MATHEMATICS OF OPERATIONS RESEARCH
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0364-765X | EISSN 1526-5471 | 00 | 0000 | 0001

INFORMS
DOI 10.1287/xxxx.0000.0000

Hardness and Approximation Results for L_p–Ball Constrained Homogeneous Polynomial Optimization Problems

Ke Hou
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, khou@se.cuhk.edu.hk

Anthony Man–Cho So
Department of Systems Engineering and Engineering Management, and by courtesy, CUHK–BGI Innovation Institute of Trans–omics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, manchoso@se.cuhk.edu.hk

In this paper, we establish hardness and approximation results for various L_p–ball constrained homogeneous polynomial optimization problems, where $p \in [2, \infty]$. Specifically, we prove that for any given $d \geq 3$ and $p \in [2, \infty]$, both the problem of optimizing a degree–d homogeneous polynomial over the L_p–ball and the problem of optimizing a degree–d multilinear form (regardless of its super–symmetry) over L_p–balls are NP–hard. On the other hand, we show that these problems can be approximated to within a factor of $\Omega\left((\log n)^{(d-2)/p}/n^{d/2-1}\right)$ in deterministic polynomial time, where n is the number of variables. We further show that with the help of randomization, the approximation guarantee can be improved to $\Omega((\log n/n)^{d/2-1})$, which is independent of p and is currently the best for the aforementioned problems. Our results unify and generalize those in the literature, which focus either on the quadratic case or the case where $p \in \{2, \infty\}$. We believe that the wide array of tools used in this paper will have further applications in the study of polynomial optimization problems.

Key words: polynomial optimization; approximation algorithms; algorithmic convex geometry; convex programming
MSC2000 subject classification: Primary: 90C26, 90C59; secondary: 15A69, 52C07
OR/MS subject classification: Primary: programming; secondary: mathematics

1. Introduction. Motivated by its diverse applications and profound connections to various branches of mathematics, polynomial optimization has been the focus of much research effort during the past decade or so. From an algorithmic perspective, polynomial optimization problems are generally intractable. Thus, a fundamental research issue is to determine their approximability. One important class of problems whose approximability has been extensively investigated in recent years is that of homogeneous polynomial optimization with L_2–norm constraints. The first results in this direction were obtained by de Klerk et al. [9] and Barvinok [4], who showed that certain specially structured L_2–sphere constrained polynomial optimization problems admit polynomial–time approximation schemes (PTASes). These were then followed by the work of Luo and Zhang [21], in which an approximation algorithm was developed for homogeneous quartic optimization problems with quadratic constraints (which includes the L_2–ball as a special case). Around the same time, Ling et al. [19] considered the problem of approximately optimizing a biquadratic function over the Cartesian product of two L_2–spheres; while Zhang et al. [33] studied the hardness and approximability of certain L_2–sphere constrained homogeneous cubic optimization problems. Since then, there have been significant activities in this line of research. For instance, in [32, 20, 31], various researchers derived approximation results for the problem of optimizing a biquadratic function over quadratic constraints, thereby extending the results in [19]. In [12], He et al. improved and substantially extended the results in [21] by providing approximation algorithms for optimizing a general homogeneous polynomial over quadratic constraints (see also [17] for some latest developments). It is worth noting that most of the aforementioned results were obtained using semidefinite relaxation techniques, and that most of the algorithms are randomized. Recently, in a marked departure from the semidefinite relaxation paradigm, So [28] employed techniques from algorithmic convex geometry to design deterministic approximation algorithms for various L_2–sphere constrained homogeneous polynomial optimization problems. The algorithms in [28] have a worst–case approximation guarantee of $\Omega((\log n/n)^{d/2-1})$, where n is the number of variables and d is the degree of the polynomial. Roughly speaking, this means that given any problem instance, the algorithms will produce a feasible solution whose objective value is at least $\Omega((\log n/n)^{d/2-1})$ times the optimum. This improves upon the $\Omega((1/n)^{d/2-1})$ bound established in [19, 12, 33] and is currently the best for general L_2–sphere constrained homogeneous polynomial and multiquadratic optimization problems. Such development raises a natural question: Can the approach in [28] be applied to other classes of polynomial optimization problems?

In this paper, we address the above question by extending the approach in [28] to study the L_p–ball constrained homogeneous polynomial optimization problem; i.e., problem of the form

$$\max\{f(x) : \|x\|_p \leq 1\}, \tag{1}$$

where $p \in [2,\infty]$ and $f : \mathbb{R}^n \to \mathbb{R}$ is a homogeneous polynomial of (fixed) degree $d \geq 3$. Our motivation for studying Problem (1) is twofold. First, it is a natural extension of the matrix norm problem in [5, 30] and the L_p–Grothendieck problem in [16]—both of which concern quadratic f's with certain structure—as well as the L_∞–ball constrained trilinear optimization problem in [14] and the L_2–ball constrained homogeneous polynomial optimization problem in [21, 12]. However, there is no prior hardness or approximation result for Problem (1) in its full generality. Secondly, Problem (1) lies at the heart of many applications. For instance, Baratchart et al. [3] demonstrated that many labeling problems in pattern recognition and image processing can be tackled by maximizing a certain polynomial over an L_p–ball. In addition, the L_p–singular value and singular vector of a tensor—which have been extensively studied in the spectral theory of tensors and play an important role in signal processing, automatic control and data analysis—can be defined as the optimal value of and optimal solution to an L_p–ball constrained homogeneous polynomial optimization problem, respectively [18, 26]. As our main contribution, we obtain both hardness and approximation results for Problem (1). Specifically, on the hardness side, we show that Problem (1) is NP–hard for any given $d \geq 3$ and $p \in [2,\infty]$. To the best of our knowledge, this is the first hardness result for Problem (1) that holds for any given $d \geq 3$ and $p \in [2,\infty]$. By contrast, existing hardness results for Problem (1), such as those in [22, 33, 1, 13], hold only for certain values of d and p. A key tool we used to prove the hardness result is a tensor symmetrization procedure introduced by Ragnarsson and Van Loan [27], which allows us to establish the equivalence between multilinear optimization problems and certain homogeneous polynomial optimization problems. On the approximation side, we show that Problem (1) can be approximated to within a factor of $\Omega\left((\log n)^{(d-2)/p}/n^{d/2-1}\right)$ by a deterministic polynomial–time algorithm. Furthermore, if one allows randomization, then the approximation bound can be improved to $\Omega((\log n/n)^{d/2-1})$, independent of p. In the process of deriving these results, we also establish the hardness of and develop approximation algorithms for certain L_p–ball constrained multilinear optimization problems, which could be of independent interest. We remark that the aforementioned results apply only to the case where $p \in [2,\infty]$. The case where $p \in [1,2)$, which is not covered in this paper, does not seem to be well understood, even when f is quadratic. We refer the interested reader to [30, 9, 6] for some results in this direction.

Before describing in detail our approximation algorithms for Problem (1), let us give an overview of our approach and highlight some of the key technical issues. To fix ideas, let us first consider the case where $d = 3$; i.e., $f(x) = \sum_{i,j,k=1}^n a_{ijk}x_ix_jx_k$ for some order–3 tensor $\mathcal{A} = (a_{ijk}) \in \mathbb{R}^{n\times n\times n}$.

Using by–now standard techniques (see, e.g., [12, 28]), one can show that the optimal value of Problem (1) is within a constant factor of that of its multilinear relaxation, which in the case of $d=3$ is given by

$$\max\left\{\sum_{i,j,k=1}^{n} a_{ijk}x_iy_jz_k : \|x\|_p \le 1, \|y\|_p \le 1, \|z\|_p \le 1\right\}. \tag{2}$$

Thus, as far as approximating Problem (1) is concerned, it suffices to focus on Problem (2). Although the latter generally remains NP–hard (see Proposition 1 and Theorem 3), intuitively it should be easier to handle because of the decoupling of variables. Indeed, following the ideas in [14, 28], one can show that the optimal value of Problem (2) is equal to half times the L_q–diameter of a certain convex body $\mathcal{K}_p$, where $q=p/(p-1)\in[1,2]$ is the conjugate of p. However, the latter quantity is known to be efficiently approximable only when $p=2$. To tackle the case where $p>2$, we do not work on $\mathcal{K}_p$ directly as in [28]. Instead, we construct another convex body $\mathcal{K}_p'$ whose L_q–diameter is within a constant factor of the optimal value of Problem (2) but can be approximated efficiently. The validity of our construction is established using Grothendieck's inequality—a tool that originates from functional analysis and has since found many applications in optimization and theoretical computer science; see, e.g., [29, 15, 25]. Consequently, we are able to approximate Problem (2) and hence also Problem (1) in polynomial time for the case where $d=3$.

To extend the above results to the case where $d>3$, a natural idea is to apply recursion. We will present two implementations of this idea, which will lead to two algorithms with different characteristics. The first is based on the following crucial observation (see Proposition 8): Suppose that we have a deterministic approximation algorithm $\mathscr{A}_d$ for optimizing a degree–d multilinear form over L_p–balls, where $d\ge 3$. Consider a degree–$(d+1)$ multilinear form F. For any $\bar{x}^1\in\mathbb{R}^n$, let $G_d(\bar{x}^1)$ be the value returned by $\mathscr{A}_d$ when applied to the degree–d multilinear optimization problem

$$\max\left\{F(\bar{x}^1,x^2,\dots,x^{d+1}) : \|x^i\|_p \le 1 \quad \text{for } i=2,3,\dots,d+1\right\}.$$

Then, the function G_d essentially defines a norm on $\mathbb{R}^n$. Such a property, which was first established in [28] for the case where $p=2$, is extremely useful and can be of independent interest. In particular, it allows us to utilize existing polytopal approximations of L_p–balls [8] to design a deterministic $\Omega\left((\log n)^{(d-2)/p}\big/n^{d/2-1}\right)$–approximation algorithm for Problem (1).

The second approach to implementing the recursion idea is by randomization. Specifically, consider a degree–d multilinear form F, where $d>3$. It is known that if $x^2,\dots,x^d\in\mathbb{R}^n$ are arbitrary and $\xi\in\mathbb{R}^n$ is a random vector uniformly distributed on the L_q–sphere, then

$$F(\xi,x^2,\dots,x^d)\ge\Omega\left(\sqrt{\frac{\log n}{n}}\right)\cdot\left[\max_{\|x\|_p\le 1}F(x,x^2,\dots,x^d)\right]$$

holds with a probability that is at least inversely proportional to a polynomial in n; cf. [14, Lemma 3.3]. Using this result, it is not hard to show that any β_{d-1}–approximation algorithm for optimizing a degree–$(d-1)$ multilinear form over L_p–balls will yield an $\Omega(\beta_{d-1}\sqrt{\log n/n})$–approximation algorithm for Problem (1). To complete the argument, we show by induction that β_{d-1} can be taken as $\beta_{d-1}=\Omega((\log n/n)^{(d-1)/2-1})$. This gives an $\Omega((\log n/n)^{d/2-1})$–approximation algorithm for Problem (1). It should be noted that unlike the deterministic algorithm described above, the algorithm obtained using this approach is randomized and thus will only attain the stated approximation ratio with high probability. However, it is much easier to implement than its deterministic counterpart.

The rest of the paper is organized as follows. Section 2 contains the preliminaries. In Section 3, we show that the problem of optimizing a homogeneous polynomial of fixed degree over an L_p–ball is NP–hard. Then, in Section 4, we introduce a multilinear relaxation of the L_p–ball constrained homogeneous polynomial optimization problem and show that it is equivalent to the latter from an approximation perspective. We also discuss the hardness of the multilinear relaxation. In Section 5, we develop both deterministic and randomized polynomial–time approximation algorithms for the problem of optimizing a multilinear form over L_p–balls by relating it to the problem of determining the diameters of certain convex bodies. Finally, we conclude with some closing remarks in Section 6.

2. Preliminaries. We begin by introducing the notation and definitions used in this paper. A *tensor* is a multidimensional array, and the *order* of a tensor is the number of dimensions. Let $\mathcal{A} = (a_{i_1 i_2 \cdots i_d}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ be a tensor of order d. We denote its $(i_1, i_2, \ldots, i_d)$–th element by either $a_{i_1 i_2 \cdots i_d}$ or $[\mathcal{A}]_{i_1 i_2 \cdots i_d}$. We say that $\mathcal{A}$ is *non–zero* if at least one of its elements is non–zero, and is *cubical* if $n_1 = n_2 = \cdots = n_d$. A cubical tensor is said to be *super–symmetric* if every element $a_{i_1 i_2 \cdots i_d}$ is invariant under any permutation of the indices.

Let K and $j_1, j_2, \ldots, j_K$ be integers such that $1 \le K \le d$ and $1 \le j_1 < j_2 < \cdots < j_K \le d$. Furthermore, let $x^{j_k} \in \mathbb{R}^{n_{j_k}}$, where $k = 1, \ldots, K$, be given vectors. We use $\mathcal{A}(x^{j_1}, x^{j_2}, \ldots, x^{j_K})$ to denote the order–$(d-K)$ tensor obtained by "summing out" the indices $j_1, j_2, \ldots, j_K$ from the order–d tensor $\mathcal{A} = (a_{i_1 i_2 \cdots i_d}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ using $x^{j_1}, x^{j_2}, \ldots, x^{j_K}$. For instance, if $K = 2$, $j_1 = 2$ and $j_2 = 4$, then

$$\mathcal{A}(x^2, x^4)_{i_1 i_3 i_5 i_6 \cdots i_d} = \sum_{i_2=1}^{n_2} \sum_{i_4=1}^{n_4} a_{i_1 i_2 \cdots i_d} x^2_{i_2} x^4_{i_4}.$$

Given an order–d tensor $\mathcal{A} = (a_{i_1 i_2 \cdots i_d}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$, we can associate it with a multilinear form $F_{\mathcal{A}} : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \cdots \times \mathbb{R}^{n_d} \to \mathbb{R}$ via

$$F_{\mathcal{A}}(x^1, x^2, \ldots, x^d) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} a_{i_1 i_2 \cdots i_d} x^1_{i_1} x^2_{i_2} \cdots x^d_{i_d}.$$

If $\mathcal{A}$ is super–symmetric with $n_1 = n_2 = \cdots = n_d = n$, then we can further associate it with a homogeneous degree–d polynomial $f_{\mathcal{A}} : \mathbb{R}^n \to \mathbb{R}$ via

$$f_{\mathcal{A}}(x) = F_{\mathcal{A}}(x, x, \ldots, x) = \sum_{1 \le i_1, \ldots, i_d \le n} a_{i_1 i_2 \cdots i_d} x_{i_1} x_{i_2} \cdots x_{i_d}.$$

In general, even if $\mathcal{A}$ is not super–symmetric or even cubical, it is still possible to relate the multilinear form $F_{\mathcal{A}}$ to a certain homogeneous degree–d polynomial via symmetrization [27]. To introduce this procedure, we need some preliminary definitions. Let $\pi = (\pi_1, \pi_2, \ldots, \pi_d)$ be a permutation of the set $\{1, 2, \ldots, d\}$. The *π–transpose* of $\mathcal{A} = (a_{i_1 i_2 \cdots i_d}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is the order–d tensor $\mathcal{A}^\pi = (\bar{a}_{i_{\pi_1} i_{\pi_2} \cdots i_{\pi_d}}) \in \mathbb{R}^{n_{\pi_1} \times n_{\pi_2} \times \cdots \times n_{\pi_d}}$ whose elements are given by

$$\bar{a}_{i_{\pi_1} i_{\pi_2} \cdots i_{\pi_d}} = a_{i_1 i_2 \cdots i_d} \quad \text{for } i_j = 1, \ldots, n_j;\ j = 1, \ldots, d.$$

Let $N = n_1 + n_2 + \cdots + n_d$ and partition the index set $\{1, \ldots, N\}$ into sets of consecutive integers as follows:

$$\{1, \ldots, N\} = \bigcup_{j=1}^{d} B_j, \quad \text{where } B_j = \left\{ \sum_{i=1}^{j-1} n_i + 1, \ldots, \sum_{i=1}^{j} n_i \right\}. \tag{3}$$

Given an arbitrary cubical order–d tensor $\mathcal{B} \in \mathbb{R}^{N^d}$ and $\chi_i \in \{1, \ldots, d\}$ for $i = 1, \ldots, d$, the $(\chi_1, \ldots, \chi_d)$*–th block* of $\mathcal{B}$ is defined as the sub–tensor

$$\mathcal{B}_{\chi_1 \chi_2 \cdots \chi_d} = (b_{i_1 i_2 \cdots i_d})_{i_j \in B_{\chi_j};\, j=1,\ldots,d} \in \mathbb{R}^{n_{\chi_1} \times n_{\chi_2} \times \cdots \times n_{\chi_d}}.$$

Armed with these definitions, we define the *symmetrization* of $\mathcal{A}=(a_{i_1 i_2 \cdots i_d}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ as the order–d cubical tensor $\text{sym}(\mathcal{A}) \in \mathbb{R}^{N^d}$ whose blocks are given by

$$[\text{sym}(\mathcal{A})]_{\chi_1 \chi_2 \cdots \chi_d} = \begin{cases} \mathcal{A}^{\chi} & \text{if } \chi=(\chi_1,\chi_2,\ldots,\chi_d) \text{ is a permutation of } \{1,2,\ldots,d\}, \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

For instance, when $d=2$, $\mathcal{A}$ is an $n_1 \times n_2$ matrix, and its symmetrization is given by the well–known construction

$$\text{sym}(\mathcal{A}) = \begin{bmatrix} \mathbf{0} & \mathcal{A} \\ \mathcal{A}^T & \mathbf{0} \end{bmatrix}.$$

More generally, it is known that the tensor $\text{sym}(\mathcal{A})$ enjoys the following properties [27]:

1. $\text{sym}(\mathcal{A})$ is super–symmetric. In particular, it can be associated with a homogeneous degree–d polynomial $f_{\text{sym}(\mathcal{A})}$.
2. For every $x = \left[(x^1)^T \, (x^2)^T \, \cdots \, (x^d)^T \right]^T \in \mathbb{R}^N$, where $x^i \in \mathbb{R}^{n_i}$ and $i=1,\ldots,d$, we have

$$f_{\text{sym}(\mathcal{A})}(x) = d! \cdot F_{\mathcal{A}}(x^1, x^2, \ldots, x^d). \tag{4}$$

Now, let $d \geq 3$ and $p \in [2,\infty]$ be given. Let $\mathcal{A}=(a_{i_1 i_2 \cdots i_d}) \in \mathbb{R}^{n^d}$ be an arbitrary non–zero super–symmetric tensor of order d, and let $f_{\mathcal{A}} : \mathbb{R}^n \to \mathbb{R}$ be the corresponding homogeneous polynomial. Our main objective in this paper is to study the algorithmic aspects of the following L_p–ball constrained homogeneous polynomial optimization problem:

$$(\mathsf{HP}) \qquad \begin{array}{rll} \bar{v} = & \text{maximize} & f_{\mathcal{A}}(x) \equiv \displaystyle\sum_{1 \leq i_1,\ldots,i_d \leq n} a_{i_1 i_2 \cdots i_d} x_{i_1} x_{i_2} \cdots x_{i_d} \\ & \text{subject to} & \|x\|_p \leq 1, \, x \in \mathbb{R}^n. \end{array}$$

3. Hardness of L_p–ball constrained homogeneous polynomial optimization.

Our first result is the following theorem, which concerns the complexity of Problem (HP):

THEOREM 1. *Problem* (HP) *is NP–hard for any given $d \geq 3$ and $p \in [2,\infty]$.*

The proof of Theorem 1 consists of two steps. First, we show that the problem of maximizing a degree–d *multilinear form* over L_p–balls is NP–hard for any given $d \geq 3$ and $p \in [2,\infty]$. Then, we give a polynomial–time reduction of this problem to Problem (HP) using the symmetrization procedure introduced in Section 2, thereby proving the NP–hardness of the latter.

To begin, let us formally define the problem used in the first step.

(ML) Let $\mathcal{A}=(a_{i_1 i_2 \cdots i_d}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ be an arbitrary non–zero order–d tensor, and let $F_{\mathcal{A}} : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \cdots \times \mathbb{R}^{n_d} \to \mathbb{R}$ be the corresponding multilinear form. Solve

$$\begin{array}{rll} v_{\mathsf{ML}}(\mathcal{A},d) = & \text{maximize} & F_{\mathcal{A}}(x^1,x^2,\ldots,x^d) \\ & \text{subject to} & \|x^i\|_p \leq 1, \, x^i \in \mathbb{R}^{n_i} \quad \text{for } i=1,\ldots,d. \end{array}$$

PROPOSITION 1. *Problem* (ML) *is NP–hard for any given $d \geq 3$ and $p \in [2,\infty]$.*

Proof. Let $d \geq 3$ and $p \in [2,\infty]$ be fixed. Consider first the case where $p \in (2,\infty]$. Our plan is to reduce the following problem—which is known to be NP–hard [30]—to Problem (ML):

(NORM) Let $B \in \mathbb{R}^{m \times n}$ and $p \in (2,\infty]$ be given. Let $q = p/(p-1)$ be the conjugate of p. Compute $\|B\|_{p\to q}$, the $p \to q$ norm of B, where

$$\|B\|_{p\to q} = \max\{\|By\|_q : \|y\|_p \leq 1\}.$$

Towards that end, suppose that we are given an instance of Problem (NORM). By Hölder's inequality, we have

$$\begin{aligned}\max_{\|y\|_p\le 1}\|By\|_q &= \max_{\|x\|_p\le 1,\,\|y\|_p\le 1} x^T By \\ &= \max_{\substack{\|x\|_p\le 1,\,\|y\|_p\le 1\\ |z_1|,\ldots,|z_{d-2}|\le 1}} \left(\prod_{i=1}^{d-2} z_i\right) x^T By \\ &= \max_{\substack{\|x\|_p\le 1,\,\|y\|_p\le 1\\ |z_1|,\ldots,|z_{d-2}|\le 1}} F_{\mathcal{A}}\left(z_1,z_2,\ldots,z_{d-2},x,y\right),\end{aligned}$$

where $\mathcal{A}=(a_{1,\ldots,1,i,j})\in\mathbb{R}^{1\times\cdots\times 1\times m\times n}$ is the order–d tensor with $a_{1,\ldots,1,i,j}=b_{ij}$ for $i=1,\ldots,m$ and $j=1,\ldots,n$; $F_{\mathcal{A}}:\mathbb{R}\times\cdots\times\mathbb{R}\times\mathbb{R}^m\times\mathbb{R}^n\to\mathbb{R}$ is the multilinear form associated with $\mathcal{A}$. This establishes the NP–hardness of Problem (ML) when $d\ge 3$ and $p\in(2,\infty]$.

Next, consider the case where $p=2$. It has been shown in [12, Proposition 2] that Problem (ML) is NP–hard when $d=3$ and $p=2$. Now, let $\mathcal{B}=(b_{ijk})\in\mathbb{R}^{n_1\times n_2\times n_3}$ be an arbitrary non–zero order–3 tensor, and let $F_{\mathcal{B}}:\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\mathbb{R}^{n_3}\to\mathbb{R}$ be the corresponding multilinear form. Using similar argument as above, for any given $d\ge 4$, we have

$$\begin{aligned}\max_{\|w\|_2,\,\|x\|_2,\,\|y\|_2\le 1} F_{\mathcal{B}}(w,x,y) &= \max_{\substack{\|w\|_2,\,\|x\|_2,\,\|y\|_2\le 1\\ |z_1|,\ldots,|z_{d-3}|\le 1}} \left(\prod_{i=1}^{d-3} z_i\right) F_{\mathcal{B}}(w,x,y) \\ &= \max_{\substack{\|w\|_2,\,\|x\|_2,\,\|y\|_2\le 1\\ |z_1|,\ldots,|z_{d-3}|\le 1}} F_{\mathcal{A}}(z_1,z_2,\ldots,z_{d-3},w,x,y),\end{aligned}$$

where $\mathcal{A}=(a_{1,\ldots,1,i,j,k})\in\mathbb{R}^{1\times\cdots\times 1\times n_1\times n_2\times n_3}$ is the order–d tensor with $a_{1,\ldots,1,i,j,k}=b_{ijk}$ for $i=1,\ldots,n_1$, $j=1,\ldots,n_2$ and $k=1,\ldots,n_3$; $F_{\mathcal{A}}:\mathbb{R}\times\cdots\times\mathbb{R}\times\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\mathbb{R}^{n_3}\to\mathbb{R}$ is the multilinear form associated with $\mathcal{A}$. Thus, we conclude that when $p=2$, Problem (ML) remains NP–hard for each fixed $d\ge 3$. □

Next, we have the following proposition, which links the optimization of the multilinear form associated with a tensor $\mathcal{A}$ to that of the homogeneous polynomial associated with $\mathrm{sym}(\mathcal{A})$.

PROPOSITION 2. *Let $d\ge 2$ and $p\in[2,\infty]$ be given, and let $\mathcal{A}\in\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$ be an arbitrary non–zero order–d tensor. Set $N=n_1+n_2+\cdots+n_d$. Consider the optimization problems*

$$\begin{array}{ll}\text{maximize} & d!\cdot F_{\mathcal{A}}(x^1,x^2,\ldots,x^d)\\ \text{subject to} & \|x^i\|_p\le 1,\ x^i\in\mathbb{R}^{n_i}\quad \text{for } i=1,\ldots,d\end{array}\tag{5}$$

and

$$\begin{array}{ll}\text{maximize} & f_{sym(\mathcal{A})}(z)\\ \text{subject to} & \|z\|_p\le d^{1/p},\ z\in\mathbb{R}^N,\end{array}\tag{6}$$

where $F_{\mathcal{A}}$ is the multilinear form associated with $\mathcal{A}$ and $f_{sym(\mathcal{A})}$ is the homogeneous polynomial associated with the symmetrization of $\mathcal{A}$ (see Section 2). Let $(\bar{x}^1,\bar{x}^2,\ldots,\bar{x}^d)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\cdots\times\mathbb{R}^{n_d}$ and $\bar{z}=\left[\,(\bar{z}^1)^T\,(\bar{z}^2)^T\,\cdots\,(\bar{z}^d)^T\,\right]^T$, where $\bar{z}^i\in\mathbb{R}^{n_i}$ for $i=1,\ldots,d$, be optimal solutions to problems (5) and (6), respectively. Then, the following hold:

(a) $\|\bar{z}^i\|_p=1$ for $i=1,\ldots,d$.

(b) $(\bar{z}^1,\bar{z}^2,\ldots,\bar{z}^d)\in\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$ and $\bar{x}=\left[\,(\bar{x}^1)^T\,(\bar{x}^2)^T\,\cdots\,(\bar{x}^d)^T\,\right]^T\in\mathbb{R}^N$ are optimal solutions to problems (5) and (6), respectively.

Proof. Let us first consider the case where $p=\infty$. By (4), Problem (6) is equivalent to

$$\begin{array}{ll} \text{maximize} & d!\cdot F_{\mathcal{A}}(z^1,z^2,\ldots,z^d) \\ \text{subject to} & \|z^i\|_\infty\le 1,\ z^i\in\mathbb{R}^{n_i} \quad \text{for } i=1,\ldots,d, \end{array}$$

which has exactly the same form as Problem (5). Thus, the desired results follow immediately.

Now, consider the case where $p\in[2,\infty)$. To prove (a), we again appeal to (4), which implies the equivalence of Problem (6) and the following problem:

$$\begin{array}{ll} \text{maximize} & d!\cdot F_{\mathcal{A}}(z^1,z^2,\ldots,z^d) \\ \text{subject to} & \displaystyle\sum_{i=1}^d \|z^i\|_p^p\le d, \\ & z^i\in\mathbb{R}^{n_i} \quad \text{for } i=1,\ldots,d. \end{array}$$

Since $\mathcal{A}$ is non–zero, we must have $\sum_{i=1}^d\|\bar{z}^i\|_p^p=d$ and $\|\bar{z}^i\|_p>0$ for $i=1,\ldots,d$. Now, suppose that $\|\bar{z}^j\|_p^p=\theta\neq 1$ for some $j\in\{1,\ldots,d\}$. Then, we have $\sum_{i\neq j}\|\bar{z}^i\|_p^p=d-\theta>0$. Upon setting

$$\hat{z}^i=\begin{cases} \left(\dfrac{d-1}{d-\theta}\right)^{1/p}\bar{z}^i & \text{if } i\neq j, \\ \theta^{-1/p}\bar{z}^j & \text{otherwise,} \end{cases}$$

we obtain

$$\|\hat{z}^j\|_p^p=1,\quad \sum_{i=1}^d\|\hat{z}^i\|_p^p=\frac{d-1}{d-\theta}\sum_{i\neq j}\|\bar{z}^i\|_p^p+\|\hat{z}^j\|_p^p=d$$

and

$$F_{\mathcal{A}}(\hat{z}^1,\hat{z}^2,\ldots,\hat{z}^d)=(d-1)^{\frac{d-1}{p}}\cdot\left((d-\theta)^{d-1}\theta\right)^{-1/p}\cdot F_{\mathcal{A}}(\bar{z}^1,\bar{z}^2,\ldots,\bar{z}^d). \tag{7}$$

In particular, we see that $\hat{z}=\left[(\hat{z}^1)^T\,(\hat{z}^2)^T\,\cdots\,(\hat{z}^d)^T\right]^T\in\mathbb{R}^N$ is feasible for (6). It is easy to verify that the function $t\mapsto((d-t)^{d-1}t))^{-1/p}$ is strictly convex on $(0,d)$ and is minimized at $t=1$. Since $\theta\neq 1$ and $\mathcal{A}$ is non–zero, it follows from (7) that $F_{\mathcal{A}}(\hat{z}^1,\hat{z}^2,\ldots,\hat{z}^d)>F_{\mathcal{A}}(\bar{z}^1,\bar{z}^2,\ldots,\bar{z}^d)$, which contradicts the optimality of $\bar{z}$. Thus, we have $\|\bar{z}^i\|_p^p=1$ for $i=1,\ldots,d$, as desired.

To prove (b), we first observe that since $\bar{x}=\left[(\bar{x}^1)^T\,(\bar{x}^2)^T\,\cdots\,(\bar{x}^d)^T\right]^T\in\mathbb{R}^N$ is feasible for Problem (6) and $f_{\mathrm{sym}(\mathcal{A})}(\bar{x})=d!\cdot F_{\mathcal{A}}(\bar{x}^1,\bar{x}^2,\ldots,\bar{x}^d)$ by (4), we have $f_{\mathrm{sym}(\mathcal{A})}(\bar{z})\ge d!\cdot F_{\mathcal{A}}(\bar{x}^1,\bar{x}^2,\ldots,\bar{x}^d)$. Now, the result in (a) implies that $(\bar{z}^1,\bar{z}^2,\ldots,\bar{z}^d)\in\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$ is feasible for Problem (5), and hence using (4) we obtain $f_{\mathrm{sym}(\mathcal{A})}(\bar{z})=d!\cdot F_{\mathcal{A}}(\bar{z}^1,\bar{z}^2,\ldots,\bar{z}^d)\le d!\cdot F_{\mathcal{A}}(\bar{x}^1,\bar{x}^2,\ldots,\bar{x}^d)$. This completes the proof. □

We are now ready to complete the proof of Theorem 1.

Proof of Theorem 1. Proposition 2 implies that Problem (ML) is equivalent to

$$\begin{array}{ll} \text{maximize} & f_{\mathrm{sym}(\mathcal{A})}(z) \\ \text{subject to} & \|z\|_p\le 1,\ z\in\mathbb{R}^N, \end{array}$$

where $N=n_1+n_2+\cdots+n_d$. The latter is clearly an instance of Problem (HP). Moreover, when $d\ge 3$ is fixed, the size of $\mathrm{sym}(\mathcal{A})$ is polynomial in $n_1,n_2,\ldots,n_d$. Thus, for any given $d\ge 3$ and $p\in[2,\infty]$, we can reduce Problem (ML) to Problem (HP) in polynomial time, which implies that the latter is NP–hard, as desired. □

4. L_p–ball constrained homogeneous polynomial optimization and its multilinear relaxation. In view of Theorem 1, we now turn our attention to the task of designing polynomial–time approximation algorithms for Problem (HP) with provable guarantees. Towards that end, consider the following multilinear relaxation of Problem (HP):

$$(\mathsf{MR}) \quad \begin{array}{rll} v^* = & \text{maximize} & F_{\mathcal{A}}(x^1, x^2, \ldots, x^d) \equiv \displaystyle\sum_{1 \le i_1, \ldots, i_d \le n} a_{i_1 i_2 \cdots i_d} x^1_{i_1} x^2_{i_2} \cdots x^d_{i_d} \\ & \text{subject to} & \|x^i\|_p \le 1,\, x^i \in \mathbb{R}^{n_i} \quad \text{for } i = 1, \ldots, d. \end{array}$$

Since $f_{\mathcal{A}}(x) = F_{\mathcal{A}}(x, x, \ldots, x)$ for all $x \in \mathbb{R}^n$ and $x = \mathbf{0}$ is feasible for Problem (HP), we clearly have $v^* \ge \bar{v} \ge 0$. Our motivation for studying Problem (MR) comes from the following result, which essentially states that v^* and $\bar{v}$ are within a constant factor of each other when $d \ge 3$ is fixed.

Theorem 2. *Let $d \ge 3$ and $p \in [2, \infty]$ be given. Suppose there is a polynomial–time algorithm $\mathscr{A}_{\mathsf{MR}}$ that, given any instance of Problem* (MR)*, returns a feasible solution whose objective value is at least αv^* for some $\alpha \in (0, 1]$. Then, there is a polynomial–time algorithm $\mathscr{A}_{\mathsf{HP}}$ that, given any instance of Problem* (HP)*, returns a solution $\hat{x} \in \mathbb{R}^n$ with $\|\hat{x}\|_p \le 1$ and*

$$\begin{aligned} f_{\mathcal{A}}(\hat{x}) &\ge \alpha \cdot d! \cdot d^{-d} \cdot v^* \\ &\ge \alpha \cdot d! \cdot d^{-d} \cdot \bar{v} \qquad \textit{for odd } d \ge 3, \\ f_{\mathcal{A}}(\hat{x}) - \underline{v} &\ge 2\alpha \cdot d! \cdot d^{-d} \cdot v^* \\ &\ge \alpha \cdot d! \cdot d^{-d} \cdot (\bar{v} - \underline{v}) \qquad \textit{for even } d \ge 4, \end{aligned}$$

where $\underline{v} = \min_{\|x\|_p \le 1} f_{\mathcal{A}}(x)$. In other words, the algorithm $\mathscr{A}_{\mathsf{HP}}$ has an approximation guarantee *(resp.* relative approximation guarantee*) of $\alpha \cdot d! \cdot d^{-d}$ when d is odd (resp. even).*

For a proof of Theorem 2, see Appendix A. We remark that for the case where $p = 2$, an analogous result has been established in [12]; cf. [28, Theorem 1]. However, if in addition $d \ge 3$ is odd, then Problem (HP) and Problem (MR) are in fact equivalent, in the sense that one can extract a solution $\hat{x} \in \mathbb{R}^n$ with $\|\hat{x}\|_2 = 1$ and $f_{\mathcal{A}}(\hat{x}) = v^* = \bar{v}$; see, e.g., [24].

Theorem 2 shows that any algorithm for solving Problem (MR) will translate into an algorithm for approximating Problem (HP). Although it seems intuitive that Problem (MR) is NP–hard as well, such a result does not follow directly from Proposition 1, as the tensor associated with the objective function in Problem (ML) is not required to be super–symmetric or even cubical. The following theorem fills this gap:

Theorem 3. *Problem* (MR) *is NP–hard for any given $d \ge 3$ and $p \in [2, \infty]$.*

The proof of Theorem 3 is quite involved and can be found in Appendix B. We remark that Theorem 3 is, to the best of our knowledge, the first hardness result for the problem of optimizing a *super–symmetric* multilinear form that holds for *any* given $d \ge 3$ and $p \in [2, \infty]$; cf. [13].

5. L_p–ball constrained multilinear optimization and diameters of convex bodies. Given that both Problem (MR) and Problem (ML) are NP–hard, we shall study the slightly more general Problem (ML), where the focus will be on developing approximation algorithms with provable guarantees. Since the case where $p = 2$ has already been investigated in [28], we shall assume that $p \in (2, \infty]$ in the sequel.

5.1. Base case: Approximating L_p–ball constrained trilinear maximization. Let us begin by considering the case where $d=3$. Specifically, let $\mathcal{A}=(a_{ijk})\in\mathbb{R}^{n_1\times n_2\times n_3}$ be an arbitrary non–zero order–3 tensor. Without loss of generality, we assume that $1\le n_1\le n_2\le n_3$. Then, Problem (ML) becomes

$$\begin{aligned} v_{\mathsf{ML}}(\mathcal{A},3) = \ \text{maximize} \quad & \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} a_{ijk}x_i^1x_j^2x_k^3 \\ \text{subject to} \quad & \|x^i\|_p\le 1,\ x^i\in\mathbb{R}^{n_i} \quad \text{for } i=1,2,3. \end{aligned} \tag{8}$$

Using the definition of $\mathcal{A}(x^1)$ and Hölder's inequality, we can express $v_{\mathsf{ML}}(\mathcal{A},3)$ as

$$\begin{aligned} v_{\mathsf{ML}}(\mathcal{A},3) &= \max_{\|x^1\|_p\le 1}\ \max_{\|x^2\|_p\le 1,\ \|x^3\|_p\le 1} (x^2)^T\mathcal{A}(x^1)x^3 \\ &= \max_{\|x^1\|_p\le 1}\ \max_{\|x^3\|_p\le 1} \|\mathcal{A}(x^1)x^3\|_q \\ &= \max_{\|x^1\|_p\le 1} \|\mathcal{A}(x^1)\|_{p\to q}, \end{aligned} \tag{9}$$

where $q=p/(p-1)$ is the conjugate of p and $\|\mathcal{A}(x^1)\|_{p\to q}$ is the $p\to q$ norm of the $n_2\times n_3$ matrix $\mathcal{A}(x^1)$. From the above derivation, we see that Problem (8) encapsulates two difficult computational tasks: (i) the computation of $\|\mathcal{A}(x^1)\|_{p\to q}$ for any given $x^1\in\mathbb{R}^{n_1}$, and (ii) the maximization of a convex function $x^1\mapsto\|\mathcal{A}(x^1)\|_{p\to q}$ over a convex set $B_p^{n_1}=\{x\in\mathbb{R}^{n_1}:\|x\|_p\le 1\}$. To tackle these difficulties, we proceed in two steps. First, we show that $\|\mathcal{A}(x^1)\|_{p\to q}$ can be approximated by another efficiently computable norm. Then, we show that the maximization of this latter norm over $B_p^{n_1}$ is equivalent to determining the L_q–diameter of a certain convex body, a problem for which approximation algorithms are available. This would in turn yield approximation algorithms for Problem (8).

Step 1: Approximating $\|B\|_{p\to q}$ when $p\in(2,\infty]$. The task of computing $\|B\|_{p\to q}$ for any given $m\times n$ matrix B and $p\in(2,\infty]$ is an instance of the matrix norm problem, which has been extensively studied in the literature. In particular, Nesterov [23] showed that $\|B\|_{p\to q}$ can be approximated to within a factor of $\frac{2\sqrt{3}}{\pi}-\frac{2}{3}>0.435$ via a certain convex relaxation. Later, Ben–Tal and Nemirovski [5] and Steinberg [30] established the NP–hardness of the problem and gave a more refined analysis of Nesterov's relaxation scheme. However, the approximation bound they obtained is better than Nesterov's only when the parameters m,n,p belong to a certain regime. As it turns out, by considering a different convex relaxation, it is possible to obtain an approximation bound that uniformly improves upon that of Nesterov. To demonstrate this, we first observe that

$$\begin{aligned} \|B\|_{p\to q} = \ \text{maximize} \quad & \frac{1}{2}\begin{bmatrix} \mathbf{0} & B \\ B^T & \mathbf{0}\end{bmatrix} \bullet \left(\begin{bmatrix} y \\ z \end{bmatrix} \begin{bmatrix} y^T & z^T \end{bmatrix} \right) \\ \text{subject to} \quad & \|y\|_p\le 1,\ \|z\|_p\le 1, \end{aligned} \tag{10}$$

where $P\bullet Q=\text{tr}(P^TQ)$ denotes the Frobenius inner product of the matrices P and Q. Hence, by introducing the $(m+n)\times(m+n)$ positive semidefinite (psd) matrix X to replace the rank–one psd matrix $(y,z)(y,z)^T$ and denoting

$$\tilde{B}=\frac{1}{2}\begin{bmatrix} \mathbf{0} & B \\ B^T & \mathbf{0}\end{bmatrix},$$

we obtain the following relaxation of $\|B\|_{p\to q}$:

$$\mathsf{vec}_p(B) = \begin{cases} \max\left\{ \tilde{B} \bullet X : \sum_{i=1}^{m} |X_{ii}|^{p/2} \le 1,\ \sum_{j=m+1}^{m+n} |X_{jj}|^{p/2} \le 1,\ X \succeq \mathbf{0} \right\} & \text{for } p \in (2,\infty), \\ \max\left\{ \tilde{B} \bullet X : \max_{1\le i\le m} |X_{ii}| \le 1,\ \max_{m+1\le j\le m+n} |X_{jj}| \le 1,\ X \succeq \mathbf{0} \right\} & \text{for } p = \infty. \end{cases} \tag{11}$$

Note that for $p > 2$, Problem (11) is a convex program that can be solved to arbitrary accuracy in polynomial time using, e.g., the ellipsoid method [10] (cf. [16]). Moreover, the following simple observation of Khot and Naor [15] shows that the ratio between $\mathsf{vec}_p(B)$ and $\|B\|_{p\to q}$ is bounded above by the Grothendieck constant K_G, which is known to be strictly less than $\frac{\pi}{2\ln(1+\sqrt{2})} < 1.783$ [7].

Proposition 3. *The following inequalities hold:*

$$\|B\|_{p\to q} \le \mathsf{vec}_p(B) \le K_G \cdot \|B\|_{p\to q}.$$

For completeness, we include the proof of Proposition 3 here.

Proof. The first inequality follows readily from the fact that Problem (11) is a relaxation of Problem (10). To prove the second inequality, consider an optimal solution X^* to Problem (11) with $\operatorname{rank}(X^*) = r \ge 1$. Let $X^* = V^T V$, where $V \in \mathbb{R}^{r\times(m+n)}$, be the Cholesky factorization of X^*. Furthermore, let $u_i \in \mathbb{R}^r$ (where $i = 1,\dots,m$) and $v_j \in \mathbb{R}^r$ (where $j = 1,\dots,n$) be the i–th column and $(m+j)$–th column of V, respectively. Then, by the optimality of X^*, we have

$$\mathsf{vec}_p(B) = \tilde{B} \bullet X^* = \sum_{i=1}^{m}\sum_{j=1}^{n} B_{ij} u_i^T v_j.$$

Moreover, since $\operatorname{diag}(\tilde{B}) = \mathbf{0}$, we may assume that

$$\|u_i\|_2 = |X^*_{ii}|^{1/2} = 1 \quad \text{for } 1 \le i \le m,$$
$$\|v_j\|_2 = |X^*_{jj}|^{1/2} = 1 \quad \text{for } m+1 \le j \le m+n$$

in the case where $p = \infty$, or

$$\sum_{i=1}^{m} \|u_i\|_2^p = \sum_{i=1}^{m} |X^*_{ii}|^{p/2} = 1, \quad \sum_{j=1}^{n} \|v_j\|_2^p = \sum_{j=m+1}^{m+n} |X^*_{jj}|^{p/2} = 1$$

in the case where $p \in (2,\infty)$. Now, define an $m \times n$ matrix Q by $Q_{ij} = B_{ij} \cdot \|u_i\|_2 \cdot \|v_j\|_2$, where $i = 1,\dots,m$ and $j = 1,\dots,n$. By the Grothendieck inequality (see, e.g., [2, 15]), there exist vectors $\eta \in \{-1,1\}^m$, $\gamma \in \{-1,1\}^n$ such that

$$\mathsf{vec}_p(B) = \sum_{i=1}^{m}\sum_{j=1}^{n} B_{ij} u_i^T v_j = \sum_{i=1}^{m}\sum_{j=1}^{n} Q_{ij} \frac{u_i^T v_j}{\|u_i\|_2 \cdot \|v_j\|_2} \le K_G \sum_{i=1}^{m}\sum_{j=1}^{n} Q_{ij}\eta_i\gamma_j. \tag{12}$$

Upon letting $\bar{y}_i = \eta_i \cdot \|u_i\|_2$ for $i = 1,\dots,m$ and $\bar{z}_j = \gamma_j \cdot \|v_j\|_2$ for $j = 1,\dots,n$, we see that $\|\bar{y}\|_p = \|\bar{z}\|_p = 1$ for $p \in (2,\infty]$; i.e., $(\bar{y},\bar{z}) \in \mathbb{R}^m \times \mathbb{R}^n$ is feasible for Problem (10). Moreover, we obtain from (12) that

$$\mathsf{vec}_p(B) \le K_G \sum_{i=1}^{m}\sum_{j=1}^{n} B_{ij}\bar{y}_i\bar{z}_j \le K_G \cdot \|B\|_{p\to q}.$$

This completes the proof. □

The proof of Proposition 3 reveals that known algorithmic implementations of the Grothendieck inequality (see, e.g., [2, 7]) can be used to deliver vectors $\bar{y} \in \mathbb{R}^m$ and $\bar{z} \in \mathbb{R}^n$ that are feasible for Problem (10) and whose associated objective value $\bar{y}^T B \bar{z}$ is within a constant factor of $\|B\|_{p\to q}$. It should be noted, however, that the precise constant will depend on the particular implementation used. For our purposes, we shall consider two different implementations of the Grothendieck inequality. The first is a deterministic procedure introduced in [2], which is based on the construction of small sample spaces with many four–wise independent random variables. It guarantees that $K_G \le 27$, and hence by Proposition 3 there is a deterministic $(1/27)$–approximation algorithm for computing $\|B\|_{p\to q}$. Although the above procedure does not yield the best approximation bound for $\|B\|_{p\to q}$ (in fact, it is even worse than Nesterov's bound), it will allow us to design a deterministic approximation algorithm for Problem (ML). The second one is based on the so–called Krivine rounding scheme in [7]. The resulting procedure is randomized and guarantees that $K_G < \frac{\pi}{2\ln(1+\sqrt{2})}$, which is currently the best bound on K_G. Consequently, we can approximate $\|B\|_{p\to q}$ to within a factor that is strictly larger than $\frac{2\ln(1+\sqrt{2})}{\pi} > 0.561$, which is better than Nesterov's bound of 0.435.

Based on the above discussion, we summarize our procedure for approximating $\|B\|_{p\to q}$ as follows:

Algorithm 1 Procedure for approximating $\|B\|_{p\to q}$ when $p \in (2,\infty]$ and $q = p/(p-1)$

Input: An $m \times n$ matrix B, a rational number $p \in (2,\infty]$.
Output: A feasible solution $(\bar{y},\bar{z}) \in \mathbb{R}^m \times \mathbb{R}^n$ to Problem (10).

1: Solve the convex relaxation (11) and let $X^* = V^T V$ be an optimal solution. Let $u_1,\dots,u_m$ and $v_1,\dots,v_n$ be the first m and last n columns of V, respectively.
2: Apply either the **deterministic** rounding procedure in [2] or the **randomized** rounding procedure in [7] to the vectors $\{u_i/\|u_i\|_2\}_{i=1}^m$ and $\{v_j/\|v_j\|_2\}_{j=1}^n$ to obtain vectors $\eta \in \{-1,1\}^m$ and $\gamma \in \{-1,1\}^n$ that satisfy (12), where $K_G \le 27$ if the deterministic procedure in [2] is used and $K_G < \frac{\pi}{2\ln(1+\sqrt{2})}$ if the randomized procedure in [7] is used.
3: Set $\bar{y}_i = \eta_i \cdot \|u_i\|_2$ for $i = 1,\dots,m$ and $\bar{z}_j = \gamma_j \cdot \|v_j\|_2$ for $j = 1,\dots,n$. Return $(\bar{y},\bar{z}) \in \mathbb{R}^m \times \mathbb{R}^n$.

Step 2: Norm maximization and diameters of convex bodies. In view of (9) and Proposition 3, we see that any α–approximation to

$$\max_{\|x^1\|_p \le 1} \mathsf{vec}_p\left(\mathcal{A}(x^1)\right) \tag{13}$$

will yield an (α/K_G)–approximation to $v_{\mathsf{ML}}(\mathcal{A},3)$. Hence, it suffices to focus on Problem (13). The following result shows that Problem (13) is in fact equivalent to maximizing a certain norm over the L_p–ball.

Proposition 4. *Let $\mathcal{A} = (a_{ijk}) \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ be an arbitrary non–zero order–3 tensor. Consider the $(n_2 \times n_3) \times n_1$ matrix A given by*

$$A_{(j,k),i} = a_{ijk} \quad \text{for } i = 1,\dots,n_1;\ j = 1,\dots,n_2;\ k = 1,\dots,n_3. \tag{14}$$

Suppose that A has full column rank. Then, the function $x^1 \mapsto \mathsf{vec}_p\left(\mathcal{A}(x^1)\right)$ defines a norm on $\mathbb{R}^{n_1}$.

Proof. Using the definition of $\mathcal{A}(x^1)$ and the derivation in the proof of Proposition 3, we have

$$\begin{aligned}
\mathsf{vec}_p\left(\mathcal{A}(x^1)\right) = \ \text{maximize} \quad & \sum_{i=1}^{n_1} \left(\sum_{j=1}^{n_2} \sum_{k=1}^{n_3} a_{ijk} u_j^T v_k \right) x_i^1 \\
\text{subject to} \quad & \|\mathbf{u}\|_p \le 1,\ \|\mathbf{v}\|_p \le 1, \\
& \mathbf{u} = (\|u_1\|_2,\dots,\|u_{n_2}\|_2) \in \mathbb{R}^{n_2}, \\
& \mathbf{v} = (\|v_1\|_2,\dots,\|v_{n_3}\|_2) \in \mathbb{R}^{n_3}
\end{aligned}$$

for any $p \in (2,\infty]$. In particular, $\mathsf{vec}_p(\mathcal{A}(\cdot))$ is the pointwise supremum of a collection of linear functions, which implies that $\mathsf{vec}_p(\mathcal{A}(\cdot))$ is convex. Moreover, it is clear that $\mathsf{vec}_p\left(\mathcal{A}(kx^1)\right) = |k| \cdot \mathsf{vec}_p\left(\mathcal{A}(x^1)\right)$ for any $k \in \mathbb{R}$ and $x^1 \in \mathbb{R}^{n_1}$, which together with the convexity of $\mathsf{vec}_p(\mathcal{A}(\cdot))$ implies that $\mathsf{vec}_p(\mathcal{A}(\cdot))$ satisfies the triangle inequality. Finally, let $x^1 \in \mathbb{R}^{n_1}\backslash\{\mathbf{0}\}$ be arbitrary. Note that A has full column rank if and only if $\sum_{i=1}^{n_1} a_{ijk}x_i \neq 0$ for some $j = 1,\ldots,n_2$ and $k = 1,\ldots,n_3$ if and only if

$$\mathsf{vec}_p\left(\mathcal{A}(x^1)\right) \geq \max_{1\leq j\leq n_2,\,1\leq k\leq n_3} \left|\sum_{i=1}^{n_1} a_{ijk}x_i\right| > 0.$$

This shows that $x^1 = \mathbf{0}$ whenever $\mathsf{vec}_p\left(\mathcal{A}(x^1)\right) = 0$, and the proof is completed. □

Using the argument in [28, Section 3.1], we may assume without loss that A has full column rank; i.e., $\mathsf{vec}_p(\mathcal{A}(\cdot))$ defines a norm on $\mathbb{R}^{n_1}$. We shall denote this norm by $\|\cdot\|_{\mathcal{A}}$ in the sequel.

To proceed, consider the unit ball of the norm $\|\cdot\|_{\mathcal{A}}$ and its polar, which are given by

$$B_{\mathcal{A}} = \{x \in \mathbb{R}^{n_1} : \|x\|_{\mathcal{A}} \leq 1\}$$

and

$$B_{\mathcal{A}}^{\circ} = \left\{y \in \mathbb{R}^{n_1} : x^T y \leq 1 \text{ for all } x \in B_{\mathcal{A}}\right\},$$

respectively. Note that both $B_{\mathcal{A}}$ and $B_{\mathcal{A}}^{\circ}$ are centrally symmetric and convex. Now, using the dual characterization of norms and Hölder's inequality, we can write Problem (13) as

$$\max_{\|x\|_p\leq 1} \|x\|_{\mathcal{A}} = \max_{\|x\|_p\leq 1} \max_{y\in B_{\mathcal{A}}^{\circ}} x^T y = \max_{y\in B_{\mathcal{A}}^{\circ}} \|y\|_q = \frac{1}{2}\mathrm{diam}_q(B_{\mathcal{A}}^{\circ}), \tag{15}$$

where $q = p/(p-1)$ is the conjugate of p and $\mathrm{diam}_q(B_{\mathcal{A}}^{\circ})$ is the L_q–diameter of $B_{\mathcal{A}}^{\circ}$. In particular, our original problem of approximating $v_{\mathsf{ML}}(\mathcal{A},3)$ (see (8)) is reduced to that of approximating $\mathrm{diam}_q(B_{\mathcal{A}}^{\circ})$, which is well studied in the literature. In the following, we shall present two algorithms for approximating $\mathrm{diam}_q(B_{\mathcal{A}}^{\circ})$. The first is deterministic and implements an idea of Brieden et al. [8]. The second is based on a probabilistic argument of Khot and Naor [14]. Although the latter is randomized, it is much simpler to implement and achieves a better approximation ratio than the former.

5.1.1. Approximating the L_q–diameter of $B_{\mathcal{A}}^{\circ}$ when $q \in [1,2)$.

Deterministic approximation of $\mathrm{diam}_q(B_{\mathcal{A}}^{\circ})$. The key observation underlying the deterministic approximation algorithm is that the diameter of a convex body with respect to a *polytopal norm* can be computed to arbitrary accuracy in deterministic polynomial time under certain conditions [8]. Thus, in order to approximate the L_q–diameter of $B_{\mathcal{A}}^{\circ}$, it suffices to first construct a centrally symmetric polytope $\mathcal{P}$ that approximates the unit L_q–ball, and then compute the diameter of $B_{\mathcal{A}}^{\circ}$ with respect to the polytopal norm induced by $\mathcal{P}$. Before we describe the algorithm in more detail, let us recall some definitions from the algorithmic theory of convex bodies (see [10] for further details). For $p > 2$, let $B_p^n(r) = \{x \in \mathbb{R}^n : \|x\|_p \leq r\}$ denote the n–dimensional L_p–ball centered at the origin with radius $r > 0$. Let $\mathcal{K}$ be a centrally symmetric convex body in $\mathbb{R}^n$. For any $\epsilon \geq 0$, the *outer parallel body* and *inner parallel body* of $\mathcal{K}$ are given by

$$\mathcal{K}(\epsilon) = \mathcal{K} + B_2^n(\epsilon) \quad \text{and} \quad \mathcal{K}(-\epsilon) = \{x \in \mathbb{R}^n : x + B_2^n(\epsilon) \subset \mathcal{K}\},$$

respectively. We say that $\mathcal{K}$ is *well–bounded* if there exist rational numbers $0 < r \leq R < \infty$ such that $B_2^n(r) \subset \mathcal{K} \subset B_2^n(R)$. The *weak membership problem* associated with $\mathcal{K}$ is defined as follows:

WEAK MEMBERSHIP PROBLEM. Given a vector $y \in \mathbb{Q}^n$ and a rational number $\epsilon > 0$, either (i) assert that $y \in \mathcal{K}(\epsilon)$, or (ii) assert that $y \notin \mathcal{K}(-\epsilon)$.

A *weak membership oracle* for $\mathcal{K}$ is a black box that solves the weak membership problem associated with $\mathcal{K}$.

The starting point of our algorithm for approximating $\text{diam}_q(B^\circ_{\mathcal{A}})$ is the following result of Brieden et al. [8]:

THEOREM 4. *Given an integer $n \geq 1$ and a rational number $q \in (1,2]$, one can construct in deterministic polynomial time a centrally symmetric polytope $\mathcal{P}$ in $\mathbb{R}^n$ such that (i) $B^n_q(1) \subset \mathcal{P} \subset B^n_q\left(O\left(n^{1/2}/(\log n)^{1/p}\right)\right)$, where $p = q/(q-1)$ is the conjugate of q, and (ii) for any well–bounded centrally symmetric convex body $\mathcal{K}$ in $\mathbb{R}^n$, one has*

$$\Omega\left(\frac{(\log n)^{1/p}}{n^{1/2}}\right) \cdot diam_q(\mathcal{K}) \leq diam_{\mathcal{P}}(\mathcal{K}) \leq diam_q(\mathcal{K}),$$

where $diam_{\mathcal{P}}(\mathcal{K})$ is the diameter of $\mathcal{K}$ with respect to the polytopal norm $\|\cdot\|_{\mathcal{P}}$ induced by $\mathcal{P}$ (i.e., for any $x \in \mathbb{R}^n$, one has $\|x\|_{\mathcal{P}} = \min\{\lambda \geq 0 : x \in \lambda\mathcal{P}\}$, and $\mathcal{P}$ is the unit ball of the induced norm). Moreover, if $\mathcal{K}$ is equipped with a weak membership oracle, then for any given rational number $\epsilon > 0$, the quantity $diam_{\mathcal{P}}(\mathcal{K})$ can be computed to an accuracy of ϵ in deterministic oracle–polynomial time,[1] and a vector $x \in \mathcal{K}(\epsilon)$ is delivered with $\|x\|_{\mathcal{P}} \geq (1/2) \cdot diam_{\mathcal{P}}(\mathcal{K}) - \epsilon$.

Armed with Theorem 4, we see that in order to design a deterministic polynomial–time algorithm for approximating $\text{diam}_q(B^\circ_{\mathcal{A}})$, it remains to show that $B^\circ_{\mathcal{A}}$ is well–bounded, and that there is a deterministic polynomial–time algorithm for solving the weak membership problem associated with $B^\circ_{\mathcal{A}}$. This is done in the following proposition:

PROPOSITION 5. *Let $\mathcal{A} = (a_{ijk}) \in \mathbb{Q}^{n_1 \times n_2 \times n_3}$ be an arbitrary non–zero order–3 tensor, and let A be the $(n_2 \times n_3) \times n_1$ matrix given by (14). Suppose that A has full column rank. Then, the following hold for the centrally symmetric convex body $B^\circ_{\mathcal{A}}$:*

(a) $B^\circ_{\mathcal{A}}$ is well–bounded. Specifically, there exist rational numbers $0 < r \leq R < \infty$, whose encoding lengths are polynomially bounded by the input size of Problem (8), such that $B^{n_1}_2(r) \subset B^\circ_{\mathcal{A}} \subset B^{n_1}_2(R)$.

(b) The weak membership problem associated with $B^\circ_{\mathcal{A}}$ can be solved in deterministic polynomial time.

Proof. (a) By polarity, we have $B^{n_1}_2(r) \subset B^\circ_{\mathcal{A}} \subset B^{n_1}_2(R)$ if and only if $B^{n_1}_2(1/R) \subset B_{\mathcal{A}} \subset B^{n_1}_2(1/r)$. Thus, it suffices to show that $B_{\mathcal{A}}$ is well–bounded. Now, using the argument in [28, Proposition 2] and the assumption that A has full column rank, one can show that $B^{n_1}_2(r') \subset B_{\mathcal{A}} \subset B^{n_1}_2(R')$, where

$$r' = \frac{1}{\lceil\sqrt{n_1}\rceil \cdot m}, \quad m = \max_{1 \leq i \leq n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} |a_{ijk}|$$

and

$$R' = \left\lceil \sqrt{\frac{n_2 n_3}{\lambda_{min}(A^T A)}} \right\rceil$$

are rational numbers and satisfy $0 < r' \leq R' < \infty$. Moreover, the encoding lengths of r' and R' can be polynomially bounded by the input size of Problem (8); see [10]. This establishes (a).

(b) By the well–boundedness of $B_{\mathcal{A}}$ and the results in [10, Chapter 4], it suffices to show that the weak membership problem associated with $B_{\mathcal{A}}$ can be solved in deterministic polynomial time. However, this follows directly from the argument in [28, Proposition 3] and the observation that $\|x\|_{\mathcal{A}}$ can be computed to arbitrary accuracy in deterministic polynomial time (see (11) and the remarks following it). □

[1] An algorithm has oracle–polynomial time complexity if its runtime is polynomial in both the input size and the number of calls to the oracle [10].

Using (15), Proposition 5 and Theorem 4, we conclude that the optimal value of Problem (13) can be approximated to within a factor of $\Omega\left((\log n_1)^{1/p}/n_1^{1/2}\right)$ in deterministic polynomial time. Thus, by (9) and Proposition 3, the optimal value of Problem (8) can also be approximated to within a factor of $\Omega\left((\log n_1)^{1/p}/n_1^{1/2}\right)$ in deterministic polynomial time. To extract a feasible solution to Problem (8) with the stated approximation guarantee, we just need to unwind our sequence of reductions. For simplicity, let us assume that all computations can be done exactly. Then, by Proposition 5 and Theorem 4, we can find a centrally symmetric polytope $\mathcal{P}$ and a vector $\bar{y}\in B_{\mathcal{A}}^\circ$ such that

$$\|\bar{y}\|_q \ge \|\bar{y}\|_{\mathcal{P}} = \frac{1}{2}\mathrm{diam}_{\mathcal{P}}(B_{\mathcal{A}}^\circ) \ge \Omega\left(\frac{(\log n_1)^{1/p}}{n_1^{1/2}}\right)\cdot \mathrm{diam}_q(B_{\mathcal{A}}^\circ). \tag{16}$$

Now, define the vector $\bar{x}^1\in\mathbb{R}^{n_1}$ by

$$\bar{x}_i^1 = \frac{\mathrm{sgn}(\bar{y}_i)\cdot|\bar{y}_i|^{q-1}}{\|\bar{y}\|_q^{q-1}} \quad \text{for } i=1,\ldots,n_1.$$

It is easy to verify that $\|\bar{x}^1\|_p = 1$ and

$$\mathsf{vec}_p\left(\mathcal{A}(\bar{x}^1)\right) = \|\bar{x}^1\|_{\mathcal{A}} = (\bar{x}^1)^T\bar{y} = \|\bar{y}\|_q. \tag{17}$$

In particular, by applying the deterministic version of Algorithm 1 to the $n_2\times n_3$ matrix $\mathcal{A}(\bar{x}^1)$, we can extract two vectors $\bar{x}^2\in\mathbb{R}^{n_2}$ and $\bar{x}^3\in\mathbb{R}^{n_3}$ such that $\|\bar{x}^2\|_p=\|\bar{x}^3\|_p=1$ and

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} a_{ijk}\bar{x}_i^1\bar{x}_j^2\bar{x}_k^3 \ge \frac{1}{27}\mathsf{vec}_p\left(\mathcal{A}(\bar{x}^1)\right). \tag{18}$$

Finally, since (9), (15) and Proposition 3 together imply

$$\frac{1}{2}\mathrm{diam}_q(B_{\mathcal{A}}^\circ) = \max_{\|x\|_p\le 1}\|x\|_{\mathcal{A}} \ge \max_{\|x^1\|_p\le 1}\|\mathcal{A}(x^1)\|_{p\to q} = v_{\mathsf{ML}}(\mathcal{A},3),$$

we conclude from (16)–(18) that $(\bar{x}^1,\bar{x}^2,\bar{x}^3)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\mathbb{R}^{n_3}$ is an $\Omega\left((\log n_1)^{1/p}/n_1^{1/2}\right)$–approximate solution to Problem (8).

Recall that the above conclusion is obtained under the assumption that all computations are exact. However, it can be shown via a similar but more tedious calculation that the same conclusion holds when the computations are inexact; cf. [28]. Thus, we have proven the following theorem:

Theorem 5. *For any given $p\in(2,\infty]$, there is a deterministic polynomial–time approximation algorithm for Problem (8) with approximation ratio $\Omega\left((\log n_1)^{1/p}/n_1^{1/2}\right)$.*

The following corollary is a direct consequence of Theorems 2 and 5:

Corollary 1. *For $d=3$ and any given $p\in(2,\infty]$, there is a deterministic polynomial–time approximation algorithm for Problem* (HP) *with approximation ratio $\Omega\left((\log n)^{1/p}/n^{1/2}\right)$.*

Randomized approximation of $\mathrm{diam}_q(B_{\mathcal{A}}^\circ)$. In this section, we consider an alternative approach to approximating $\mathrm{diam}_q(B_{\mathcal{A}}^\circ)$, namely, via randomization. The theoretical underpinning of this approach is the following probabilistic results due to Khot and Naor [14]:

Proposition 6. *The following hold:*

(a) Let $\zeta_1, \ldots, \zeta_n$ be i.i.d. Bernoulli random variables and set $\zeta = (\zeta_1, \ldots, \zeta_n) \in \mathbb{R}^n$. Then, there exist universal constants $\delta_0, c_0 > 0$ such that for every $w \in \mathbb{R}^n$,

$$\Pr\left(w^T \zeta \geq \sqrt{\frac{\delta_0 \log n}{n}} \cdot \|w\|_1 \right) \geq \frac{c_0}{n^{\delta_0}}.$$

(b) Suppose that $q \in (1,2)$, and let $p = q/(q-1)$. Let $\xi_1, \ldots, \xi_n$ be i.i.d. random variables with density $p \cdot \exp(-|t|^p)/(2\Gamma(1/p))$ and set $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$.[2] Then, there exist universal constants $\delta_1, c_1, c_2, \bar{n} > 0$ such that for all $n \geq \bar{n}$, we have

$$\Pr\left(\frac{w^T \xi}{\|\xi\|_p} \geq \sqrt{\frac{\delta_1 \log n}{n}} \cdot \|w\|_q \right) \geq \frac{c_1}{n^{c_2}}$$

for every $w \in \mathbb{R}^n$.

Remark 1. An inspection of the proofs in [14] reveals that one can take

$$\delta_0 = \frac{1}{48}, \quad c_0 = \frac{1}{72}, \tag{19}$$

$$\delta_1 = \frac{\mathbb{E}\left[\xi_1^2\right]}{160 \times 2^{2/q}} > \frac{3}{6400}, \quad c_1 = \frac{1}{144}, \quad c_2 = \frac{1}{40}, \quad \bar{n} = 41. \tag{20}$$

Using Proposition 6, we can prove the following result:

Proposition 7. *For any given $q \in [1,2)$, there is a randomized polynomial–time algorithm that returns a vector $v \in \mathbb{R}^{n_1}$ with the following property:*

$$\Pr\left[\Omega\left(\sqrt{\frac{\log n_1}{n_1}} \right) \cdot diam_q(B_{\mathcal{A}}^\circ) \leq 2\|v\|_{\mathcal{A}} \leq diam_q(B_{\mathcal{A}}^\circ) \right] \geq \frac{1}{2}.$$

Proof. Since $B_{\mathcal{A}}^\circ$ is compact and $x \mapsto \|x\|_q$ is continuous, there exists a $\bar{y} \in B_{\mathcal{A}}^\circ$ such that $\|\bar{y}\|_q = \mathrm{diam}_q(B_{\mathcal{A}}^\circ)/2$. We consider two cases:

Case 1: $q = 1$. Let δ_0, c_0 be as in (19) and set $M = (\ln 2) n_1^{\delta_0}/c_0$. Consider a collection $\{\zeta_j^i : i = 1, \ldots, M; j = 1, \ldots, n_1\}$ of i.i.d. Bernoulli random variables. Define

$$\zeta^i = (\zeta_1^i, \ldots, \zeta_{n_1}^i) \in \mathbb{R}^{n_1} \quad \text{for } i = 1, \ldots, M,$$
$$i^* = \arg\max_{1 \leq i \leq M} \|\zeta^i\|_{\mathcal{A}}, \quad v = \zeta^{i^*}, \quad \tau = 2\|v\|_{\mathcal{A}}.$$

We claim that v has the desired property. Indeed, it is clear from (15) that $\tau \leq \mathrm{diam}_1(B_{\mathcal{A}}^\circ)$. Moreover, upon recalling that $\|\zeta^i\|_{\mathcal{A}} = \max_{y \in B_{\mathcal{A}}^\circ} y^T \zeta^i$ and using Proposition 6(a), we have

$$\begin{aligned}
\Pr\left(\tau \geq \sqrt{\frac{\delta_0 \log n_1}{n_1}} \cdot \mathrm{diam}_1(B_{\mathcal{A}}^\circ) \right) &\geq 1 - \Pr\left(\bigcap_{i=1}^M \left\{ \bar{y}^T \zeta^i < \sqrt{\frac{\delta_0 \log n_1}{n_1}} \cdot \|\bar{y}\|_1 \right\} \right) \\
&\geq 1 - \left(1 - \frac{c_0}{n_1^{\delta_0}} \right)^M \\
&\geq \frac{1}{2},
\end{aligned}$$

[2] Recall that $\Gamma : \mathbb{R}_{++} \to \mathbb{R}$ is the gamma function defined by $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt$.

which establishes the claim.

Case 2: $q \in (1,2)$**.** Let δ_1, c_1, c_2 be as in (20) and set $M = (\ln 2) n_1^{c_2}/c_1$. Consider a collection $\{\xi_j^i : i = 1,\ldots,M;\, j = 1,\ldots,n_1\}$ of i.i.d. random variables with density $p \cdot \exp(-|t|^p)/(2\Gamma(1/p))$, where $p = q/(q-1)$. Define

$$\xi^i = (\xi_1^i,\ldots,\xi_{n_1}^i) \in \mathbb{R}^{n_1}, \quad \bar{\xi}^i = \frac{\xi^i}{\|\xi^i\|_p} \quad \text{for } i = 1,\ldots,M,$$
$$i^* = \arg\max_{1 \le i \le M} \|\bar{\xi}^i\|_{\mathcal{A}}, \quad v = \bar{\xi}^{i^*}, \quad \tau = 2\|v\|_{\mathcal{A}}.$$

Using Proposition 6(b) and our previous argument, we have $\tau \le \mathrm{diam}_q(B_{\mathcal{A}}^\circ)$ and

$$\begin{aligned} \Pr\left(\tau \ge \sqrt{\frac{\delta_1 \log n_1}{n_1}} \cdot \mathrm{diam}_q(B_{\mathcal{A}}^\circ)\right) &\ge 1 - \Pr\left(\bigcap_{i=1}^M \left\{ \bar{y}^T \bar{\xi}^i < \sqrt{\frac{\delta_1 \log n_1}{n_1}} \cdot \|\bar{y}\|_q \right\}\right) \\ &\ge 1 - \left(1 - \frac{c_1}{n_1^{c_2}}\right)^M \\ &\ge \frac{1}{2}. \end{aligned}$$

This completes the proof of Proposition 7. □

By combining Proposition 7 with the procedure outlined in the paragraph above Theorem 5, we can extract an $\Omega(\sqrt{\log n_1/n_1})$–approximate solution to Problem (8). Thus, we have proven the following theorem:

Theorem 6. *For any given $p \in (2,\infty]$, there is a randomized polynomial–time approximation algorithm for Problem (8) with approximation ratio $\Omega(\sqrt{\log n_1/n_1})$. In particular, for $d = 3$ and any given $p \in (2,\infty]$, there is a randomized polynomial–time approximation algorithm for Problem* (HP) *with approximation ratio $\Omega(\sqrt{\log n/n})$.*

5.2. General case: Approximating L_p–ball constrained multilinear maximization via recursion. Now, let us consider the problem of maximizing a degree–d multilinear form over L_p–balls, where $d \ge 4$ and $p \in (2,\infty]$ are fixed. Our approach is based on the following simple observation: Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ be an arbitrary non–zero order–d tensor. Then,

$$v_{\mathsf{ML}}(\mathcal{A}, d) = \max_{\|x^1\|_p \le 1} v_{\mathsf{ML}}(\mathcal{A}(x^1), d-1).$$

This suggests that it may be possible to approximate the degree–d problem $v_{\mathsf{ML}}(\mathcal{A}, d)$ if we have an algorithm for approximating the degree–$(d-1)$ problem $v_{\mathsf{ML}}(\mathcal{B}, d-1)$, where $\mathcal{B}$ is an arbitrary non–zero order–$(d-1)$ tensor. To implement this idea, we proceed as follows. Let $\mathcal{H}$ be an arbitrary Hilbert space. Given an arbitrary non–zero order–d tensor $\mathcal{A} = (a_{i_1 i_2 \cdots i_d}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$, let $F_{\mathcal{A}}$ be the associated multilinear form, and define a function $\tilde{F}_{\mathcal{A}} : \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_{d-2}} \times \mathcal{H}^{n_{d-1}} \times \mathcal{H}^{n_d} \to \mathbb{R}$ by

$$\tilde{F}_{\mathcal{A}}\left(x^1,\ldots,x^{d-2},\{u_j\}_{j=1}^{n_{d-1}},\{v_k\}_{k=1}^{n_d}\right) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_{d-2}=1}^{n_{d-2}} \sum_{j=1}^{n_{d-1}} \sum_{k=1}^{n_d} a_{i_1 \cdots i_{d-2} jk} \cdot x_{i_1}^1 \cdots x_{i_{d-2}}^{d-2} \cdot u_j^T v_k.$$

By Proposition 3, for any given $\bar{x}^i \in \mathbb{R}^{n_i}$, where $i = 1,\ldots,d-2$, we have

$$\frac{1}{K_G} \cdot \mathsf{vec}_p\left(\mathcal{A}(\bar{x}^1,\ldots,\bar{x}^{d-2})\right) \le \left\|\mathcal{A}(\bar{x}^1,\ldots,\bar{x}^{d-2})\right\|_{p \to q} \le \mathsf{vec}_p\left(\mathcal{A}(\bar{x}^1,\ldots,\bar{x}^{d-2})\right).$$

Since

$$\left\|\mathcal{A}(\bar{x}^1,\ldots,\bar{x}^{d-2})\right\|_{p\to q} = \max\left\{F_{\mathcal{A}}(\bar{x}^1,\ldots,\bar{x}^{d-2},x^{d-1},x^d) : \|x^{d-1}\|_p \le 1,\ \|x^d\|_p \le 1\right\},$$

it follows that

$$\frac{1}{K_G}\cdot r_{\mathsf{ML}}(\mathcal{A},d) \le \max_{\|x^i\|_p\le 1,\, i=1,\ldots,d} F_{\mathcal{A}}(x^1,\ldots,x^d) = v_{\mathsf{ML}}(\mathcal{A},d) \le r_{\mathsf{ML}}(\mathcal{A},d),$$

where

$$\begin{array}{rll} r_{\mathsf{ML}}(\mathcal{A},d) = & \text{maximize} & \tilde{F}_{\mathcal{A}}\left(x^1,\ldots,x^{d-2},\{u_j\}_{j=1}^{n_{d-1}},\{v_k\}_{k=1}^{n_d}\right) \\ & \text{subject to} & \|x^i\|_p \le 1 \quad \text{for } i=1,\ldots,d-2, \\ & & \|\mathbf{u}\|_p \le 1,\ \|\mathbf{v}\|_p \le 1, \\ & & \mathbf{u} = (\|u_1\|_2,\ldots,\|u_{n_{d-1}}\|_2) \in \mathbb{R}^{n_{d-1}}, \\ & & \mathbf{v} = (\|v_1\|_2,\ldots,\|v_{n_d}\|_2) \in \mathbb{R}^{n_d}. \end{array} \tag{21}$$

In particular, $v_{\mathsf{ML}}(\mathcal{A},d)$ and $r_{\mathsf{ML}}(\mathcal{A},d)$ are equivalent from the approximation perspective. In the sequel, we shall focus on designing approximation algorithms for the latter using both deterministic and randomized approaches.

5.2.1. Deterministic approximation of $r_{\mathsf{ML}}(\mathcal{A},d)$. Our deterministic approach is motivated by the results developed in [28]. Before delving into the details, let us give an overview of the approach. Suppose there is a deterministic algorithm that can approximate the problem $r_{\mathsf{ML}}(\mathcal{B},d-1)$ for any non–zero order–$(d-1)$ tensor $\mathcal{B}$, where $d\ge 4$ is fixed. Then, given an arbitrary $x^1\in\mathbb{R}^{n_1}$, since $\mathcal{A}(x^1)$ is an order–$(d-1)$ tensor, we can apply the algorithm to the problem

$$\begin{array}{rll} r_{\mathsf{ML}}(\mathcal{A}(x^1),d-1) = & \text{maximize} & \tilde{F}_{\mathcal{A}}\left(x^1,\ldots,x^{d-2},\{u_j\}_{j=1}^{n_{d-1}},\{v_k\}_{k=1}^{n_d}\right) \\ & \text{subject to} & \|x^i\|_p \le 1 \quad \text{for } i=2,\ldots,d-2, \\ & & \|\mathbf{u}\|_p \le 1,\ \|\mathbf{v}\|_p \le 1, \\ & & \mathbf{u} = (\|u_1\|_2,\ldots,\|u_{n_{d-1}}\|_2) \in \mathbb{R}^{n_{d-1}}, \\ & & \mathbf{v} = (\|v_1\|_2,\ldots,\|v_{n_d}\|_2) \in \mathbb{R}^{n_d} \end{array}$$

and obtain a value $G_{d-1}(x^1)$ that satisfies $\beta_{d-1}\cdot r_{\mathsf{ML}}(\mathcal{A}(x^1),d-1) \le G_{d-1}(x^1) \le r_{\mathsf{ML}}(\mathcal{A}(x^1),d-1)$, where $\beta_{d-1}\in(0,1)$ is the approximation ratio of the algorithm. Since this holds for any $x^1\in\mathbb{R}^{n_1}$, it follows that

$$\beta_{d-1}\cdot r_{\mathsf{ML}}(\mathcal{A},d) \le \max_{\|x^1\|_p\le 1} G_{d-1}(x^1) \le r_{\mathsf{ML}}(\mathcal{A},d).$$

Now, if we can show that the function G_{d-1} defines a norm on $\mathbb{R}^{n_1}$, then $\max_{\|x^1\|_p\le 1} G_{d-1}(x^1)$ is a norm maximization problem, which can be approximated using the techniques outlined in Section 5.1. This would then yield an approximation algorithm for the problem $r_{\mathsf{ML}}(\mathcal{A},d)$.

To carry out this plan, we need the following result:

PROPOSITION 8. *Let $d\ge 3$ and $p\in(2,\infty]$ be given. For $i=1,\ldots,d-3$, let $\mathcal{P}_i$ be a centrally symmetric polytope in $\mathbb{R}^{n_{i+1}}$ satisfying the properties stated in Theorem 4. Furthermore, let $\mathcal{A}=(a_{i_1i_2\cdots i_d})\in\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$ be an arbitrary non–zero order–d tensor. Define the functions $\Lambda_i^{\mathcal{A},d}:\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\cdots\times\mathbb{R}^{n_i}\to\mathbb{R}_+$ for $i=1,\ldots,d-2$ inductively as follows:*

$$\begin{aligned} \Lambda_{d-2}^{\mathcal{A},d}(x^1,x^2,\ldots,x^{d-2}) &= \mathsf{vec}_p\left(\mathcal{A}(x^1,x^2,\ldots,x^{d-2})\right), \\ \Lambda_i^{\mathcal{A},d}(x^1,x^2,\ldots,x^i) &= diam_{\mathcal{P}_i}\left[\left\{y\in\mathbb{R}^{n_{i+1}} : \Lambda_{i+1}^{\mathcal{A},d}(x^1,x^2,\ldots,x^i,y)\le 1\right\}^\circ\right] \end{aligned} \tag{22}$$

for $i=d-3,d-4,\ldots,1$. Then, the following hold:

(a) For $j=1,\ldots,d-2$ and for any $\bar{x}^1,\ldots,\bar{x}^{k-1},\bar{x}^{k+1},\ldots,\bar{x}^j$, where $\bar{x}^i\in\mathbb{R}^{n_i}$, the function $\bar{\Lambda}_{j,k}^{\mathcal{A},d}:\mathbb{R}^{n_k}\to\mathbb{R}_+$ given by

$$\bar{\Lambda}_{j,k}^{\mathcal{A},d}(x)=\Lambda_j^{\mathcal{A},d}(\bar{x}^1,\ldots,\bar{x}^{k-1},x,\bar{x}^{k+1},\ldots,\bar{x}^j)$$

is a semi–norm on $\mathbb{R}^{n_k}$ for any $k\in\{1,\ldots,j\}$.

(b) Let A be the $(n_2\times\cdots\times n_d)\times n_1$ matrix given by

$$A_{(i_2,\ldots,i_d),i_1}=a_{i_1i_2\cdots i_d}\quad \text{for } i_j=1,\ldots,n_j;\ j=1,\ldots,d. \tag{23}$$

Suppose that A has full column rank. Then, the function $\Lambda_1^{\mathcal{A},d}$ defines a norm on $\mathbb{R}^{n_1}$.

(c) We have

$$\Lambda_{i-1}^{\mathcal{A}(x^1),d-1}(x^2,x^3,\ldots,x^i)=\Lambda_i^{\mathcal{A},d}(x^1,x^2,\ldots,x^i)\quad \text{for } i=2,3,\ldots,d-2.$$

Proof. Both (a) and (b) are essentially adaptations of the corresponding claims in [28, Proposition 4]. To prove (c), we proceed by backward induction on i. For $i=d-2$, we have, by definition,

$$\begin{aligned}\Lambda_{d-3}^{\mathcal{A}(x^1),d-1}(x^2,x^3,\ldots,x^{d-2}) &= \mathsf{vec}_p\left([\mathcal{A}(x^1)](x^2,x^3,\ldots,x^{d-2})\right)\\ &= \mathsf{vec}_p\left(\mathcal{A}(x^1,x^2,\ldots,x^{d-2})\right)\\ &= \Lambda_{d-2}^{\mathcal{A},d}(x^1,x^2,\ldots,x^{d-2}).\end{aligned}$$

For the inductive step, we use both the definition in (22) and the inductive hypothesis to obtain

$$\begin{aligned}\Lambda_{i-1}^{\mathcal{A}(x^1),d-1}(x^2,x^3,\ldots,x^i) &= \mathrm{diam}_{\mathcal{P}_i}\left[\left\{y\in\mathbb{R}^{n_{i+1}}:\Lambda_i^{\mathcal{A}(x^1),d-1}(x^2,x^3,\ldots,x^i,y)\le 1\right\}^\circ\right]\\ &= \mathrm{diam}_{\mathcal{P}_i}\left[\left\{y\in\mathbb{R}^{n_{i+1}}:\Lambda_{i+1}^{\mathcal{A},d}(x^1,x^2,\ldots,x^i,y)\le 1\right\}^\circ\right]\\ &= \Lambda_i^{\mathcal{A},d}(x^1,x^2,\ldots,x^i).\end{aligned}$$

This completes the proof. $\square$

We are now ready to prove the main result of this section:

THEOREM 7. *Let $d\ge 3$ and $p\in(2,\infty]$ be given. Let $\mathcal{A}=(a_{i_1i_2\cdots i_d})\in\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$ be an arbitrary non–zero order–d tensor. Consider the functions $\{\Lambda_i^{\mathcal{A},d}\}_{i=1}^{d-2}$ defined in (22) and the $(n_2\times\cdots\times n_d)\times n_1$ matrix A defined in (23). Suppose that A has full column rank. Then, the following hold:*

(a) For any given $x\in\mathbb{R}^{n_1}$, the norm $\Lambda_1^{\mathcal{A},d}(x)$ is efficiently computable*; i.e., it can be computed to any desired accuracy by a deterministic algorithm whose runtime is polynomial in the input size of Problem (21) and the level of accuracy.*

(b) There exist rational numbers $0<r\le R<\infty$, whose encoding lengths are polynomially bounded by the input size of Problem (21), such that

$$B_2^{n_1}(r)\subset\left\{x\in\mathbb{R}^{n_1}:\Lambda_1^{\mathcal{A},d}(x)\le 1\right\}\subset B_2^{n_1}(R).$$

Consequently, the quantity $\mathrm{diam}_{\mathcal{P}_0}\left(\left\{x\in\mathbb{R}^{n_1}:\Lambda_1^{\mathcal{A},d}(x)\le 1\right\}^\circ\right)$ can be efficiently computed, where $\mathcal{P}_0$ is a centrally symmetric polytope in $\mathbb{R}^{n_1}$ satisfying the properties stated in Theorem 4.

(c) We have

$$\Omega\left(\prod_{i=1}^{d-2}\frac{(\log n_i)^{1/p}}{n_i^{1/2}}\right)\cdot r_{\mathsf{ML}}(\mathcal{A},d)\le\frac{1}{2}\mathrm{diam}_{\mathcal{P}_0}\left(\left\{x\in\mathbb{R}^{n_1}:\Lambda_1^{\mathcal{A},d}(x)\le 1\right\}^\circ\right)\le r_{\mathsf{ML}}(\mathcal{A},d).$$

In particular, there is a deterministic polynomial–time algorithm for Problem (ML) *with approximation ratio* $\Omega\left(\prod_{i=1}^{d-2}(\log n_i)^{1/p}/n_i^{1/2}\right)$.

Proof. We proceed by induction on $d \ge 3$. The base case follows from (15), Proposition 5 and Theorem 4. Now, suppose that $d > 3$. Let $x_1 \in \mathbb{R}^{n_1}\backslash\{\mathbf{0}\}$ be arbitrary, and consider the order–$(d-1)$ tensor $\mathcal{A}(x^1) \in \mathbb{R}^{n_2\times n_3\times\cdots\times n_d}$. Without loss of generality, we may assume that the $(n_3\times\cdots\times n_d)\times n_2$ matrix $A(x^1)$, where $[A(x^1)]_{(i_3,\ldots,i_d),i_2} = [\mathcal{A}(x^1)]_{i_2i_3\cdots i_d}$, has full column rank. By the inductive hypothesis, $\Lambda_1^{\mathcal{A}(x^1),d-1}$ is an efficiently computable norm on $\mathbb{R}^{n_2}$ and the set $\left\{x\in\mathbb{R}^{n_2} : \Lambda_1^{\mathcal{A}(x^1),d-1}(x)\le 1\right\}$ is well–bounded. Moreover, using (22) and Proposition 8(c), we have

$$\Lambda_1^{\mathcal{A},d}(x^1) = \mathrm{diam}_{\mathcal{P}_1}\left[\left\{x\in\mathbb{R}^{n_2} : \Lambda_2^{\mathcal{A},d}(x^1,x)\le 1\right\}^\circ\right] = \mathrm{diam}_{\mathcal{P}_1}\left[\left\{x\in\mathbb{R}^{n_2} : \Lambda_1^{\mathcal{A}(x^1),d-1}(x)\le 1\right\}^\circ\right].$$

Hence, by arguing as in the proof of Proposition 5 and applying Theorem 4, we conclude that $\Lambda_1^{\mathcal{A},d}$ is an efficiently computable norm on $\mathbb{R}^{n_1}$.

Let $B_{\Lambda_1^{\mathcal{A},d}} = \left\{x\in\mathbb{R}^{n_1} : \Lambda_1^{\mathcal{A},d}(x)\le 1\right\}$ be the unit ball of $\Lambda_1^{\mathcal{A},d}$. Using the argument in the proof of [28, Theorem 4], one can show that $B_{\Lambda_1^{\mathcal{A},d}}$ is well–bounded. As a corollary, we see that $B^\circ_{\Lambda_1^{\mathcal{A},d}}$ is also well–bounded, and that the weak membership problem associated with $B^\circ_{\Lambda_1^{\mathcal{A},d}}$ can be solved in deterministic polynomial time. This implies that $\mathrm{diam}_{\mathcal{P}_0}\left(\left\{x\in\mathbb{R}^{n_1} : \Lambda_1^{\mathcal{A},d}(x)\le 1\right\}^\circ\right)$ can be efficiently computed.

Now, the inductive hypothesis, the definition of $\Lambda_1^{\mathcal{A},d}$ in (22) and Proposition 8(c) yield

$$\begin{aligned}\Omega\left(\prod_{i=2}^{d-2}\frac{(\log n_i)^{1/p}}{n_i^{1/2}}\right)\cdot r_{\mathsf{ML}}(\mathcal{A}(x^1),d-1) &\le \frac{1}{2}\mathrm{diam}_{\mathcal{P}_1}\left(\left\{x\in\mathbb{R}^{n_2} : \Lambda_1^{\mathcal{A}(x^1),d-1}(x)\le 1\right\}^\circ\right)\\ &= \frac{1}{2}\Lambda_1^{\mathcal{A},d}(x^1)\\ &\le r_{\mathsf{ML}}(\mathcal{A}(x^1),d-1).\end{aligned}$$

Since $r_{\mathsf{ML}}(\mathcal{A},d) = \max_{\|x^1\|_p\le 1} r_{\mathsf{ML}}(\mathcal{A}(x^1),d)$, it follows that

$$\Omega\left(\prod_{i=2}^{d-2}\frac{(\log n_i)^{1/p}}{n_i^{1/2}}\right)\cdot r_{\mathsf{ML}}(\mathcal{A},d) \le \frac{1}{2}\max_{\|x^1\|_p\le 1}\Lambda_1^{\mathcal{A},d}(x^1) \le r_{\mathsf{ML}}(\mathcal{A},d). \tag{24}$$

By mimicking the derivation of (15), one can show that

$$\max_{\|x^1\|_p\le 1}\Lambda_1^{\mathcal{A},d}(x^1) = \frac{1}{2}\mathrm{diam}_q\left(B^\circ_{\Lambda_1^{\mathcal{A},d}}\right). \tag{25}$$

Moreover, since $B^\circ_{\Lambda_1^{\mathcal{A},d}}$ is well–bounded, Theorem 4 and the definition of $\mathcal{P}_0$ imply that

$$\Omega\left(\frac{(\log n_1)^{1/p}}{n_1^{1/2}}\right)\cdot \mathrm{diam}_q\left(B^\circ_{\Lambda_1^{\mathcal{A},d}}\right) \le \mathrm{diam}_{\mathcal{P}_0}\left(B^\circ_{\Lambda_1^{\mathcal{A},d}}\right) \le \mathrm{diam}_q\left(B^\circ_{\Lambda_1^{\mathcal{A},d}}\right). \tag{26}$$

It then follows from (24)–(26) that

$$\Omega\left(\prod_{i=1}^{d-2}\frac{(\log n_i)^{1/p}}{n_i^{1/2}}\right)\cdot r_{\mathsf{ML}}(\mathcal{A},d) \le \frac{1}{2}\mathrm{diam}_{\mathcal{P}_0}\left(\left\{x\in\mathbb{R}^{n_1} : \Lambda_1^{\mathcal{A},d}(x)\le 1\right\}^\circ\right) \le r_{\mathsf{ML}}(\mathcal{A},d).$$

This completes the proof of Theorem 7. □

The following is an immediate corollary of Theorems 2 and 7:

COROLLARY 2. *For any given $d \geq 3$ and $p \in (2,\infty]$, there is a deterministic polynomial–time algorithm for* (HP) *with approximation ratio (resp. relative approximation ratio) $\Omega\left((\log n)^{(d-2)/p}/n^{d/2-1}\right)$ when $d \geq 3$ is odd (resp. even).*

5.2.2. Randomized approximation of $r_{\mathsf{ML}}(\mathcal{A},d)$. As in the case where $d=3$, we can approximate $r_{\mathsf{ML}}(\mathcal{A},d)$ using a randomized approach. Such an approach is based on the following result, which states that every optimal solution to Problem (21) satisfies certain probability inequality:

PROPOSITION 9. *Let*

$$\left(\bar{x}^1,\ldots,\bar{x}^{d-2},\{\bar{u}_j\}_{j=1}^{n_{d-1}},\{\bar{v}_k\}_{k=1}^{n_d}\right)$$

be an optimal solution to Problem (21).

(a) Let $\zeta \in \mathbb{R}^{n_1}$ be a vector of i.i.d. Bernoulli random variables. Then,

$$\Pr\left(\tilde{F}_{\mathcal{A}}\left(\zeta,\bar{x}^2,\ldots,\bar{x}^{d-2},\{\bar{u}_j\}_{j=1}^{n_{d-1}},\{\bar{v}_k\}_{k=1}^{n_d}\right) \geq \sqrt{\frac{\delta_0\log n_1}{n_1}}\cdot r_{\mathsf{ML}}(\mathcal{A},d)\right) \geq \frac{c_0}{n_1^{\delta_0}},$$

where the constants δ_0, c_0 are given by (19).

(b) Suppose that $p \in (2,\infty)$. Let $\xi \in \mathbb{R}^{n_1}$ be a vector of i.i.d. random variables with density $p\cdot\exp(-|t|^p)/(2\Gamma(1/p))$, and set $\bar{\xi} = \xi/\|\xi\|_p$. Then,

$$\Pr\left(\tilde{F}_{\mathcal{A}}\left(\bar{\xi},\bar{x}^2,\ldots,\bar{x}^{d-2},\{\bar{u}_j\}_{j=1}^{n_{d-1}},\{\bar{v}_k\}_{k=1}^{n_d}\right) \geq \sqrt{\frac{\delta_1\log n_1}{n_1}}\cdot r_{\mathsf{ML}}(\mathcal{A},d)\right) \geq \frac{c_1}{n_1^{c_2}}$$

for all $n \geq \bar{n}$, where the constants $\delta_1, c_1, c_2, \bar{n}$ are given by (20).

Proof. Let $w \in \mathbb{R}^{n_1}$ be the vector defined by

$$w_{i_1} = \sum_{i_2=1}^{n_2}\cdots\sum_{i_{d-2}=1}^{n_{d-2}}\sum_{j=1}^{n_{d-1}}\sum_{k=1}^{n_d} a_{i_1 i_2\cdots i_{d-2}jk}\cdot\bar{x}_{i_2}^2\cdots\bar{x}_{i_{d-2}}^{d-2}\cdot\bar{u}_j^T\bar{v}_k \quad \text{for } i_1 = 1,\ldots,n_1.$$

Then, for any $x \in \mathbb{R}^{n_1}$, we have

$$w^Tx = \tilde{F}_{\mathcal{A}}\left(x,\bar{x}^2,\ldots,\bar{x}^{d-2},\{\bar{u}_j\}_{j=1}^{n_{d-1}},\{\bar{v}_k\}_{k=1}^{n_d}\right).$$

Moreover, by the definition of $r_{\mathsf{ML}}(\mathcal{A},d)$ and Hölder's inequality, we have

$$\begin{aligned} r_{\mathsf{ML}}(\mathcal{A},d) &= \tilde{F}_{\mathcal{A}}\left(\bar{x}^1,\bar{x}^2,\ldots,\bar{x}^{d-2},\{\bar{u}_j\}_{j=1}^{n_{d-1}},\{\bar{v}_k\}_{k=1}^{n_d}\right) \\ &= \max_{\|x^1\|_p\leq 1}\tilde{F}_{\mathcal{A}}\left(x^1,\bar{x}^2,\ldots,\bar{x}^{d-2},\{\bar{u}_j\}_{j=1}^{n_{d-1}},\{\bar{v}_k\}_{k=1}^{n_d}\right) \\ &= \|w\|_q. \end{aligned}$$

Thus, the desired result follows from Proposition 6. □

THEOREM 8. *For any given $d \geq 3$ and $p \in (2,\infty]$, there is a randomized polynomial–time algorithm for Problem (21) that returns vectors $\hat{x}^1,\ldots,\hat{x}^{d-2},\{\hat{u}_j\}_{j=1}^{n_{d-1}},\{\hat{v}_k\}_{k=1}^{n_d}$ with the following property:*

$$\Pr\left[\kappa^{d/2-1}\prod_{i=1}^{d-2}\sqrt{\frac{\log n_i}{n_i}}\cdot r_{\mathsf{ML}}(\mathcal{A},d) \leq G(\mathcal{A},d) \leq r_{\mathsf{ML}}(\mathcal{A},d)\right] \geq \frac{1}{2}.$$

Here, $G(\mathcal{A},d)=\tilde{F}_{\mathcal{A}}\left(\hat{x}^1,\ldots,\hat{x}^{d-2},\{\hat{u}_j\}_{j=1}^{n_{d-1}},\{\hat{v}_k\}_{k=1}^{n_d}\right)$ *and*

$$\kappa=\begin{cases}\delta_0 & \text{if } p\in(2,\infty),\\ \delta_1 & \text{if } p=\infty,\end{cases}$$

where the constants δ_0,δ_1 *are given by (19) and (20), respectively. In particular, there is a randomized polynomial–time algorithm for Problem* (ML) *with approximation ratio* $\Omega\left(\prod_{i=1}^{d-2}\sqrt{\log n_i/n_i}\right)$.

Remark 2. In the recent work [11], He et al. established a result similar to Theorem 8, but only for the case where $p\in\{2,\infty\}$. We note that the proof of Theorem 8 can be easily extended to cover the case where $p=2$. However, we shall not pursue such extension here, as there already exists a *deterministic* $\Omega\left(\prod_{i=1}^{d-2}\sqrt{\log n_i/n_i}\right)$–approximation algorithm for Problem (ML) when $p=2$ [28].

Proof. We shall prove the theorem only for the case where $p\in(2,\infty)$; the case where $p=\infty$ will be similar. The proof proceeds by induction on $d\geq 3$. The base case follows from Proposition 7. Now, set $M=(2\ln 2)n_1^{c_2}/c_1$, where the constants c_1,c_2 are given by (20). Consider a collection $\{\xi_j^i : i=1,\ldots,M;\ j=1,\ldots,n_1\}$ of i.i.d. random variables with density $p\cdot\exp(-|t|^p)/(2\Gamma(1/p))$. Define

$$\xi^i=(\xi_1^i,\ldots,\xi_{n_1}^i)\in\mathbb{R}^{n_1},\quad \bar{\xi}^i=\frac{\xi^i}{\|\xi^i\|_p}\quad\text{for } i=1,\ldots,M.$$

By the inductive hypothesis, there is a randomized polynomial–time algorithm that can compute, for each $i=1,\ldots,M$, a number $G(\mathcal{A}(\bar{\xi}^i),d-1)$ satisfying

$$\Pr\left[\delta_1^{(d-3)/2}\prod_{i=2}^{d-2}\sqrt{\frac{\log n_i}{n_i}}\cdot r_{\mathsf{ML}}(\mathcal{A}(\bar{\xi}^i),d-1)\leq G(\mathcal{A}(\bar{\xi}^i),d-1)\leq r_{\mathsf{ML}}(\mathcal{A}(\bar{\xi}^i),d-1)\right]\geq\frac{1}{2}.$$

Now, consider the events

$$E_i=\left\{G(\mathcal{A}(\bar{\xi}^i),d-1)\geq\delta_1^{(d-3)/2}\prod_{i=2}^{d-2}\sqrt{\frac{\log n_i}{n_i}}\cdot r_{\mathsf{ML}}(\mathcal{A}(\bar{\xi}^i),d-1)\right\}\quad\text{for } i=1,\ldots,M$$

and let

$$\left(\bar{x}^1,\ldots,\bar{x}^{d-2},\{\bar{u}_j\}_{j=1}^{n_{d-1}},\{\bar{v}_k\}_{k=1}^{n_d}\right)$$

be an optimal solution to Problem (21). Note that $\Pr(E_i)\geq 1/2$ for $i=1,\ldots,M$. We compute

$$\begin{aligned}
&\Pr\left(G(\mathcal{A}(\bar{\xi}^i),d-1)\geq\delta_1^{d/2-1}\prod_{i=1}^{d-2}\sqrt{\frac{\log n_i}{n_i}}\cdot r_{\mathsf{ML}}(\mathcal{A},d)\right)\\
\geq\ &\Pr\left(G(\mathcal{A}(\bar{\xi}^i),d-1)\geq\delta_1^{d/2-1}\prod_{i=1}^{d-2}\sqrt{\frac{\log n_i}{n_i}}\cdot r_{\mathsf{ML}}(\mathcal{A},d)\,\middle|\,E_i\right)\times\Pr(E_i)\\
\geq\ &\frac{1}{2}\Pr\left(r_{\mathsf{ML}}(\mathcal{A}(\bar{\xi}^i),d-1)\geq\sqrt{\frac{\delta_1\log n_1}{n_1}}\cdot r_{\mathsf{ML}}(\mathcal{A},d)\right) &(27)\\
\geq\ &\frac{1}{2}\Pr\left(\tilde{F}_{\mathcal{A}}\left(\bar{\xi}^i,\bar{x}^2,\ldots,\bar{x}^{d-2},\{\bar{u}_j\}_{j=1}^{n_{d-1}},\{\bar{v}_k\}_{k=1}^{n_d}\right)\geq\sqrt{\frac{\delta_1\log n_1}{n_1}}\cdot r_{\mathsf{ML}}(\mathcal{A},d)\right)\\
\geq\ &\frac{c_1}{2n_1^{c_2}}, &(28)
\end{aligned}$$

where (27) follows from the fact that $\bar{\xi}^i$ is independent of the randomizations used to compute $G(\mathcal{A}(\bar{\xi}^i), d-1)$, and (28) follows from Proposition 9(b). Upon setting

$$G(\mathcal{A}, d) = \max_{1 \le i \le M} G(\mathcal{A}(\bar{\xi}^i), d-1),$$

we conclude that

$$\begin{aligned}
&\Pr\left(G(\mathcal{A}, d) \ge \delta_1^{d/2-1} \prod_{i=1}^{d-2} \sqrt{\frac{\log n_i}{n_i}} \cdot r_{\mathsf{ML}}(\mathcal{A}, d) \right) \\
&\ge 1 - \prod_{i=1}^{M} \Pr\left(G(\mathcal{A}(\bar{\xi}^i), d-1) < \delta_1^{d/2-1} \prod_{i=1}^{d-2} \sqrt{\frac{\log n_i}{n_i}} \cdot r_{\mathsf{ML}}(\mathcal{A}, d) \right) \\
&\ge 1 - \left(1 - \frac{c_1}{2n_1^{c_2}} \right)^M \\
&\ge \frac{1}{2}.
\end{aligned}$$

This completes the proof. $\square$

Corollary 3. *For any given $d \ge 3$ and $p \in (2, \infty]$, there is a randomized polynomial–time algorithm for Problem* (HP) *with approximation ratio (resp. relative approximation ratio) $\Omega\left((\log n/n)^{d/2-1}\right)$ when d is odd (resp. even).*

6. Conclusion. In this paper, we studied the hardness and approximability of homogeneous polynomial optimization and related multilinear optimization problems with L_p–ball constraints. A crucial first step in our proofs is to relate the polynomial optimization problem at hand to a suitable multilinear optimization problem. To obtain approximation results, we further showed that the L_p–ball constrained multilinear optimization problem is equivalent, from an approximation perspective, to that of determining the diameters of certain convex bodies. Such equivalence was established using the Grothendieck inequality (see, e.g., [15, 25]) and an argument of Khot and Naor [14] (cf. [28]). Consequently, by extending the approaches in [14, 28] and applying results from algorithmic convex geometry, we were able to develop both deterministic and randomized polynomial–time approximation algorithms for various L_p–ball constrained polynomial optimization problems, whose approximation guarantees are currently the best known in the literature. We believe that the wide array of tools used in this paper will have further applications in the study of polynomial optimization problems. As an immediate illustration, consider the following variant of Problem (ML):

$$(\widehat{\mathsf{ML}}) \qquad \begin{array}{ll} \text{maximize} & F_{\mathcal{A}}(x^1, x^2, \ldots, x^d) \equiv \displaystyle\sum_{1 \le i_1, \ldots, i_d \le n} a_{i_1 i_2 \cdots i_d} x^1_{i_1} x^2_{i_2} \cdots x^d_{i_d} \\ \text{subject to} & \|x^i\|_{p_i} \le 1, \; x^i \in \mathbb{R}^{n_i} \quad \text{for } i = 1, \ldots, d, \end{array}$$

where $2 \le p_1 \le p_2 \le \cdots \le p_d \le \infty$. By a straightforward modification of the arguments in Section 5, it can be shown that Problem $(\widehat{\mathsf{ML}})$ admits a deterministic polynomial–time approximation algorithm with approximation ratio $\Omega\left(\prod_{i=1}^{d-2} (\log n_i)^{1/p_i} / n_i^{1/2}\right)$, as well as a randomized polynomial–time approximation algorithm with approximation ratio $\Omega\left(\prod_{i=1}^{d-2} \sqrt{\log n_i / n_i}\right)$. Finally, it would be interesting to find more applications of the optimization models studied in this paper.

Appendix A: Proof of Theorem 2. The proof of Theorem 2 relies on the following polarization formula, whose proof can be found, e.g., in [12, Lemma 3.5]:

PROPOSITION 10. *Let $x^1, x^2, \ldots, x^d \in \mathbb{R}^n$ be arbitrary, and let $\xi_1, \xi_2, \ldots, \xi_d$ be i.i.d. Bernoulli random variables (i.e., $\Pr(\xi_i = 1) = \Pr(\xi_i = -1) = 1/2$ for $i = 1, \ldots, d$). Then, we have*

$$\mathbb{E}\left[\left(\prod_{i=1}^{d}\xi_i\right) f_{\mathcal{A}}\left(\sum_{j=1}^{d}\xi_j x^j\right)\right] = d! \cdot F_{\mathcal{A}}(x^1, x^2, \ldots, x^d). \tag{29}$$

Armed with Proposition 10, we proceed as follows. Let $(x^1, x^2, \ldots, x^d)$ be the feasible solution to Problem (MR) returned by $\mathscr{A}_{\mathsf{MR}}$. By assumption, we have $\|x^i\|_p \le 1$ for $i = 1, \ldots, d$ and $F_{\mathcal{A}}(x^1, x^2, \ldots, x^d) \ge \alpha v^*$. When $d \ge 3$ is odd, we can rewrite (29) as

$$d! \cdot F_{\mathcal{A}}(x^1, x^2, \ldots, x^d) = \mathbb{E}\left[\left(\prod_{i=1}^{d}\xi_i\right) f_{\mathcal{A}}\left(\sum_{j=1}^{d}\xi_j x^j\right)\right] = \mathbb{E}\left[f_{\mathcal{A}}\left(\sum_{j=1}^{d}\left(\prod_{i\ne j}\xi_i\right) x^j\right)\right].$$

In particular, since $d \ge 3$ is assumed to be fixed, we can find in constant time a vector $\beta = (\beta_1, \beta_2, \ldots, \beta_d) \in \{-1, 1\}^d$ that satisfies

$$f_{\mathcal{A}}\left(\sum_{j=1}^{d}\left(\prod_{i\ne j}\beta_i\right) x^j\right) \ge d! \cdot F_{\mathcal{A}}(x^1, x^2, \ldots, x^d).$$

Now, set $\hat{x} = \sum_{j=1}^{d}\left(\prod_{i\ne j}\beta_i\right) x^j \Big/ \left\|\sum_{j=1}^{d}\left(\prod_{i\ne j}\beta_i\right) x^j\right\|_p$. Then, we have $\|\hat{x}\|_p = 1$; i.e., it is feasible for Problem (HP). Moreover, since

$$\left\|\sum_{j=1}^{d}\left(\prod_{i\ne j}\beta_i\right) x^j\right\|_p \le \sum_{j=1}^{d}\|x^j\|_p \le d,$$

we conclude that

$$f_{\mathcal{A}}(\hat{x}) \ge \frac{d! F_{\mathcal{A}}(x^1, x^2, \ldots, x^d)}{\left\|\sum_{j=1}^{d}\left(\prod_{i\ne j}\beta_i\right) x^j\right\|_p^d} \ge \alpha \cdot d! \cdot d^{-d} \cdot v^* \ge \alpha \cdot d! \cdot d^{-d} \cdot \bar{v},$$

as required.

Next, consider the case when $d \ge 4$ is even. Observe that every realization of the random vector $\xi = (\xi_1, \xi_2, \ldots, \xi_d) \in \{-1, 1\}^d$ satisfies

$$\left\|\frac{1}{d}\sum_{j=1}^{d}\xi_j x^j\right\|_p \le \frac{1}{d}\sum_{j=1}^{d}\|x^j\|_p \le 1;$$

i.e., $\frac{1}{d}\sum_{j=1}^{d}\xi_j x^j$ is feasible for Problem (HP). Now, using the identity (29), we compute

$$\begin{aligned} d! \cdot F_{\mathcal{A}}(x^1, x^2, \ldots, x^d) &= \mathbb{E}\left[\left(\prod_{i=1}^{d}\xi_i\right) f_{\mathcal{A}}\left(\sum_{j=1}^{d}\xi_j x^j\right)\right] \\ &= \frac{d^d}{2}\mathbb{E}\left[f_{\mathcal{A}}\left(\frac{1}{d}\sum_{j=1}^{d}\xi_j x^j\right) - \underline{v}\,\middle|\, \prod_{i=1}^{d}\xi_i = 1\right] \end{aligned}$$

$$\begin{aligned} & - \frac{d^d}{2}\mathbb{E}\left[f_{\mathcal{A}}\left(\frac{1}{d}\sum_{j=1}^d \xi_j x^j\right) - \underline{v} \,\middle|\, \prod_{i=1}^d \xi_i = -1 \right] \\ \le \;& \frac{d^d}{2}\mathbb{E}\left[f_{\mathcal{A}}\left(\frac{1}{d}\sum_{j=1}^d \xi_j x^j\right) - \underline{v} \,\middle|\, \prod_{i=1}^d \xi_i = 1 \right], \end{aligned}$$

where the last inequality follows from the fact that $f_{\mathcal{A}}\left(\frac{1}{d}\sum_{j=1}^d \xi_j x^j\right) - \underline{v}$ is always non–negative. In particular, we can find in constant time a vector $\beta = (\beta_1, \beta_2, \ldots, \beta_d) \in \{-1,1\}^d$ that satisfies $\prod_{i=1}^d \beta_i = 1$ and

$$f_{\mathcal{A}}\left(\frac{1}{d}\sum_{j=1}^d \beta_j x^j\right) - \underline{v} \ge \frac{2d!}{d^d} \cdot F_{\mathcal{A}}(x^1, x^2, \ldots, x^d).$$

Upon setting $\hat{x} = \frac{1}{d}\sum_{j=1}^d \beta_j x^j$ and observing that $v^* \ge \bar{v} \ge \underline{v} \ge -v^*$, we obtain

$$f_{\mathcal{A}}(\hat{x}) - \underline{v} \ge 2\alpha \cdot d! \cdot d^{-d} \cdot v^* \ge \alpha \cdot d! \cdot d^{-d} \cdot (\bar{v} - \underline{v}).$$

Moreover, we have $\|\hat{x}\|_p \le 1$. This completes the proof of Theorem 2. □

Appendix B: Proof of Theorem 3. Let $d \ge 3$ and $p \in [2, \infty]$ be fixed. We shall reduce Problem (ML) to Problem (MR), again by using the symmetrization procedure introduced in Section 2. Towards that end, let us first establish some preparatory results.

PROPOSITION 11. *Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ be an arbitrary order–d tensor and sym$(\mathcal{A}) \in \mathbb{R}^{N^d}$ be its symmetrization, where $N = n_1 + n_2 + \cdots + n_d$. Moreover, let $z^i = [\,(z^{i,1})^T\,(z^{i,2})^T \cdots (z^{i,d})^T\,]^T \in \mathbb{R}^N$ be given, where $z^{i,j} \in \mathbb{R}^{n_j}$ for $i,j = 1,\ldots,d$. Then,*

$$F_{sym(\mathcal{A})}(z^1, z^2, \ldots, z^d) = \sum_{(\pi_1,\pi_2,\ldots,\pi_d) \in S_d} F_{\mathcal{A}}(z^{\pi_1,1}, z^{\pi_2,2}, \ldots, z^{\pi_d,d}),$$

where S_d is the set of permutations of $\{1,2,\ldots,d\}$.

Proof. Using the sets $B_1, \ldots, B_d$ defined in (3) and the definition of sym$(\mathcal{A})$, we have

$$\begin{aligned} F_{\mathrm{sym}(\mathcal{A})}(z^1,\ldots,z^d) &= \sum_{\pi=(\pi_1,\ldots,\pi_d)\in S_d} \sum_{\substack{i_j \in B_{\pi_j} \\ j=1,\ldots,d}} [\mathrm{sym}(\mathcal{A})]_{i_1\cdots i_d}\, z^1_{i_1} \cdots z^d_{i_d} \\ &= \sum_{\pi=(\pi_1,\ldots,\pi_d)\in S_d} \sum_{i_1=1}^{n_{\pi_1}} \cdots \sum_{i_d=1}^{n_{\pi_d}} [\mathcal{A}^\pi]_{i_1\cdots i_d}\, z^{1,\pi_1}_{i_1} \cdots z^{d,\pi_d}_{i_d} \\ &= \sum_{\pi=(\pi_1,\ldots,\pi_d)\in S_d} \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} [\mathcal{A}]_{i_1\cdots i_d}\, z^{\pi_1^{-1},1}_{i_1} \cdots z^{\pi_d^{-1},d}_{i_d}, \end{aligned}$$

where $\pi^{-1} = (\pi_1^{-1}, \ldots, \pi_d^{-1}) \in S_d$ is the inverse of π; i.e., $\pi_{\pi_j^{-1}} = j$ for $j = 1, \ldots, d$. Consequently, we obtain

$$F_{\mathrm{sym}(\mathcal{A})}(z^1,\ldots,z^d) = \sum_{(\pi_1,\ldots,\pi_d)\in S_d} F_{\mathcal{A}}(z^{\pi_1^{-1},1}, \ldots, z^{\pi_d^{-1},d}) = \sum_{(\pi_1,\ldots,\pi_d)\in S_d} F_{\mathcal{A}}(z^{\pi_1,1}, \ldots, z^{\pi_d,d}),$$

as desired. □

PROPOSITION 12. *Let $p \in [2,\infty)$ be fixed. Given an integer $n \in [2,d]$, define the function $f_n : [0,d]^n \to \mathbb{R}$ by*

$$f_n(x_1,\ldots,x_n) = \sum_{i=1}^n x_i^{1/p} \prod_{j\neq i} (d-x_j)^{1/p}.$$

Then, for any $(x_1,\ldots,x_n) \in [0,d]^n$, we have

$$f_n(x_1,\ldots,x_n) \le f_n\left(\frac{d}{n},\ldots,\frac{d}{n}\right) = d^{n/p}\cdot n^{1-1/p}\cdot\left(1-\frac{1}{n}\right)^{(n-1)/p}. \tag{30}$$

Proof. We prove (30) by induction on n. For the base case (i.e., $n=2$), consider the problem

$$(P_2) \qquad \max\{f_2(x_1,x_2) : 0 \le x_i \le d \text{ for } i=1,2\}.$$

Note that an optimal solution to (P_2) must either lie on the boundary of $[0,d]^2$, or lie in the interior of $[0,d]^2$ and be a solution to the following first–order necessary conditions:

$$x_1^{(1/p)-1}(d-x_2)^{1/p} = x_2^{1/p}(d-x_1)^{(1/p)-1}, \tag{31}$$

$$x_2^{(1/p)-1}(d-x_1)^{1/p} = x_1^{1/p}(d-x_2)^{(1/p)-1}. \tag{32}$$

Consider an arbitrary $\bar{x} = (\bar{x}_1,\bar{x}_2) \in [0,d]^2$. If $\bar{x}$ is a boundary point of $[0,d]^2$, then the structure of f_2 implies that

$$f_2(\bar{x}_1,\bar{x}_2) \le \max\{f_2(0,d), f_2(d,0)\} = d^{2/p}.$$

On the other hand, suppose that $\bar{x} \in (0,d)^2$ satisfies (31) and (32). Then,

$$\left(\frac{\bar{x}_1}{d-\bar{x}_1}\right)^{(1/p)-1} = \left(\frac{\bar{x}_2}{d-\bar{x}_2}\right)^{1/p} \quad \text{and} \quad \left(\frac{\bar{x}_2}{d-\bar{x}_2}\right)^{(1/p)-1} = \left(\frac{\bar{x}_1}{d-\bar{x}_1}\right)^{1/p}, \tag{33}$$

which together yield $\bar{x}_1\bar{x}_2 = (d-\bar{x}_1)(d-\bar{x}_2)$, or equivalently, $\bar{x}_1+\bar{x}_2 = d$. If $p=2$, then any $(x_1,x_2) \in [0,d]^2$ satisfying $x_1+x_2=d$ will be an optimal solution to (P_2). In particular, we have $f_2(x_1,x_2) \le f_2(d/2,d/2) = d$ for all $(x_1,x_2) \in [0,d]^2$ in this case. If $p>2$, then upon substituting $\bar{x}_1+\bar{x}_2=d$ into (33), we obtain a unique solution $\bar{x}_1=\bar{x}_2=d/2$. Since $f_2(\bar{x}_1,\bar{x}_2) = 2(d/2)^{2/p} > d^{2/p}$ for any $p>2$, we conclude that $(d/2,d/2)$ is the optimal solution to (P_2). This establishes the base case.

For the inductive step, consider the problem

$$(P_n) \qquad \max\{f_n(x_1,\ldots,x_n) : 0 \le x_i \le d \text{ for } i=1,\ldots,n\},$$

where $2<n\le d$. Again, an optimal solution to (P_n) must either lie on the boundary of $[0,d]^n$, or lie in the interior of $[0,d]^n$ and be a solution to the following first–order necessary conditions:

$$\left(\frac{x_i}{d-x_i}\right)^{(1/p)-1} = \sum_{j\neq i}\left(\frac{x_j}{d-x_j}\right)^{1/p} \quad \text{for } i=1,\ldots,n. \tag{34}$$

Consider an arbitrary $\bar{x} = (\bar{x}_1,\ldots,\bar{x}_n) \in [0,d]^n$. Suppose first that $\bar{x} \in (0,d)^n$ satisfies (34). Let $u_i = \bar{x}_i/(d-\bar{x}_i)$ for $i=1,\ldots,n$. Then, we obtain from (34) that $(u_i)^{(1/p)-1} - (u_j)^{(1/p)-1} = u_j^{1/p} - u_i^{1/p}$, or equivalently,

$$(u_i)^{(1/p)-1}(1+u_i) = (u_j)^{(1/p)-1}(1+u_j) \quad \text{for } 1 \le i < j \le n. \tag{35}$$

It is easy to verify that the function $t \mapsto t^{(1/p)-1}(1+t)$ is strictly decreasing on $(0,p-1]$ and strictly increasing on $[p-1,\infty)$. Thus, if we let $I_1 = \{i : u_i < p-1\}$ and $I_2 = \{i : u_i \ge p-1\}$, then (35) implies that $u_i = u_j = u$ for all $i,j \in I_1$ and $u_i = u_j = v$ for all $i,j \in I_2$; i.e., $\bar{x}_i = \bar{x}_j$ whenever $i,j \in I_1$

or $i,j \in I_2$. We claim that in fact $I_2 = \emptyset$. To prove this, let us first show that $|I_2| \le 1$. Suppose to the contrary that $|I_2| \ge 2$. Let $i,j \in I_2$ be such that $i \ne j$. Then, from (34) and the fact that $u_i > 0$ for $i = 1,\ldots,n$, we have

$$u_i^{(1/p)-1} - u_j^{1/p} = \sum_{k \ne i,j} u_k^{1/p} > 0.$$

However, since $u_i, u_j \ge p-1 \ge 1$, we have $u_i^{(1/p)-1} - u_j^{1/p} \le 0$, which is a contradiction. It follows that $|I_2| \le 1$.

Now, suppose that $|I_2| = 1$. Then, from (34), we have

$$v^{(1/p)-1} = (n-1)u^{1/p}. \tag{36}$$

This, together with (35), implies that

$$u^{(1/p)-1}(1+u) = (n-1)u^{1/p}\left[1 + (n-1)^{p/(1-p)}u^{1/(1-p)}\right],$$

or equivalently,

$$(n-2)u + (n-1)^{1/(1-p)}u^{(p-2)/(p-1)} = 1. \tag{37}$$

Since both summands in (37) are non–negative, we clearly have $u \le 1/(n-2) < 1$. We claim that

$$u \ge \beta \equiv \frac{1-(n-1)^{1/(1-p)}}{n-2}. \tag{38}$$

Indeed, suppose to the contrary that $u < \beta$. Then, using the fact that $u < 1$, we have

$$\begin{aligned}(n-2)u + (n-1)^{1/(1-p)}u^{(p-2)/(p-1)} &< 1 - (n-1)^{1/(1-p)} + (n-1)^{1/(1-p)}u^{(p-2)/(p-1)} \\ &= 1 + (n-1)^{1/(1-p)}\left[u^{(p-2)/(p-1)} - 1\right] \\ &\le 1,\end{aligned}$$

which contradicts (37). This establishes (38).

We now show that (38) leads to $v < 1$, which would contradict the definition of v. Indeed, using (38), we have

$$(n-1)^p u \ge \frac{(n-1)^p}{n-2}\left[1-(n-1)^{1/(1-p)}\right]. \tag{39}$$

Consider the function $h : [2,\infty) \to \mathbb{R}$ given by

$$h(p) = (n-1)^p\left[1-(n-1)^{1/(1-p)}\right].$$

By a routine computation, we have

$$h'(p) = \ln(n-1)\cdot(n-1)^p\cdot\left[1-(n-1)^{1/(1-p)} - \frac{1}{(1-p)^2}(n-1)^{1/(1-p)}\right].$$

Observe that

$$\begin{aligned}& 1-(n-1)^{1/(1-p)} - \frac{1}{(1-p)^2}(n-1)^{1/(1-p)} \ge 0 \\ \Longleftrightarrow\ & \left(\frac{1}{n-1}\right)^{1/(p-1)}\left[1+\frac{1}{(p-1)^2}\right] \le 1 \\ \Longleftarrow\ & 1+\frac{1}{(p-1)^2} \le 2^{1/(p-1)}.\end{aligned} \tag{40}$$

It is straightforward to show that (40) holds for $p \in [2,3]$ by comparing the slopes of the functions $p \mapsto 1+(p-1)^{-2}$ and $p \mapsto 2^{1/(p-1)}$. For $p \geq 3$, we have

$$\left[1+\frac{1}{(p-1)^2}\right]^{(p-1)^2} \leq e < 2^2 \leq 2^{p-1},$$

which implies that (40) holds. Thus, we see that h is increasing on $p \in [2,\infty)$, and from (39) we obtain

$$(n-1)^p u \geq \frac{(n-1)^p}{n-2}\left[1-(n-1)^{1/(1-p)}\right] \geq \frac{(n-1)^2}{n-2}\left(1-\frac{1}{n-1}\right) = n-1 > 1.$$

This, together with (36), implies that

$$v = \left((n-1)^p u\right)^{1/(1-p)} < 1,$$

which is the desired contradiction.

Thus, we have shown that $|I_2| = 0$. Using (34), we then have $u^{(1/p)-1} = (n-1)u^{1/p}$, or equivalently, $u = 1/(n-1)$. It follows that $\bar{x} = (d/n,\ldots,d/n)$ is the unique solution to (34).

Next, we show that if $\bar{x}$ lies on the boundary of $[0,d]^n$, then $f_n(\bar{x}_1,\ldots,\bar{x}_n) \leq f_n(d/n,\ldots,d/n)$. Towards that end, we first note that

$$\begin{aligned} f_n\left(\frac{d}{n},\ldots,\frac{d}{n}\right) &= d^{n/p}\cdot n^{1-1/p}\cdot\left(1-\frac{1}{n}\right)^{(n-1)/p} \\ &= d^{n/p}\cdot n\cdot\left(\frac{1}{n-1}\right)^{1/p}\cdot\left(1-\frac{1}{n}\right)^{n/p} \\ &\geq d^{n/p}\cdot n\cdot\left(\frac{1}{n-1}\right)^{1/p}\cdot e^{-1/p}\cdot\left(1-\frac{1}{n}\right)^{1/p} \\ &\geq d^{n/p}\cdot\sqrt{n/e}, \end{aligned} \tag{41}$$

where the last inequality follows from the fact that $p \geq 2$. Now, suppose that $\bar{x}_i = d$ for some $i = 1,\ldots,n$. Since the function f_n is symmetric in its arguments, we may assume without loss that $i = n$. Then, using (41) and the fact that $n \geq 3$, we have

$$f_n(\bar{x}_1,\ldots,\bar{x}_{n-1},d) \leq d^{n/p} < f_n\left(\frac{d}{n},\ldots,\frac{d}{n}\right).$$

On the other hand, suppose that $\bar{x}_n = 0$. Then, by the inductive hypothesis,

$$\begin{aligned} f_n(\bar{x}_1,\ldots,\bar{x}_{n-1},0) &= d^{1/p}\cdot f_{n-1}(\bar{x}_1,\ldots,\bar{x}_{n-1}) \\ &\leq d^{1/p}\cdot d^{(n-1)/p}\cdot(n-1)^{1-1/p}\cdot\left(1-\frac{1}{n-1}\right)^{(n-2)/p} \\ &= d^{n/p}(n-1)\left(\frac{1}{n-2}\right)^{1/p}\left(1-\frac{1}{n-1}\right)^{(n-1)/p}. \end{aligned} \tag{42}$$

Since $n \geq 3$ and $p \geq 2$, we have

$$(n-1)\left(\frac{1}{n-2}\right)^{1/p}\left(1-\frac{1}{n-1}\right)^{(n-1)/p} < n\left(\frac{1}{n-1}\right)^{1/p}\left(1-\frac{1}{n}\right)^{n/p}$$

$$\Longleftrightarrow \frac{n}{n-1}\left(1-\frac{1}{n-1}\right)^{2/p}\left[1+\frac{1}{n(n-2)}\right]^{n/p} > 1$$
$$\Longleftarrow \frac{n}{n-1}\left(1-\frac{1}{n-1}\right)^{2/p}\left[1+\frac{1}{n(n-2)}\right]^{2/p} \geq 1$$
$$\Longleftrightarrow \left(\frac{n}{n-1}\right)^{1-2/p} \geq 1.$$

Hence, we obtain from (42) that

$$f_n(\bar{x}_1,\ldots,\bar{x}_{n-1},0) < d^{n/p}\cdot n\cdot\left(\frac{1}{n-1}\right)^{1/p}\cdot\left(1-\frac{1}{n}\right)^{n/p} = f_n\left(\frac{d}{n},\ldots,\frac{d}{n}\right).$$

This completes the proof of Proposition 12. □

PROPOSITION 13. *Let $\mathcal{A}\in\mathbb{R}^{n_1\times n_2\times\cdots\times n_d}$ be an arbitrary order–d tensor and $sym(\mathcal{A})\in\mathbb{R}^{N^d}$ be its symmetrization, where $N=n_1+n_2+\cdots+n_d$. Consider the optimization problems*

$$\begin{aligned}\tau(A_d) = \ &\textit{maximize} \quad d!\cdot F_{\mathcal{A}}(x^1,x^2,\ldots,x^d)\\ &\textit{subject to} \quad \|x^i\|_p\leq 1,\ x^i\in\mathbb{R}^{n_i} \quad \textit{for } i=1,\ldots,d\end{aligned} \qquad (A_d)$$

and

$$\begin{aligned}\tau(B_d) = \ &\textit{maximize} \quad F_{sym(\mathcal{A})}(z^1,z^2,\ldots,z^d)\\ &\textit{subject to} \quad \|z^i\|_p\leq d^{1/p},\ z^i\in\mathbb{R}^{N} \quad \textit{for } i=1,\ldots,d,\end{aligned} \qquad (B_d)$$

where $F_{\mathcal{A}}$ (resp. $F_{sym}(\mathcal{A})$) is the multilinear form associated with $\mathcal{A}$ (resp. $sym(\mathcal{A})$). Then, the following hold:

(a) $\tau(A_d)=\tau(B_d)$.

(b) Let $(\bar{x}^1,\bar{x}^2,\ldots,\bar{x}^d)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\cdots\times\mathbb{R}^{n_d}$ be an optimal solution to Problem (A_d). Set

$$\hat{z}^i = \left[\,(\bar{x}^1)^T\,(\bar{x}^2)^T\,\cdots\,(\bar{x}^d)^T\,\right]^T\in\mathbb{R}^N \quad \textit{for } i=1,\ldots,d. \qquad (43)$$

Then, $(\hat{z}^1,\hat{z}^2,\ldots,\hat{z}^d)$ constitutes an optimal solution to Problem (B_d).

(c) Let $(\tilde{z}^1,\tilde{z}^2,\ldots,\tilde{z}^d)\in\mathbb{R}^N\times\mathbb{R}^N\times\cdots\times\mathbb{R}^N$ be an optimal solution to Problem (B_d) with

$$\tilde{z}^i = \left[\,(\tilde{z}^{i,1})^T\,(\tilde{z}^{i,2})^T\,\cdots\,(\tilde{z}^{i,d})^T\,\right]^T\in\mathbb{R}^N, \quad \tilde{z}^{i,j}\in\mathbb{R}^{n_j} \quad \textit{for } i,j=1,\ldots,d. \qquad (44)$$

Then, $\|\tilde{z}^{i,j}\|_p=1$ for $i,j=1,\ldots,d$. Moreover, there exists a vector $\bar{z}\in\mathbb{R}^N$ with

$$\bar{z} = \left[\,(\hat{x}^1)^T\,(\hat{x}^2)^T\,\cdots\,(\hat{x}^d)^T\,\right]^T, \quad \hat{x}^i\in\mathbb{R}^{n_i} \quad \textit{for } i=1,\ldots,d,$$

such that $(\bar{z},\bar{z},\ldots,\bar{z})$ is an optimal solution to Problem (B_d) and $(\hat{x}^1,\hat{x}^2,\ldots,\hat{x}^d)$ is an optimal solution to Problem (A_d).

Proof. By Proposition 11, Problem (B_d) is equivalent to

$$\begin{aligned}\tau(B_d) = \ &\text{maximize} \quad \sum_{\pi\in S_d} F_{\mathcal{A}}(z^{\pi_1,1},\ldots,z^{\pi_d,d})\\ &\text{subject to} \quad \|z^i\|_p\leq d^{1/p},\ z^i\in\mathbb{R}^N \quad \text{for } i=1,\ldots,d.\end{aligned} \qquad (B_d')$$

If $(\bar{x}^1,\ldots,\bar{x}^d)\in\mathbb{R}^{n_1}\times\cdots\times\mathbb{R}^{n_d}$ is an optimal solution to Problem (A_d), then the solution $(\hat{z}^1,\ldots,\hat{z}^d)\in\mathbb{R}^N\times\cdots\times\mathbb{R}^N$ as defined in (43) is feasible for Problem (B_d). Moreover, we have

$$\tau(B_d)\geq\sum_{\pi\in S_d}F_{\mathcal{A}}(\hat{z}^{\pi_1,1},\ldots,\hat{z}^{\pi_d,d}) = d!\cdot F_{\mathcal{A}}(\bar{x}^1,\ldots,\bar{x}^d) = \tau(A_d).$$

Hence, (b) is implied by (a).

To prove (a) and (c), we consider two cases:

Case 1: $p=\infty$**.** Consider an optimal solution $(\tilde{z}^1,\ldots,\tilde{z}^d)\in\mathbb{R}^N\times\cdots\times\mathbb{R}^N$ to Problem (B_d) with $\tilde{z}^i$ given by (44). Then, $(\tilde{z}^1,\ldots,\tilde{z}^d)$ is also optimal for Problem (B_d'). Let $\tau\in S_d$ be a permutation of $\{1,\ldots,d\}$ satisfying

$$F_{\mathcal{A}}(\tilde{z}^{\tau_1,1},\ldots,\tilde{z}^{\tau_d,d})\ge F_{\mathcal{A}}(\tilde{z}^{\pi_1,1},\ldots,\tilde{z}^{\pi_d,d})\quad\text{for all }\pi\in S_d. \tag{45}$$

By the optimality of $(\tilde{z}^1,\ldots,\tilde{z}^d)$ for Problem (B_d') and the fact that $\|\tilde{z}^i\|_\infty=\max_{1\le j\le d}\|\tilde{z}^{i,j}\|_\infty$, we have $\|\tilde{z}^{i,j}\|_\infty=1$ for $i,j=1,\ldots,d$. Now, set $\hat{x}^i=\tilde{z}^{\tau_i,i}$ for $i=1,\ldots,d$ and form $\bar{z}=\left[(\hat{x}^1)^T\cdots(\hat{x}^d)^T\right]^T\in\mathbb{R}^N$. By construction, we have $\|\bar{z}\|_\infty=1$ and hence $(\bar{z},\ldots,\bar{z})$ is feasible for Problem (B_d'). Using (45), we compute

$$\sum_{\pi\in S_d}F_{\mathcal{A}}(\hat{x}^1,\ldots,\hat{x}^d)=\sum_{\pi\in S_d}F_{\mathcal{A}}(\tilde{z}^{\tau_1,1},\ldots,\tilde{z}^{\tau_d,d})\ge\sum_{\pi\in S_d}F_{\mathcal{A}}(\tilde{z}^{\pi_1,1},\ldots,\tilde{z}^{\pi_d,d})=\tau(B_d), \tag{46}$$

which certifies the optimality of $(\bar{z},\ldots,\bar{z})$ for Problem (B_d') and hence also for Problem (B_d). Moreover, since $\|\hat{x}^i\|_\infty\le\|\bar{z}\|_\infty=1$ for $i=1,\ldots,d$, $(\hat{x}^1,\ldots,\hat{x}^d)$ is feasible for Problem (A_d). This implies that

$$\sum_{\pi\in S_d}F_{\mathcal{A}}(\hat{x}^1,\ldots,\hat{x}^d)=d!\cdot F_{\mathcal{A}}(\hat{x}^1,\ldots,\hat{x}^d)\le\tau(A_d). \tag{47}$$

Upon combining (46) and (47), we have $\tau(A_d)\ge\tau(B_d)$, and that $(\hat{x}^1,\ldots,\hat{x}^d)$ is an optimal solution to Problem (A_d). This establishes (a) and (c) for this case.

Case 2: $p\in[2,\infty)$**.** We prove (a) and (c) by induction on $d\ge2$. For the base case (i.e., $d=2$), we have $\mathcal{A}\in\mathbb{R}^{n_1\times n_2}$. Hence, we can write

$$\tau(A_2)=2\max\left\{x^T\mathcal{A}y:\|x\|_p\le1,\ \|y\|_p\le1\right\} \tag{A_2}$$

and

$$\tau(B_2)=\max\left\{(v^1)^T\mathcal{A}w^2+(w^1)^T\mathcal{A}v^2:\|v^1\|_p^p+\|v^2\|_p^p\le2,\ \|w^1\|_p^p+\|w^2\|_p^p\le2\right\}. \tag{B_2}$$

Let $(\tilde{z}^1,\tilde{z}^2)\in\mathbb{R}^{n_1+n_2}\times\mathbb{R}^{n_1+n_2}$, where $\tilde{z}^1=\left[(\tilde{v}^1)^T\,(\tilde{v}^2)^T\right]$ and $\tilde{z}^2=\left[(\tilde{w}^1)^T\,(\tilde{w}^2)^T\right]$, be an optimal solution to Problem (B_2). Suppose that $\|\tilde{v}^1\|_p^p=k_1$ and $\|\tilde{w}^1\|_p^p=k_2$. Then,

$$\begin{aligned}\tau(B_2)&\le\max\left\{(v^1)^T\mathcal{A}w^2:\|v^1\|_p^p\le k_1,\ \|w^2\|_p^p\le2-k_2\right\}\\&\quad+\max\left\{(w^1)^T\mathcal{A}v^2:\|w^1\|_p^p\le k_2,\ \|v^2\|_p^p\le2-k_1\right\}\\&=\frac{\tau(A_2)}{2}\left[k_1^{1/p}(2-k_2)^{1/p}+k_2^{1/p}(2-k_1)^{1/p}\right]\\&\le\tau(A_2),\end{aligned}$$

where the last inequality follows from Proposition 12 and the fact that $0\le k_1,k_2\le2$. This establishes (a). Moreover, since all the above inequalities hold as equalities, from the proof of Proposition 12, both k_1,k_2 must equal to 1 when $p>2$ and can be taken as 1 when $p=2$. This, together with the optimality of $(\tilde{z}^1,\tilde{z}^2)$, implies that we can take $\tilde{v}^1=\tilde{w}^1$ and $\tilde{v}^2=\tilde{w}^2$. Upon setting $\hat{x}^i=\tilde{v}^i$ for $i=1,2$ and forming $\bar{z}=\left[(\hat{x}^1)^T\,(\hat{x}^2)^T\right]$, it can be verified that (c) holds. Thus, the base case is established.

Next, consider an optimal solution $(\tilde{z}^1,\ldots,\tilde{z}^d)$ to Problem (B_d) with $\tilde{z}^i$ given by (44). Suppose that $\|\tilde{z}^{i,d}\|_p^p = k_i$ for $i=1,\ldots,d$. Then,

$$\begin{aligned}
\tau(B_d) &= \sum_{\pi\in S_d} F_{\mathcal{A}}(\tilde{z}^{\pi_1,1},\ldots,\tilde{z}^{\pi_d,d}) \\
&= \sum_{i=1}^d \sum_{\pi\in S_d:\pi_d=i} F_{\mathcal{A}(\tilde{z}^{i,d})}(\tilde{z}^{\pi_1,1},\ldots,\tilde{z}^{\pi_{d-1},d-1}) \\
&\le \sum_{i=1}^d \max_{\substack{\|w^j\|_p^p\le d-k_j,\, w^j\in\mathbb{R}^{N-n_d} \\ j=1,\ldots,i-1,i+1,\ldots,d}} \sum_{\pi\in S_d:\pi_d=i} F_{\mathcal{A}(\tilde{z}^{i,d})}(w^{\pi_1,1},\ldots,w^{\pi_{d-1},d-1}) \\
&= \sum_{i=1}^d \prod_{j\ne i}\left(\frac{d-k_j}{d-1}\right)^{1/p} \max_{\substack{\|w^j\|_p^p\le d-1,\, w^j\in\mathbb{R}^{N-n_d} \\ j=1,\ldots,d-1}} \sum_{\pi\in S_{d-1}} F_{\mathcal{A}(\tilde{z}^{i,d})}(w^{\pi_1,1},\ldots,w^{\pi_{d-1},d-1}) \\
&= (d-1)!\sum_{i=1}^d \prod_{j\ne i}\left(\frac{d-k_j}{d-1}\right)^{1/p} \max_{\substack{\|x^j\|_p\le 1,\, x^j\in\mathbb{R}^{n_j} \\ j=1,\ldots,d-1}} F_{\mathcal{A}(\tilde{z}^{i,d})}(x^1,\ldots,x^{d-1}) \qquad (48) \\
&\le (d-1)!\sum_{i=1}^d \prod_{j\ne i}\left(\frac{d-k_j}{d-1}\right)^{1/p} \max_{\substack{\|x^j\|_p\le 1,\, x^j\in\mathbb{R}^{n_j} \\ j=1,\ldots,d-1}} \max_{\|x^d\|_p^p\le k_i,\, x^d\in\mathbb{R}^{n_d}} F_{\mathcal{A}}(x^1,\ldots,x^{d-1},x^d) \\
&= \frac{\tau(A_d)}{d(d-1)^{(n-1)/p}} \sum_{i=1}^d k_i^{1/p} \prod_{j\ne i}(d-k_j)^{1/p} \\
&\le \tau(A_d), \qquad (49)
\end{aligned}$$

where (48) follows from Proposition 11 and the inductive hypothesis, and (49) follows from Proposition 12 and the fact that $0\le k_i\le d$ for $i=1,\ldots,d$. This establishes (a). Moreover, since all the above inequalities hold as equalities, the proof of Proposition 12 shows that we must have $k_i=1$ for $i=1,\ldots,d$. This implies that $\|\tilde{z}^{i,d}\|_p = 1$ for $i=1,\ldots,d$. By repeating the above argument using the group $\{\tilde{z}^{i,j}: i=1,\ldots,d\}$ in place of $\{\tilde{z}^{i,d}: i=1,\ldots,d\}$ for each $j=1,\ldots,d-1$, we conclude that $\|\tilde{z}^{i,j}\|_p=1$ for $i,j=1,\ldots,d$. Now, as in Case 1, let $\tau\in S_d$ be a permutation of $\{1,\ldots,d\}$ satisfying (45). Set $\hat{x}^i = \tilde{z}^{\tau_i,i}$ for $i=1,\ldots,d$ and form $\bar{z} = \left[\,(\hat{x}^1)^T \cdots (\hat{x}^d)^T\,\right]^T \in\mathbb{R}^N$. By construction, we have

$$\|\bar{z}\|_p^p = \sum_{i=1}^d \|\hat{x}^i\|_p^p = \sum_{i=1}^d \|\tilde{z}^{\tau_i,i}\|_p^p = d,$$

and hence $(\bar{z},\ldots,\bar{z})$ is feasible for Problem (B_d). It remains to argue as in Case 1 to complete the inductive step and also the proof of Proposition 13. □

Proposition 13 implies that for any given $d\ge 3$ and $p\in[2,\infty]$, any instance of Problem (ML) can be converted into an instance of Problem (MR) in polynomial time. Since Problem (ML) is NP–hard by Proposition 1, it follows that Problem (MR) is also NP–hard. This completes the proof of Theorem 3. □

Acknowledgments. This research was supported by the Hong Kong Research Grants Council (RGC) General Research Fund (GRF) Project CUHK 419409, and was partially completed while the second author was visiting the Institute for Mathematical Sciences, National University of Singapore in 2012. We thank the two anonymous reviewers for their careful reading of the manuscript.

References

[1] Ahmadi, A. A., A. Olshevsky, P. A. Parrilo, J. N. Tsitsiklis. 2013. NP–Hardness of Deciding Convexity of Quartic Polynomials and Related Problems. *Math. Programming Ser. A* **137**(1–2) 453–476.

[2] Alon, N., A. Naor. 2006. Approximating the Cut–Norm via Grothendieck's Inequality. *SIAM J. Comput.* **35**(4) 787–803.

[3] Baratchart, L., M. Berthod, L. Pottier. 1998. Optimization of Positive Generalized Polynomials under ℓ^p Constraints. *J. Convex Anal.* **5**(2) 353–379.

[4] Barvinok, A. 2007. Integration and Optimization of Multivariate Polynomials by Restriction onto a Random Subspace. *Foundations Comput. Math.* **7**(2) 229–244.

[5] Ben-Tal, A., A. Nemirovski. 2001. On Approximating Matrix Norms. Manuscript.

[6] Bhaskara, A., A. Vijayaraghavan. 2011. Approximating Matrix p–Norms. *Proc. 22nd Annual ACM–SIAM Sympos. Discrete Algorithms (SODA 2011)*. 497–511.

[7] Braverman, M., K. Makarychev, Y. Makarychev, A. Naor. 2011. The Grothendieck Constant is Strictly Smaller than Krivine's Bound. *Proc. 52nd Annual IEEE Sympos. Foundations Comput. Sci. (FOCS 2011)*. 453–462.

[8] Brieden, A., P. Gritzmann, R. Kannan, V. Klee, L. Lovász, M. Simonovits. 2001. Deterministic and Randomized Polynomial–Time Approximation of Radii. *Mathematika* **48** 63–105.

[9] de Klerk, E., M. Laurent, P. A. Parrilo. 2006. A PTAS for the Minimization of Polynomials of Fixed Degree over the Simplex. *Theoretical Comput. Sci.* **361**(2–3) 210–225.

[10] Grötschel, M., L. Lovász, A. Schrijver. 1993. *Geometric Algorithms and Combinatorial Optimization*, *Algorithms and Combinatorics*, vol. 2. Second corrected ed. Springer–Verlag, Berlin Heidelberg.

[11] He, S., B. Jiang, Z. Li, S. Zhang. 2013. Probability Bounds for Polynomial Functions in Random Variables. Preprint.

[12] He, S., Z. Li, S. Zhang. 2010. Approximation Algorithms for Homogeneous Polynomial Optimization with Quadratic Constraints. *Math. Programming Ser. B* **125**(2) 353–383.

[13] Hillar, C. J., L.-H. Lim. 2013. Most Tensor Problems are NP–Hard. *J. ACM* **60**(6) Article 45.

[14] Khot, S., A. Naor. 2008. Linear Equations Modulo 2 and the L_1 Diameter of Convex Bodies. *SIAM J. Comput.* **38**(4) 1448–1463.

[15] Khot, S., A. Naor. 2012. Grothendieck–Type Inequalties in Combinatorial Optimization. *Comm. Pure Appl. Math.* **65**(7) 992–1035.

[16] Kindler, G., A. Naor, G. Schechtman. 2010. The UGC Hardness Threshold of the L_p Grothendieck Problem. *Math. Oper. Res.* **35**(2) 267–283.

[17] Li, Z., S. He, S. Zhang. 2012. *Approximation Methods for Polynomial Optimization*. SpringerBriefs in Optimization, Springer Science+Business Media, LLC, New York.

[18] Lim, L.-H. 2005. Singular Values and Eigenvalues of Tensors: A Variational Approach. *Proc. 1st IEEE Internat. Workshop Comput. Adv. Multi–Sensor Adaptive Processing (CAMSAP 2005)*. 129–132.

[19] Ling, C., J. Nie, L. Qi, Y. Ye. 2009. Biquadratic Optimization over Unit Spheres and Semidefinite Programming Relaxations. *SIAM J. Optim.* **20**(3) 1286–1310.

[20] Ling, C., X. Zhang, L. Qi. 2012. Semidefinite Relaxation Approximation for Multivariate Bi–Quadratic Optimization with Quadratic Constraints. *Numer. Linear Algebra Appl.* **19**(1) 113–131.

[21] Luo, Z.-Q., S. Zhang. 2010. A Semidefinite Relaxation Scheme for Multivariate Quartic Polynomial Optimization with Quadratic Constraints. *SIAM J. Optim.* **20**(4) 1716–1736.

[22] Nesterov, Yu. 2003. Random Walk in a Simplex and Quadratic Optimization over Convex Polytopes. CORE Discussion Paper 2003071, Université Catholique de Louvain, Belgium.

[23] Nesterov, Yu. 2000. Global Quadratic Optimization via Conic Relaxation. H. Wolkowicz, R. Saigal, L. Vandenberghe, eds., *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, *International Series in Operations Research and Management Science*, vol. 27. Kluwer Academic Publishers, Boston, Massachusetts, 363–387.

[24] Pappas, A., Y. Sarantopoulos, A. Tonge. 2007. Norm Attaining Polynomials. *Bull. London Math. Soc.* **39**(2) 255–264.

[25] Pisier, G. 2012. Grothendieck's Theorem, Past and Present. *Bull. Amer. Math. Soc. (N.S.)* **49**(2) 237–323.

[26] Qi, L. 2005. Eigenvalues of a Real Supersymmetric Tensor. *J. Symbolic Comput.* **40**(6) 1302–1324.

[27] Ragnarsson, S., C. F. Van Loan. 2013. Block Tensors and Symmetric Embeddings. *Linear Algebra Appl.* **438**(2) 853–874.

[28] So, A. M.-C. 2011. Deterministic Approximation Algorithms for Sphere Constrained Homogeneous Polynomial Optimization Problems. *Math. Programming Ser. B* **129**(2) 357–382.

[29] So, A. M.-C., J. Zhang, Y. Ye. 2007. On Approximating Complex Quadratic Optimization Problems via Semidefinite Programming Relaxations. *Math. Programming Ser. B* **110**(1) 93–110.

[30] Steinberg, D. 2005. Computation of Matrix Norms with Applications to Robust Optimization. Master's thesis, Technion—Israel Institute of Technology, Technion City, Haifa 32000, Israel.

[31] Yang, Y., Q. Yang. 2012. On Solving Biquadratic Optimization via Semidefinite Relaxation. *Comput. Optim. Appl.* **53**(3) 845–867.

[32] Zhang, X., C. Ling, L. Qi. 2011. Semidefinite Relaxation Bounds for Bi–Quadratic Optimization Problems with Quadratic Constraints. *J. Global Optim.* **49**(2) 293–311.

[33] Zhang, X., L. Qi, Y. Ye. 2012. The Cubic Spherical Optimization Problems. *Math. Comput.* **81**(279) 1513–1525.