{\left(\mathbb{E}X^2\right)^2}{\mathbb{E}X^4} \ge \frac{(1-\theta)^2}{9}, \tag{4}$$

where we used the well known (and easy) fact that $\mathbb{E}X^4 \leq 9\left(\mathbb{E}X^2\right)^2$.

The inclusion of events $\left\{X^2 \geq \theta\right\} \subseteq \left\{X \geq \sqrt{\theta}\right\} \cup \left\{-X \geq \sqrt{\theta}\right\}$, and the fact that X is symmetric, implies that $\Pr\left(X \geq \sqrt{\theta}\right) \geq \frac{1}{2}\Pr\left(X^2 \geq \theta\right)$. Since $\delta < \frac{1}{2}$ there is $n_0(\delta) \in \mathbb{N}$ such that for every $n \geq n_0(\delta)$ we have $\frac{48\delta \log n}{n^{\frac{1}{2}-\delta}} < \frac{1}{2}$. For such n we deduce that

$$\Pr\left(\sum_{j=1}^{n} a_j\varepsilon_j \geq \sqrt{\frac{\delta \log n}{n}} \cdot \|a\|_1\right) \geq \frac{1}{2}\Pr\left(X^2 \geq \frac{\delta \log n}{n}\|a\|_1^2\right) \overset{(3)}{\geq} \frac{1}{2}\Pr\left(X^2 \geq \frac{48\delta \log n}{n^{\frac{1}{2}-\delta}} \cdot \mathbb{E}X^2\right) \overset{(4)}{\geq} \frac{1}{72}.$$

Hence Lemma 3.2 holds assuming (3) and $n \geq n_0(\delta)$. By adjusting $c(\delta)$ the required result holds also when $n \leq n_0(\delta)$, so it remains to deal with the case

$$\frac{4\|a\|_2^2\sqrt{n}}{\|a\|_1^2} \leq \frac{1}{12n^{\delta}}. \tag{5}$$

Assuming (5) we define $S := \left\{j \in \{1,\ldots,n\} :\ |a_j| < \frac{2\|a\|_2^2}{\|a\|_1}\right\}$. Then

$$\|a\|_1 = \sum_{j\notin S}|a_j| + \sum_{j\in S}|a_j| \leq \frac{\|a\|_1}{2\|a\|_2^2}\sum_{j\notin S}a_j^2 + \sqrt{|S|\sum_{j\in S}a_j^2} \leq \frac{\|a\|_1}{2} + \sqrt{n\sum_{j\in S}a_j^2}.$$

Hence,

$$\sqrt{\sum_{j\in S}a_j^2} \geq \frac{\|a\|_1}{2\sqrt{n}}. \tag{6}$$

Write $Y = \sum_{j\in S} a_j\varepsilon_j$ and $Z = X - Y$. For every $t \in \mathbb{R}$ we have $\{Y \geq 2t\} \subseteq \{Y + Z \geq t\} \cup \{Y - Z \geq t\}$. Since $Y + Z$ and $Y - Z$ have the same distribution as X, it follows that $\Pr(X \geq t) \geq \frac{1}{2}\Pr(Y \geq 2t)$. Let g be a standard Gaussian random variable. By the Berry-Esseen inequality (see [20]. The constant we use below follows from [30]) we know that

$$\Pr(Y \geq 2t) = \Pr\left(\frac{Y}{\sqrt{\mathbb{E}Y^2}} \geq \frac{2t}{\sqrt{\mathbb{E}Y^2}}\right) \geq \Pr\left(\frac{Y}{\sqrt{\mathbb{E}Y^2}} \geq \frac{t\sqrt{n}}{\|a\|_1}\right) \geq \Pr\left(g \geq \frac{t\sqrt{n}}{\|a\|_1}\right) - \max_{j\in S}\frac{|a_j|}{\sqrt{\sum_{k\in S}a_k^2}}$$
$$\overset{(6)}{\geq} \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{t^2 n}{\|a\|_1^2}\right) - \frac{4\|a\|_2^2\sqrt{n}}{\|a\|_1^2}.$$

Plugging $t = \sqrt{\frac{\delta \log n}{n}} \cdot \|a\|_1$ we get that

$$\Pr\left(\sum_{j=1}^{n} a_j\varepsilon_j \geq \sqrt{\frac{\delta \log n}{n}} \cdot \|a\|_1\right) \geq \frac{1}{2}\Pr(Y \geq 2t) \geq \frac{1}{6n^{\delta}} - \frac{4\|a\|_2^2\sqrt{n}}{\|a\|_1^2} \overset{(5)}{\geq} \frac{1}{12n^{\delta}},$$

as required. $\square$

Proof of Theorem 3.1. Let $\left\{\varepsilon_{ij} : \ i \in \{1, \dots, m\}, \ j \in \{1, \dots, n\}\right\}$ be i.i.d. symmetric Bernoulli random variables. Compute the number

$$\mathrm{Alg}(K) := 2 \max_{1 \le i \le m} \max_{a \in K} \sum_{j=1}^{n} a_j \varepsilon_{ij}.$$

Then $\mathrm{Alg}(K) \le 2 \max_{a \in K} \|a\|_1 = \mathrm{diam}_1(K)$. Moreover, M can be computed using $O(m)$ oracle calls. Now, fix $a \in K$ such that $\|a\|_1 = \frac{1}{2}\,\mathrm{diam}_1(K)$. Using Lemma 3.2 we see that there exists a universal constant $c > 0$ for which

$$\begin{aligned}\Pr\left(\mathrm{Alg}(K) > \frac{1}{2}\sqrt{\frac{\log n}{n}} \cdot \mathrm{diam}_1(K)\right) &= 1 - \Pr\left(\bigcap_{i=1}^{m} \left\{\sum_{j=1}^{n} a_j \varepsilon_{ij} < \frac{1}{2}\sqrt{\frac{\log n}{n}} \cdot \|a\|_1\right\}\right) \\ &\ge 1 - \left(1 - \frac{c}{\sqrt[4]{n}}\right)^m \ge 1 - \exp\left(-\frac{cm}{\sqrt[4]{n}}\right).\end{aligned}$$

Choosing $m = \left\lceil \frac{\sqrt[4]{n}}{c} \right\rceil$ we see that with probability at least $\frac{1}{2}$,

$$\frac{1}{2}\sqrt{\frac{\log n}{n}} \cdot \mathrm{diam}_1(K) \le \mathrm{Alg}(K) \le \mathrm{diam}_1(K),$$

as required.

The algorithmic lower bound in Theorem 3.1 is essentially already contained in [10]—the authors simply overlooked there an easy stronger upper bound on the volume of polytopes inscribed in in the cube $[-1,1]^n$. In Proposition 1.10 in [10] the authors prove that for every 0-symmetric polytope $P \subseteq [-1,1]^n$ with at most $2k$ vertices, where $20\log_2\left(\frac{k}{n} + 1\right) \le n \le k$,

$$(\mathrm{vol}(P))^{1/n} \le O(1)\sqrt{1 + \log n} \cdot \sqrt{\frac{\log\left(\frac{k}{n} + 1\right)}{n}}. \tag{7}$$

The term $\sqrt{1 + \log n}$ in (7) is precisely the reason why the lower bound in [10] for the accuracy of randomized algorithms which compute $\mathrm{diam}_1(K)$ was $\Omega\left(\frac{\sqrt{n}}{\log n}\right)$ instead of $O\left(\sqrt{\frac{n}{\log n}}\right)$. This term can be removed as follows.

Let B_2^n be the standard unit Euclidean ball of ℓ_2^n. Write $P = \mathrm{conv}\{\pm v_1, \dots, \pm v_k\}$, where $v_i \in [-1,1]^n$. Then $\frac{v_i}{\sqrt{n}} \in B_2^n$, and by the results of [6, 11, 17] we deduce that

$$\left(\frac{\mathrm{vol}\left(\frac{1}{\sqrt{n}}P\right)}{\mathrm{vol}\left(B_2^n\right)}\right)^{1/n} \le O(1)\sqrt{\frac{\log\left(\frac{k}{n} + 1\right)}{n}}.$$

Since $\left(\mathrm{vol}\left(B_2^n\right)\right)^{1/n} = \Theta\left(\frac{1}{\sqrt{n}}\right)$ it follows that

$$(\mathrm{vol}(P))^{1/n} \le O(1)\sqrt{\frac{\log\left(\frac{k}{n} + 1\right)}{n}}.$$

□

3.1 The case of the L_p diameter, $1 < p < 2$

Fix $p \in (1,2)$, and define $q = \frac{p}{p-1} > 2$. Let $h_1, \dots, h_n$ be i.i.d. random variables whose density is $\frac{q}{2\Gamma(1/q)} e^{-|t|^q}$. Let H be the random vector $(h_1, \dots, h_n) \in \mathbb{R}^n$. Then the random variables $H/\|H\|_q$ and $\|H\|_q$ are independent [29] (see [7] for more information on this phenomenon). The following lemma is analogous to Lemma 3.2.

Lemma 3.3. *There exist universal constants $\delta, c_1, c_2 > 0$ such that for every $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ we have*

$$\Pr\left(\left\langle \frac{H}{\|H\|_q}, a \right\rangle \geq \sqrt{\frac{\delta \log n}{n}} \cdot \|a\|_p\right) \geq \frac{c_1}{n^{c_2}}.$$

Proof. The random variable $\|H\|_q$ has density $\frac{q}{\Gamma(n/q)} u^{n-1} e^{-u^q}$ for $u > 0$ (see for example [26]). Hence for every $t \in (0,1)$ we have

$$\mathbb{E} e^{t\|H\|_q^q} = \frac{q}{\Gamma(n/q)} \int_0^\infty e^{tu^q} \cdot u^{n-1} e^{-u^q} du = \frac{q}{\Gamma(n/q)} \int_0^\infty u^{n-1} e^{-[(1-t)^{1/q} u]^q} du = \frac{1}{(1-t)^{n/q}}.$$

Since $q \geq 2$ it follows that

$$\Pr\left(\|H\|_q \geq n^{1/q}\right) \leq e^{-n\left(1-\frac{1}{e}\right)} \mathbb{E} e^{\left(1-\frac{1}{e}\right)\|H\|_q^q} = e^{-n\left(1-\frac{1}{e}-\frac{1}{q}\right)} \leq e^{-n\left(\frac{1}{2}-\frac{1}{e}\right)} \leq e^{-n/8}. \tag{8}$$

Using the independence of $H/\|H\|_q$ and $\|H\|_q$ we deduce that

$$\begin{aligned}
\Pr\left(\left\langle \frac{H}{\|H\|_q}, a \right\rangle \geq \sqrt{\frac{\delta \log n}{n}} \cdot \|a\|_p\right) &\geq \Pr\left(\left\langle \frac{H}{\|H\|_q}, a \right\rangle \geq \sqrt{\frac{\delta \log n}{n}} \cdot \|a\|_p\right) \Pr\left(\|H\|_q \leq n^{1/q}\right) \\
&= \Pr\left(\left\langle \frac{H}{\|H\|_q}, a \right\rangle \geq \sqrt{\frac{\delta \log n}{n}} \cdot \|a\|_p \ \wedge \ \|H\|_q \leq n^{1/q}\right) \\
&\geq \Pr\left(\langle H, a \rangle \geq n^{1/q} \sqrt{\frac{\delta \log n}{n}} \cdot \|a\|_p \ \wedge \ \|H\|_q \leq n^{1/q}\right) \\
&\geq 1 - \Pr\left(\langle H, a \rangle < n^{1/q} \sqrt{\frac{\delta \log n}{n}} \cdot \|a\|_p\right) - \Pr\left(\|H\|_q > n^{1/q}\right) \\
&\overset{(8)}{\geq} \Pr\left(\langle H, a \rangle \geq n^{1/q} \sqrt{\frac{\delta \log n}{n}} \cdot \|a\|_p\right) - e^{-n/8}.
\end{aligned} \tag{9}$$

As in the proof of Lemma 3.2 we write $X = \sum_{j=1}^n a_j h_j$. Let

$$S := \left\{ j \in \{1, \dots, n\} : \ |a_j| \leq \frac{2^{\frac{1}{2-p}} \|a\|_2^{\frac{2}{2-p}}}{\|a\|_p^{\frac{p}{2-p}}} \right\}.$$

Then, using the definition of S and Hölder's inequality, we see that

$$\|a\|_p^p = \sum_{j \notin S} |a_j|^p + \sum_{j \in S} |a_j|^p \leq \frac{\|a\|_p^p}{2\|a\|_2^2} \sum_{j \notin S} a_j^2 + |S|^{\frac{2-p}{2}} \left(\sum_{j \in S} a_j^2\right)^{\frac{p}{2}} \leq \frac{\|a\|_p^p}{2} + n^{\frac{2-p}{2}} \left(\sum_{j \in S} a_j^2\right)^{\frac{p}{2}}.$$

It follows that

$$\sqrt{\sum_{j\in S} a_j^2} \geq \frac{\|a\|_p}{2^{\frac{1}{p}} n^{\frac{1}{p}-\frac{1}{2}}}. \tag{10}$$

Set $Y := \sum_{j\in S} a_j h_j$, and note that

$$\sqrt{\mathbb{E}Y^2} = \sqrt{\sum_{j\in S} a_j^2 \mathbb{E}h_j^2} = \Omega(1)\sqrt{\sum_{j\in S} a_j^2} \overset{(10)}{\geq} c\frac{\|a\|_p}{n^{\frac{1}{p}-\frac{1}{2}}}, \tag{11}$$

where $c > 0$ is a universal constant. Using the Berry-Esseen inequality as in the proof of Lemma 3.2, we see that

$$\begin{aligned}\Pr\left(\sum_{j=1}^n a_j h_j \geq n^{\frac{1}{q}}\sqrt{\frac{\delta \log n}{n}} \cdot \|a\|_p\right) &\geq \frac{1}{2}\Pr\left(\frac{Y}{\sqrt{\mathbb{E}Y^2}} \geq 2cn^{1-\frac{1}{p}}\sqrt{\frac{\delta\log n}{n}}\cdot n^{\frac{1}{p}-\frac{1}{2}}\right)\\ &\geq \frac{1}{6n^{4c^2\delta}} - \max_{j\in S}\frac{|a_j|}{\sqrt{\sum_{k\in S} a_k^2}}\cdot \mathbb{E}|h_1|^3 = \frac{1}{6n^{4c^2\delta}} - O\left(n^{\frac{1}{p}-\frac{1}{2}}\left(\frac{\|a\|_2}{\|a\|_p}\right)^{\frac{2}{2-p}}\right) \geq \frac{1}{12n^{4c^2\delta}},\end{aligned} \tag{12}$$

provided that

$$n^{\frac{1}{p}-\frac{1}{2}}\left(\frac{\|a\|_2}{\|a\|_p}\right)^{\frac{2}{2-p}} \leq \frac{\tilde{c}}{n^{4c^2\delta}}, \tag{13}$$

for some small enough constant $\tilde{c}$. But, assuming that (13) fails, and that δ is small enough and n is large enough, we may apply the Paley-Zygmund inequality to conclude that

$$\Pr\left(\sum_{j=1}^n a_j h_j \geq n^{\frac{1}{q}}\sqrt{\frac{\delta\log n}{n}}\cdot\|a\|_p\right) \geq \Pr\left(X^2 \geq \frac{\tilde{c}^{2-p}\delta\log n}{n^{2-\frac{p}{2}-\frac{1}{p}-8c^2\delta}}\cdot \mathbb{E}X^2\right) \geq \Pr\left(X^2 \geq \frac{\delta \log n}{n^{\frac{1}{2}-8c^2\delta}}\cdot\mathbb{E}X^2\right) \geq \Omega(1), \tag{14}$$

where we used the fact that $p \in (1,2)$ and the easy bound $\sqrt[4]{\mathbb{E}X^4} = O\left(\sqrt{\mathbb{E}X^2}\right)$. Combining (12) and (14) with (9) yields the required result. $\square$

Now, arguing as in the proof of Theorem 3.1, given a 0-symmetric convex body $K \subseteq \mathbb{R}^n$ with a weak optimization oracle, we select m i.i.d. copies of H, $H_1, \ldots, H_m$, and define

$$\mathrm{Alg}(K) := 2\max_{1\leq i\leq m}\max_{a\in K}\left\langle \frac{H_i}{\|H_i\|_q}, a\right\rangle.$$

Arguing as in the proof of Theorem 3.1, with Lemma 3.2 replaced by Lemma 3.3, we see that for $m = \mathrm{poly}(n)$, with constant probability

$$\Omega(1)\sqrt{\frac{\log n}{n}}\cdot \mathrm{diam}_p(K) \leq \mathrm{Alg}(K) \leq \mathrm{diam}_p(K).$$

4 Discussion and open problems

We end this paper with some remarks and directions for future research.

• We assumed throughout that $a_{iik} = a_{ijj} = a_{iji} = 0$. This restriction, which also appeared in [2] as the condition that the Fourier support graph does not have self loops, is necessary since otherwise if $P \neq NP$ then there is no polynomial time algorithm that evaluates the maximum of $\sum_{i,j,k=1}^{n} a_{ijk}x_i x_j x_k$ over $x \in \{-1, 1\}^n$ up to any factor (even one that grows with n arbitrarily fast)—see the discussion in Remark 3.2 in [2].

• It would be interesting to investigate maximization problem of $\sum_{i,j,k=1}^{n} a_{ijk}x_i x_j x_k$ in terms of the combinatorial structure of the Fourier support *hypergraph* given by $\{\{i, j, k\} : a_{ijk} \neq 0\}$. The results of [2] suggest that it might be possible to achieve a better approximation guarantee in the presence of additional structural information of this type.

• A natural question that arises from our results is to study the maximization problem for

$$\sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=k}} a_S \prod_{j \in S} x_j$$

when $k \geq 4$. Our methods do not immediately give good bounds in this case—it is quite easy to get a $O\left(\frac{n^{\frac{k}{2}-1}}{(\log n)^{\frac{k}{2}-1}}\right)$ approximation algorithm for odd k by iterating our approach, and it would be desirable to get improved bounds. Such improvements, beyond their intrinsic interest, might have implications to the problem of refutation of random $k-CNF$ formulas [16, 15, 14] (see [13] for motivation of such questions). The connection between these two problems is explained in the following theorem.

Theorem 4.1. *Suppose for every $\ell \in \{1, \ldots, k\}$ there is a deterministic polynomial time algorithm that approximates*

$$\max_{x_1, x_2, \ldots, x_n \in \{-1,1\}} \sum_{\substack{S \subseteq \{1,\ldots,n\} \\ |S|=\ell}} a_S \prod_{j \in S} x_j$$

within factor $f(n)$. Then there is a polynomial time refutation procedure that refutes w.h.p. a random k-CNF formula with $2^{4k+1} n f(n)^2$ clauses.

Remark 4.1. Note that for $\ell = 1$, there is a (trivial) exact algorithm and for $\ell = 2$ the result of [27, 25, 12] give a $O(\log n)$-approximation. Therefore, as long as $f(n) \geq O(\log n)$, the hypothesis in Theorem 4.1 is required to hold only for $3 \leq \ell \leq k$.

The best known refutation procedure for random $3 - CNF$ formulas works when they have $O\left(n^{3/2}\right)$ clauses [16]. This can be viewed as evidence that obtaining an improvement over our approximation factor to $o\left(n^{1/4}\right)$ is likely to be difficult.

Before proving Theorem 4.1 we shall introduce some notation. Let -1 represent logical TRUE and 1 represent logical FALSE. Let $\phi = \{C_1, C_2, \ldots, C_m\}$ be a k-CNF formula on variables $x := \{x_1, x_2, \ldots, x_n\}$. Let the set of indices of variables in the clause C_i be denoted as S_i, so that $|S_i| = k$. Define $\{\sigma_{ij} : 1 \leq i \leq m, j \in S_i\}$ as follows: $\sigma_{ij} = 1$ if x_j appears in clause C_i un-negated and $\sigma_{ij} = -1$ if x_j appears in clause C_i negated. Consider the expression

$$1 - \frac{1}{2^k} \prod_{j \in S_i} \left(1 + \sigma_{ij} x_j\right).$$

For any $\{-1,1\}$-assignment to variables, this expression evaluates to 1 if the clause C_i is satisfied and to 0 if the clause C_i is not satisfied. Therefore, the fraction of satisfied clauses is

$$\frac{1}{m}\sum_{i=1}^{m}\left(1-\frac{1}{2^k}\prod_{j\in S_i}\left(1+\sigma_{ij}x_j\right)\right). \tag{15}$$

For notational convenience, think of S_i as an ordered k-tuple of indices, and for $T\subseteq\{1,\ldots,k\}$, let $S_i[T]$ denote the subset of S_i given by the co-ordinates in T. With this notation (15) can be rewritten as

$$1-\frac{1}{2^k}-\frac{1}{m}\sum_{i=1}^{m}\sum_{\emptyset\neq T\subseteq\{1,\ldots,k\}}\prod_{j\in S_i[T]}\sigma_{ij}x_j = 1-\frac{1}{2^k}+\sum_{\emptyset\neq T\subseteq\{1,\ldots,k\}}\Gamma_T(x), \tag{16}$$

Where

$$\Gamma_T(x) := -\frac{1}{m}\sum_{i=1}^{m}\prod_{j\in S_i[T]}\sigma_{ij}x_j. \tag{17}$$

Lemma 4.2. *If ϕ is satisfiable, then there exists a $\{-1,1\}$ assignment to variables of ϕ and a nonempty set $T\subseteq\{1,\ldots k\}$ such that $\Gamma_T(x)\geq\frac{1}{2^k(2^k-1)}$. In other words,*

$$\exists\,\emptyset\neq T\subseteq\{1,\ldots,k\},\qquad \max_{x\in\{-1,1\}^n}\Gamma_T(x)\geq\frac{1}{2^k(2^k-1)}.$$

Proof. Since ϕ is satisfiable, there is an assignment x that satisfies every clause. For this assignment, the expression (16) has value 1. Thus, for some $T\subseteq[k], T\neq\emptyset$, it must be the case that $\Gamma_T(x)\geq\frac{1}{2^k(2^k-1)}$. □

Lemma 4.3. *Let ϕ be a random k-CNF formula with $m\geq 2^{4k+1}nf(n)^2$ clauses. Then with probability $1-2^{-\Omega(n)}$ over the choice of the formula, for every $\{-1,1\}$-assignment to variables and every nonempty $T\subseteq\{1,\ldots,k\}$ we have $\Gamma_T(x)\leq\frac{1}{2^{2k}f(n)}$. In other words,*

$$\forall\emptyset\neq T\subseteq\{1,\ldots,k\},\qquad \max_{x\in\{-1,1\}^n}\Gamma_T(x)\leq\frac{1}{2^{2k}f(n)}.$$

Proof. Fix any $\{-1,1\}$-assignment to the variables and a nonempty set $T\subseteq\{1,\ldots,k\}$. We will show that with probability $1-e^{-n}$ over the choice of ϕ we have $\Gamma_T(x)\leq\frac{1}{2^{2k}f(n)}$. Taking the union bound over all possible $\{-1,1\}$-assignments to variables and all choices for T implies the statement of the lemma.

Note that when ϕ is random, the signs σ_{ij} are random and independent, and therefore an inspection of the definition (17) shows that $\Gamma_T(x)$ is an average of m independent Bernoulli random variables. By the Chernoff bound,

$$\Pr\left[\Gamma_T(x)\geq\frac{1}{2^{2k}f(n)}\right]\;\leq\;\exp\left(-\frac{1}{2}\cdot\frac{m}{(2^{2k}f(n))^2}\right)\leq e^{-n}.$$

□

Proof of Theorem 4.1. The refutation procedure is very simple. Given a formula ϕ, use the $f(n)$-approximation algorithm to compute, for every nonempty $T\subseteq\{1,\ldots,k\}$, a number $\mathrm{Alg}(\Gamma_T)$ such that

$$\frac{1}{f(n)}\left(\max_{x\in\{-1,1\}^n}\Gamma_T(x)\right)\leq\mathrm{Alg}(\Gamma_T)\leq\max_{x\in\{-1,1\}^n}\Gamma_T(x).$$

If there is some $T\neq\emptyset$ for which $\mathrm{Alg}(\Gamma_T)\geq\frac{1}{2^k(2^k-1)f(n)}$, then say YES. Otherwise, say NO. Lemma 4.2 shows that this procedure always says YES if ϕ is satisfiable. Lemma 4.3 shows that the procedure says NO on a $1-2^{-\Omega(n)}$ fraction of random formulas. □

References

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

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

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

[4] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience [John Wiley & Sons], New York, second edition, 2000. With an appendix on the life and work of Paul Erdős.

[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] I. Bárány and Z. Füredi. Computing the volume is difficult. *Discrete Comput. Geom.*, 2(4):319–326, 1987.

[7] F. Barthe, M. Csörnyei, and A. Naor. A note on simultaneous polar and Cartesian decomposition. In *Geometric aspects of functional analysis*, volume 1807 of *Lecture Notes in Math.*, pages 1–19. Springer, Berlin, 2003.

[8] H. L. Bodlaender, P. Gritzmann, V. Klee, and J. van Leeuwen. Computational complexity of norm-maximization. *Combinatorica*, 10(2):203–225, 1990.

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

[10] A. Brieden, P. Gritzmann, R. Kannan, V. Klee, L. Lovász, and M. Simonovits. Deterministic and randomized polynomial-time approximation of radii. *Mathematika*, 48(1-2):63–105 (2003), 2001.

[11] B. Carl and A. Pajor. Gel′fand numbers of operators with values in a Hilbert space. *Invent. Math.*, 94(3):479–504, 1988.

[12] 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.

[13] U. Feige. Relations between average case complexity and approximation complexity. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, pages 534–543 (electronic), New York, 2002. ACM.

[14] U. Feige, J. H. Kim, and E. Ofek. Witnesses for non-satisfiability of dense random 3CNF formulas. In *47th Annual Symposium on Foundations of Computer Science*, pages 497–508. IEEE Computer Society, 2006.

[15] U. Feige and E. Ofek. Easily refutable subformulas of large random 3CNF formulas. In *Automata, languages and programming*, volume 3142 of *Lecture Notes in Comput. Sci.*, pages 519–530. Springer, Berlin, 2004.

[16] J. Friedman, A. Goerdt, and M. Krivelevich. Recognizing more unsatisfiable random k-SAT instances efficiently. *SIAM J. Comput.*, 35(2):408–430 (electronic), 2005.

[17] E. D. Gluskin. Extremal properties of orthogonal parallelepipeds and their applications to the geometry of Banach spaces. *Mat. Sb. (N.S.)*, 136(178)(1):85–96, 1988.

[18] P. Gritzmann and V. Klee. Computational convexity. In *Handbook of discrete and computational geometry*, CRC Press Ser. Discrete Math. Appl., pages 491–515. CRC, Boca Raton, FL, 1997.

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

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

[21] J. Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859 (electronic), 2001.

[22] J. Håstad and S. Venkatesh. On the advantage over a random assignment. *Random Structures Algorithms*, 25(2):117–149, 2004.

[23] J.-P. Kahane. *Some random series of functions*, volume 5 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 1985.

[24] 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.

[25] 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.

[26] A. Naor. The surface measure and cone measure on the sphere of l_p^n. *Trans. Amer. Math. Soc.*, 359(3):1045–1079 (electronic), 2007.

[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] R. Paley and A. Zygmund. A note on analytic functions in the unit circle. *Proc. Camb. Phil. Soc.*, 28:266–272, 1932.

[29] G. Schechtman and J. Zinn. On the volume of the intersection of two L_p^n balls. *Proc. Amer. Math. Soc.*, 110(1):217–224, 1990.

[30] 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.

[31] K. Varadarajan, S. Venkatesh, and J. Zhang. On approximating the radii of point sets in high dimensions. In *43rd Annual Symposium on Foundations of Computer Science*, pages 561–569. IEEE Computer Society, 2002.