Commun. Optim. Theory 2023 (2023) 31 https://doi.org/10.23952/cot.2023.31

Communications in Optimization Theory
Available online at http://cot.mathres.org

SOME PROPERTIES OF BI-FORM OPTIMIZATION OVER GENERALIZED SPHERES

SHENGLONG HU

Department of Mathematics, School of Science, Hangzhou Dianzi University, Hangzhou 310018, China

Dedicated to the memory of Professor Naum Z. Shor

Abstract. In this paper, we study the bi-form optimization problem over the generalized spheres. Since bi-form optimization problem includes bi-quadratic optimization as a special case, it is NP-Hard in general. Nevertheless, we characterize the primal and the dual problems associated to the bi-form optimization problem and show that there is no duality gap between them. Based on tensor analysis and the dual characterization, we propose some methods to solve the bi-form optimization problem approximately. Particularly, we analyze the sums of powers of tensor polynomials and give a Positivstellensatz for the bi-form optimization over the spheres. We also show that there is a class of bi-form optimization problems which can be solved in polynomial-time. For a bi-form problem with nonpositive coefficients, we present a globally convergent algorithm which can compute an approximation solution with an explicit approximation ratio in terms of the degree of the bi-form and the orders of the norms.
Keywords. Bi-form optimization; Sums of squares; Sums of powers; Spheres.

2020 Mathematics Subject Classification. 65F15, 65K05, 90C90.

1. Introduction

Polynomial optimization problems have a long and rich history, see [9, 18, 20, 26, 29] and references therein. In this paper, we study the following bi-form optimization problem over the generalized spheres:

$$\begin{array}{ll} \min & Bx^s\cdot y^t := \sum_{i_1,\dots,i_s=1}^m \sum_{j_1,\dots,j_t=1}^n B_{i_1\dots i_s j_1\dots j_t} x_{i_1}\cdots x_{i_s} y_{j_1}\cdots y_{j_t} \\ \text{s.t.} & \|x\|_k = 1, \quad \|y\|_l = 1, \end{array} \tag{1.1}$$

where B is an (s,t)-order $m\times n$ dimensional tensor with entries $B_{i_1\dots i_s j_1\dots j_t}\in\mathbb{R}$ for all $i_1,\dots,i_s\in\{1,2,\dots,m\}$ and $j_1,\dots,j_t\in\{1,2,\dots,n\}$ (notions on tensors will be given in Section 2), $x\in\mathbb{R}^m$, $y\in\mathbb{R}^n$, $\|\cdot\|_p$ represents l_p norm for vectors of appropriate size for $p\geq 1$, $s\geq 2$ and $t\geq 2$ are given even integers, and $k,l\geq 1$ are given positive numbers.

E-mail address: shenglonghu@hdu.edu.cn.
Received: November 20, 2022; Accepted: March 7, 2023.

It is a *variational optimization problem* for rectangular tensors. Actually, in [4, 5], singular value problems for rectangular tensors were proposed and analyzed:

$$\left\{\begin{array}{rcl} Tx^{s-1}y^t & = & \lambda x^{[s+t-1]}, \\ Tx^s y^{t-1} & = & \lambda y^{[s+t-1]}. \end{array}\right. \tag{1.2}$$

Here T is an (s,t)-order $m\times n$ dimensional rectangular tensor in $\mathbb{R}$. If $(\lambda,x,y)\in\mathbb{R}\times\mathbb{R}^m\setminus\{0\}\times\mathbb{R}^n\setminus\{0\}$ satisfy (1.2), then it is named an H-singular triple of the tensor T (the name came from the literature [21, 22, 23], we call simply singular triple in this paper). Denote by $\sigma(T)$ the set of singular values of T and $\rho(T)$ the largest absolute value of element(s) in $\sigma(T)$. We call $\sigma(T)$ and $\rho(T)$ the spectra and the spectral radius of the tensor T, respectively. When the tensor T is partially symmetric, it is easy to see that $\rho(T)$ and its corresponding singular vectors are the global optimal value and its corresponding optimal solution of the following polynomial maximization problem, respectively

$$\begin{array}{ll} \max & |Tx^s\cdot y^t| := \left|\sum_{i_1,\dots,i_s=1}^m \sum_{j_1,\dots,j_t=1}^n T_{i_1\dots i_s j_1\dots j_t} x_{i_1}\cdots x_{i_s} y_{j_1}\cdots y_{j_t}\right| \\ \text{s.t.} & \|x\|_{s+t}=1, \quad \|y\|_{s+t}=1. \end{array} \tag{1.3}$$

In this paper, we will consider such variational optimization problems for rectangular tensors. We will assume that T is partially symmetric. Unlike the singular value problems for matrices, there are various definitions for singular values for rectangular tensors [19], hence various variational optimization problems for the spectral radii of tensors. There are also applications involved in multivariate polynomial optimization under unit spherical constraints using norms other than 2-norm or $s+t$-norm to make the system (1.2) homogeneous [27]. Actually, they form one starting point for investigations on low rank tensor approximation problems [11, 12, 13] from the perspective of projection.

Besides the singular value problem for tensors, problem (1.1) arises from various areas of applied science, such as, the strong ellipticity condition problem in solid mechanics, the entanglement problem in quantum physics, the best rank one approximation problem in signal processing and data analysis, and so on [20]. In particular, here are some special cases of the optimization problem (1.1).

- When $s,t\geq 1$, $k=l=s+t$ and B is partially symmetric (symmetric inside each of the two groups of indices), (1.1) reduces to singular value problem for the tensor B proposed in [4]. Specially, when $s=t=1$ and $k=l=2$, (1.1) reduces to the familiar singular value problem for matrices [14].
- When $s=t=k=l=2$, (1.1) reduces to the bi-quadratic optimization over unit spheres which was discussed deeply in [20].

In this paper, we will consider (1.1) with two cases: (**PI**) $k=s$ and $l=t$ for even s,t, and (**PII**) $k=l\geq s+t$ which includes bi-quadratic optimization and singular value problem for tensors as special cases, respectively. Since the bi-quadratic optimization problem is NP-Hard in general, problem (**PI**) considered in this paper is NP-Hard in general as well. Nevertheless, there exists no result on the NP-hardness of problem (**PII**) in the literature.

For the special bi-quadratic optimization problem, various approximation methods (both for theoretical analysis and practical computations) to solve a general bi-quadratic optimization problem or such an optimization problem with special structures were presented in [20].

In this paper, based on the concept of K-eigenvalues of tensors [30, 31], we will first establish some decomposition results for homogeneous polynomials of even degrees. Then, we will investigate problem (**PI**) from the prima-dual aspects in detail. Concretely, we first characterize the primal and the dual problems associated to problem (**PI**), and prove that there is no duality gap between them. Then, using the dual characterization of problem (**PI**) and the concepts of K-eigenvalues, we propose a numerical scheme for problem (**PI**). We also discuss the *sums of squares* of a tensor polynomial through its representation matrix polynomial, and the generalized *sums of powers* of a tensor polynomial. Especially, utilizing results from [25] on dual cone of the cone of positive semidefinite tensors, we give a theorem on sums of powers of tensor polynomials which is a generalization of the results in [20, 26]. At last, we point out that a class of problem (**PI**) can be solved in polynomial-time through the representation matrix of that tensor. To the best of our knowledge, such a result is new even when it reduces to the case of bi-quadratic optimization problems.

For problem (**PII**), we concentrate our attentions to nonnegative rectangular tensor T in (1.2). It turns out that the resulting maximization problem (1.3) can be solved by a globally linearly convergent algorithm with respect to Hilbert's projective metric. We point out that in [4], for the class of *irreducible* nonnegative tensors, an algorithm was proposed for solving maximization problem (1.3). Nevertheless, there is no convergent result. So, the contribution of this paper for problem (**PII**) are two folds:

- we give a convergent algorithm for the singular value problem for irreducible rectangular tensors proposed in [4], and also extend their results to general nonnegative rectangular tensors;
- for general cases when $k = l > s + t$, we give a convergent algorithm for the maximization problem (1.3) (hence its corresponding counterpart (1.1)) with an explicit approximation ratio.

The rest of the paper is organized as follows. The definitions and various related properties of K-eigenvalues and K-eigenvectors are given in Section 2. Section 3 investigates problem (**PI**) from the primal-dual aspects and several heuristic numerical schemes. In Section 4, we discuss problem (**PII**) in detail. Some final remarks are given in Section 5.

2. Preliminaries

In this section, we present a simple yet useful notion for tensors. It starts from the balanced matrix flattening of an even order tensor and the natural embedding of the space of symmetric tensors into the space of tensors. Then, we can treat a tensor as a linear mapping, and hence eigenvalue and singular value decompositions can be employed. Due to the embedding of symmetric tensors, a symmetric tensor (corresponding to a homogeneous polynomial) can be decomposed as well. More importantly, it acts well on the Veronese variety when it is viewed as a dual linear operator, and hence a meaningful decomposition of polynomials. We refer to [17] for basic notions on tensors.

These are related to K-eigenvalues of tensors discussed in the literature [30, 31].

2.1. Matrix flattening.

Given positive integers l and m, an l-th order m dimensional tensor T in a field $\mathbb{K}$ ($\mathbb{R}$ or $\mathbb{C}$) is a set of elements $T_{i_1 i_2 \dots i_l}$ for all $i_j \in \{1, 2, \dots, m\}$ and $j \in \{1, 2, \dots, l\}$, where every element $T_{i_1 i_2 \dots i_l}$ is in $\mathbb{K}$. It is denoted by $T := (T_{i_1 i_2 \dots i_l})$. For any two l-th order m

dimensional tensors $T^{(1)}$, $T^{(2)}$ and $\gamma \in \mathbb{K}$, we define

$$T^{(1)} + T^{(2)} := (T^{(1)}_{i_1 i_2 \ldots i_l} + T^{(2)}_{i_1 i_2 \ldots i_l}) \quad \text{and} \quad \gamma T := (\gamma T_{i_1 i_2 \ldots i_{2n}}).$$

Then, the set of all l-th order m dimensional tensors in $\mathbb{K}$ forms a linear space, denoted by $\mathscr{T}_{l,m}$, where the zero element under addition is denoted by $\mathbf{0}$, a tensor with all elements being zero. We will abbreviate $\mathscr{T}_{l,m}$ as $\mathscr{T}$ if there is no confusion.

For the convenience of the subsequent analysis, we introduce the following notation for tensors. For a tensor $T \in \mathscr{T}_{l,m}$ with $l = 2n$, we put its indices into two sets

$$\begin{aligned} \mathscr{I} &= \{i_1 i_2 \ldots i_n \mid i_j \in \{1,2,\ldots,m\}, j \in \{1,2,\ldots,n\}\}; \\ \mathscr{J} &= \{i_{n+1} i_{n+2} \ldots i_{2n} \mid i_j \in \{1,2,\ldots,m\}, j \in \{n+1,n+2,\ldots,2n\}\}. \end{aligned}$$

We arrange elements in $\mathscr{I}$ and $\mathscr{J}$ by the lexicographic order, respectively; and then re-label elements in $\mathscr{I}$ and $\mathscr{J}$ in turn by $1,2,\ldots,m^n$, respectively. Thus, the tensor T can be represented as a matrix $\bar{T} := (\bar{T}_{ij})$ of $m^n \times m^n$. Conversely, given an arbitrary $m^n \times m^n$ matrix $\bar{T}$, then it must be the matrix representation of some $2n$-th order m dimensional tensor in the above way. It is the balanced matrix flattening of an even order tensor.

Throughout this paper, for any tensor $T \in \mathscr{T}_{l,m}$, we denote its vector representation by $\hat{T} \in \mathbb{K}^{m^l}$; and if $l = 2n$, we denote its matrix representation by $\bar{T} \in \mathbb{K}^{m^n \times m^n}$. For any $S, Q \in \mathscr{T}_{l,m}$, define $\langle S, Q\rangle := \sum_{i=1}^{m^l} \hat{S}_i \hat{Q}_i$. Then, it is easy to show that $\langle \cdot,\cdot\rangle$ defined as above is an inner product in $\mathscr{T}_{l,m}$. The norm induced by this inner product is $\|T\| := \sqrt{\sum_{i=1}^{m^l} \hat{T}_i^2}$ for any $T \in \mathscr{T}_{l,m}$. If $l = 2n$, i.e., $S, Q \in \mathscr{T}_{2n,m}$, then the above inner product definition is equivalent to $\langle S, Q\rangle := \sum_{i=1,j=1}^{m^n} \bar{S}_{ij} \bar{Q}_{ij}$; and the norm induced by the inner product is $\|T\| := \sqrt{\sum_{i=1,j=1}^{m^n} \bar{T}_{ij}^2}$ for any $T \in \mathscr{T}_{2n,m}$, which is called the Hilbert-Schmidt norm. For any two tensors $A := (A_{i_1 i_2 \ldots i_n})$ and $B := (B_{i_{n+1} i_{n+2} \ldots i_{2n}})$ in $\mathscr{T}_{n,m}$, define $C := A \otimes B$ by $(C_{i_1 i_2 \ldots i_{2n}}) := (A_{i_1 i_2 \ldots i_n} B_{i_{n+1} i_{n+2} \ldots i_{2n}})$.

Definition 2.1. Let $\mathscr{T}_{2n,m}$ and $\mathscr{T}_{n,m}$ be the spaces of $2n$-th order m dimensional and n-th order m dimensional tensors in the field $\mathbb{K}$, respectively. For any tensor $T \in \mathscr{T}_{2n,m}$, if there is a number $\lambda \in \mathbb{C}$ and a nonzero tensor $D \in \mathscr{T}_{n,m}$, such that $\bar{T}\hat{D} = \lambda \hat{D}$. Then, λ is called a K-eigenvalue of T, and D is called a K-eigenvector of T with respect to the eigenvalue λ.

In Definition 2.1, the eigenvalues and eigenvectors of even order tensors were defined by using the methods used in matrix theory [30, 31]. Then, matrix theory can be applied [15].

For any $T \in \mathscr{T}_{2n,m}$, there exist m^n (counted with multiplicities) K-eigenvalues of T.

The space of all symmetric tensors in $\mathscr{T}_{2n,m}$ is denoted by $\mathscr{S}_{2n,m}$. For any $T \in \mathscr{S}_{2n,m}$, suppose that K-eigenvalues of T are $\{\lambda_i : i \in \{1,2,\ldots,m^n\}\}$ and the corresponding orthonormal system of K-eigenvectors is $\{D^{(1)}, D^{(2)}, \ldots, D^{(m^n)}\}$. Then, we have $T = \sum_{i=1}^{m^n} \lambda_i D^{(i)} \otimes D^{(i)}$.

2.2. **Decomposition of even degree homogenous polynomial.** For an n-th order m dimensional tensor T and a vector $x \in \mathbb{K}^m$ ($\mathbb{K}$ is $\mathbb{R}$ or $\mathbb{C}$), we define

$$Tx^n := \langle T, X\rangle,$$

where X is an n-th order m dimensional tensor defined by

$$X := \{X_{i_1 i_2 \ldots i_n} := x_{i_1} x_{i_2} \ldots x_{i_n} \mid i_j \in \{1,2,\ldots,m\}, \text{ and } j \in \{1,2,\ldots,n\}\}.$$

Actually, X can also be written as the tensor product of n vectors x, i.e., $X := x^n := \underbrace{x \otimes x \otimes \cdots \otimes x}_{n \text{ copies}}$.
It is the Veronese image of a vector of order n.

Theorem 2.2. *Let $\mathscr{T}_{2n,m}$ be the space of $2n$-th order m dimensional tensors in the field $\mathbb{K}$. For any $T \in \mathscr{T}_{2n,m}$, if there exist k numbers $\{\lambda_i : i \in \{1,2,\ldots,k\}\}$ and n-th order m dimensional tensors $\{D^{(1)}, D^{(2)}, \ldots, D^{(k)}\}$ such that $T = \sum_{i=1}^k \lambda_i D^{(i)} \otimes D^{(i)}$. Then for any $x \in \mathbb{K}^m$, we have $Tx^{2n} = \sum_{i=1}^k \lambda_i (D^{(i)} x^n)^2$.*

Proof. Let X denote the tensor generated by the tensor product of $2n$ vectors x; and Y denote the tensor generated by the tensor product of n vectors x. Suppose that the matrix representation of X is denoted by $\bar{X}$ and the vector representation of Y is denoted by $\hat{Y}$. Then, it is easy to see that $\bar{X} = \hat{Y}\hat{Y}^T$. Furthermore,

$$\begin{aligned}
Tx^{2n} &= \langle T, X \rangle = \left\langle \sum_{i=1}^k \lambda_i D^{(i)} \otimes D^{(i)}, X \right\rangle = \sum_{i=1}^k \lambda_i \left\langle D^{(i)} \otimes D^{(i)}, X \right\rangle \\
&= \sum_{i=1}^k \lambda_i \sum_{k,j=1}^{m^n} \left(\hat{D}_k^{(i)} \hat{D}_j^{(i)} \right) \bar{X}_{kj} = \sum_{i=1}^k \lambda_i \sum_{k,j=1}^{m^n} \left(\hat{D}_k^{(i)} \hat{D}_j^{(i)} \right) (\hat{Y}_k \hat{Y}_j) \\
&= \sum_{i=1}^k \lambda_i \sum_{k,j=1}^{m^n} \left(\hat{D}_k^{(i)} \hat{Y}_k \right) \left(\hat{D}_j^{(i)} \hat{Y}_j \right) = \sum_{i=1}^k \lambda_i \sum_{k=1}^{m^n} \sum_{j=1}^{m^n} \left(\hat{D}_k^{(i)} \hat{Y}_k \right) \left(\hat{D}_j^{(i)} \hat{Y}_j \right) \\
&= \sum_{i=1}^k \lambda_i \sum_{k=1}^{m^n} \left(\hat{D}_k^{(i)} \hat{Y}_k \right) \sum_{j=1}^{m^n} \left(\hat{D}_j^{(i)} \hat{Y}_j \right) = \sum_{i=1}^k \lambda_i \left(\sum_{j=1}^{m^n} (\hat{D}_j^{(i)} \hat{Y}_j) \right)^2 \\
&= \sum_{i=1}^k \lambda_i \left(D^{(i)} x^n \right)^2,
\end{aligned}$$

where the fourth equality follows from the definitions of $\otimes$ and inner product, the last follows from the definition of $D^{(i)} x^n$. The proof is complete. □

From the proof of the above theorem, we have the following corollary.

Corollary 2.3. *Let $\mathscr{T}_{2n,m}$ be the space of $2n$-th order m dimensional tensors in the field $\mathbb{K}$. For any $T \in \mathscr{T}_{2n,m}$, if there exist k numbers $\{\lambda_i : i \in \{1,2,\ldots,k\}\}$ and n-th order m dimensional tensors $\{D^{(1)}, D^{(2)}, \ldots, D^{(k)}\}$ and $\{F^{(1)}, F^{(2)}, \ldots, F^{(k)}\}$ such that $T = \sum_{i=1}^k \lambda_i D^{(i)} \otimes F^{(i)}$. Then for any $x \in \mathbb{K}^m$, we have $Tx^{2n} = \sum_{i=1}^k \lambda_i (D^{(i)} x^n)(F^{(i)} x^n)$.*

The following theorem and corollary demonstrate that a non-symmetric tensor applied to the Veronese image can be equivalently transformed into a balanced symmetric (symmetric with respect to the two groups of indices) tensor applied to the Veronese image. Actually, the conclusion of this theorem holds for the subspace of symmetric tensors, since the Veronese image is symmetric. There are situations in the sequel that need this refined conclusion. While, in most sitations, the conclusin of this theorem is sufficient.

Theorem 2.4. *Let $\mathscr{T}_{2n,m}$ be the space of $2n$-th order m dimensional tensors in the field $\mathbb{K}$ and $\mathscr{S}_{2n,m}^b$ be the corresponding balanced symmetric subspace of $\mathscr{T}_{2n,m}$. For any $T \in \mathscr{T}_{2n,m}$, there exists a tensor $S \in \mathscr{S}_{2n,m}^b$ such that $Tx^{2n} = Sx^{2n}$ for any $x \in \mathbb{K}^m$.*

Proof. For any $T \in \mathscr{T}_{2n,m}$, let $\bar{T}$ denote the matrix form of T. Suppose that S is a tensor with the matrix representation $\bar{S} := \frac{\bar{T}+\bar{T}^T}{2}$, then $S \in \mathscr{S}^b_{2n,m}$. For any $x \in \mathbb{K}^m$, let X denote the tensor generated by the tensor product of $2n$ vectors x, i.e., $\underbrace{x \otimes x \otimes \ldots \otimes x}_{2n \text{ copies}}$. It is easy to see that X is symmetric. Since $\bar{T} = \frac{\bar{T}+\bar{T}^T}{2} + \frac{\bar{T}-\bar{T}^T}{2}$, we have

$$\begin{aligned} Tx^{2n} &= \frac{1}{2}\left(2\langle \bar{S}, \bar{X}\rangle + \langle \bar{T}, \bar{X}\rangle - \langle \bar{T}^T, \bar{X}\rangle\right) \\ &= Sx^{2n} + \frac{1}{2}\left(\sum_{i,j=1}^{m^n} \bar{T}_{ij}\bar{X}_{ij} - \sum_{j,i=1}^{m^n} \bar{T}_{ji}\bar{X}_{ij}\right) \\ &= Sx^{2n} + \frac{1}{2}\left(\sum_{i,j=1}^{m^n} \bar{T}_{ij}\bar{X}_{ij} - \sum_{j,i=1}^{m^n} \bar{T}_{ji}\bar{X}_{ji}\right) \\ &= Sx^{2n}, \end{aligned}$$

where the third equality follows from the symmetry of X. The proof is complete. □

The proof actually gives the following.

Corollary 2.5. *Let $\mathscr{T}_{2n,m}$ be the space of $2n$-th order m dimensional tensors in the field $\mathbb{K}$ and $\mathscr{S}^b_{2n,m}$ be the corresponding balanced symmetric subspace of $\mathscr{T}_{2n,m}$. For any $T \in \mathscr{T}_{2n,m}$, there exists a tensor $S \in \mathscr{S}^b_{2n,m}$ such that $\langle T, D \otimes D\rangle = \langle S, D \otimes D\rangle$ for any $D \in \mathscr{T}_{n,m}$.*

A family of l-th order m dimensional real tensors $\{D^{(1)}, D^{(2)}, \ldots, D^{(k)}\}$ is said to be regular if for any $x \in \mathbb{R}^m \setminus \{0\}$, it follows that at least one of $D^{(1)}x^n, D^{(2)}x^n, \ldots, D^{(k)}x^n$ is nonzero. By Theorem 2.2, we may obtain the following theorem.

Theorem 2.6. *Suppose that the $2n$-th order m dimensional real tensor T is balanced symmetric.*

(i) *$\langle T, D \otimes D\rangle \geq 0$ holds for all n-th order m dimensional real tensor D if and only if every K-eigenvalue of T is nonnegative. If every K-eigenvalue of T is nonnegative, then there is a family of n-th order m dimensional tensors $\{D^{(1)}, D^{(2)}, \ldots, D^{(k)}\}$ such that $Tx^{2n} = \sum_{i=1}^k \left(D^{(i)}x^n\right)^2$ holds for all $x \in \mathbb{R}^m$.*

(ii) *$\langle T, D \otimes D\rangle > 0$ holds for all n-th order m dimensional real nonzero tensor D if and only if every K-eigenvalue of T is positive. If every K-eigenvalue of T is positive, then there is a regular family of n-th order m dimensional tensors $\{D^{(1)}, D^{(2)}, \ldots, D^{(m^n)}\}$ such that $Tx^{2n} = \sum_{i=1}^{m^n} \left(D^{(i)}x^n\right)^2$ holds for all $x \in \mathbb{R}^m$.*

The results established so far are for "*square*" tensors, i.e., tensors with all its indices the same dimension. However, these results can be generalized to "*rectangular*" tensor easily. An $m \times n$ dimensional rectangular tensor B of order $s+t$ means a multiarray with its entries $B_{i_1\ldots i_s j_1\ldots j_t} \in \mathbb{K}$ for all $i_1, \ldots, i_s \in \{1, \ldots, m\}$ and $j_1, \ldots, j_t \in \{1, \ldots, n\}$ [4, Section 2].

Similar to the discussions in Sections 2.1 and 2.2, using singular value decomposition of matrices [15], we can get the following results.

Theorem 2.7. *For any $m \times n$ dimensional rectangular tensor B of order $s+t$ in the field $\mathbb{R}$, we have*

(i) *there exists* $k \le \min\{m^s, n^t\}$ *positive numbers* λ_i *and tensors* $P_i \in \mathscr{T}_{s,m}$ *and* $Q_i \in \mathscr{T}_{t,n}$ *with* $i \in \{1,\dots,k\}$ *such that* $B = \sum_{i=1}^k \lambda_i P_i \otimes Q_i$*;*

(ii) *and hence,* $Bx^s \cdot y^t = \sum_{i=1}^k \lambda_i (P_i x^s)(Q_i y^t)$ *for any* $x \in \mathbb{R}^m$ *and* $y \in \mathbb{R}^n$*.*

3. A Positivstellensatz of Bi-Form Optimization over Unit Spheres

In this section, we discuss the bi-form optimization problem (1.1) with $k = s$ and $l = t$. For the convenience of reference, we state it more explicitly:

$$\begin{aligned} \min \quad & Bx^s \cdot y^t \\ \text{s.t.} \quad & \|x\|_s = 1, \quad \|y\|_t = 1, \end{aligned} \tag{3.1}$$

where B is an $m \times n$ dimensional $s+t$-th order rectangular tensor. When $s = t = 2$, we arrive at a bi-quadratic optimization problem. We denote by $B_s := \{x \in \mathbb{R}^m \mid \|x\|_s = 1\}$ and $B_t := \{y \in \mathbb{R}^n \mid \|y\|_t = 1\}$ the unit spheres. Recall that $s \ge 2$ and $t \ge 2$ are even integers throughout this section.

For the convenience of the subsequent analysis, we will denote by I_k a "diagonal" tensor in $\mathscr{T}_{k,m}$, i.e., $(I_k)_{i_1 \dots i_k} = 1$ if $i_1 = \dots = i_k$ and 0 otherwise. We will call a rectangular tensor B positive semidefinite if $Bx^s \cdot y^t \in \mathbb{R}[x,y]$ is a positive semidefinite (a.k.a. nonnegative) polynomial. Here $\mathbb{R}[x,y]$ represents the polynomial ring with the coefficients taking from the field $\mathbb{R}$. Similarly, we can define positive definite tensors. Note that, when a tensor is square, we will let $x = y$ in the above definition. Please refer to [25] for a comprehensive reference. For square tensors, we will denote by $\mathscr{P}_{l,m}$, abbreviated as $\mathscr{P}$ if there is no confusion, for the set of positive semidefinite tensors in $\mathscr{T}_{l,m}$. Consequently, we will denote by $\mathscr{P}^*$ the dual cone of $\mathscr{P}$ [25]. The positive semidefiniteness of a tensor is equivalent to its symmetrization or partial symmetrization for a rectangular tensor. While, for simplicity, the dual cones are always studied inside the subspace of symmetric tensors or the subspace of partially symmetric rectangular tensors. We will denote by $T \succeq \mathbf{0}$ for $T \in \mathscr{P}$, and $T \succ \mathbf{0}$ for $T \in \operatorname{int}(\mathscr{P})$, the interior of $\mathscr{P}$.

Using the decomposition results in the previous section, i.e., Theorem 2.7, we can decompose the objective function of (3.1) as

$$Bx^s \cdot y^t = \sum_{i=1}^k (P^i x^s)(Q^i y^t). \tag{3.2}$$

Here P^i's and Q^i's are (square) tensors of order s dimension m and order t dimension n, respectively. Note that P^i's and Q^i's can be chosen as symmetric. We first discuss the necessary conditions for optimality of (3.1).

Necessary Conditions for Optimality. Since problem (3.1) is actually minimizing a continuous function on a compact set, it has optimal solutions. Let λ_* be the optimal value of (3.1), and one of the corresponding optimal solution pairs be (x_*, y_*). Then, we must have

$$Bx^s \cdot y^t \ge Bx_*^s \cdot y_*^t = \lambda_*, \;\; \forall\, x \in B_s, \quad \forall y \in B_t.$$

This implies further that

$$Bx^s \cdot y^t \ge \lambda_* (\|x\|_s)^s = \lambda_* (\|y\|_t)^t, \;\; \forall\, x \in B_s, \quad \forall y \in B_t.$$

This, together with $\|x_*\|_s = 1$, implies that $Bx_*^s \cdot y^t - \lambda_* I_t y^t$ is a positive semidefinite polynomial in $\mathbb{R}[y]$. We also have by (3.2) that

$$Bx_*^s \cdot y^t - \lambda_* I_t y^t = \sum_{i=1}^{k} (P^i x_*^s)(Q^i y^t) - \lambda_* I_t y^t.$$

Hence, one of the necessary conditions for optimality is:

$$\sum_{i=1}^{k} (P^i x_*^s) Q^i - \lambda_* I_t \succeq \mathbf{0}. \tag{3.3}$$

A similar analysis can yield that another necessary condition is:

$$\sum_{i=1}^{k} (Q^i y_*^t) P^i - \lambda_* I_s \succeq \mathbf{0}. \tag{3.4}$$

We now summarize the above analysis into the following theorem.

Theorem 3.1. *(x_*, y_*) is an optimal solution pair of (3.1) with optimal value λ_* only if*

$$\begin{array}{c} \sum_{i=1}^{k} (P^i x_*^s) Q^i - \lambda_* I_t \succeq \mathbf{0}, \\ \sum_{i=1}^{k} (Q^i y_*^t) P^i - \lambda_* I_s \succeq \mathbf{0}, \\ \|x_*\|_s = 1, \ \textit{and} \ \|y_*\|_t = 1. \end{array} \tag{3.5}$$

In the literature, there are investigations on polynomial optimization over spheres similar to (3.1) via the eigenvalues of symmetric tensors [4, 21, 24]. The underlying principle is based on the first order necessary conditions for these polynomial optimization problems. Essentially, the necessary conditions considered there corresponding to KKT points. While, KKT points for (3.1) is different from the necessary conditions established in Theorem 3.1. In the next subsection, we will see some nice properties of (3.5).

3.1. Primal-dual characterization.

For any given $x \in B_s$ and $y \in B_t$, we can find a $\lambda \in \mathbb{R}$ such that

$$\sum_{i=1}^{k} (P^i x^s) Q^i - \lambda I_t \succeq \mathbf{0} \ \text{ and } \ \sum_{i=1}^{k} (Q^i y^t) P^i - \lambda I_s \succeq \mathbf{0}. \tag{3.6}$$

This is because x and y are bounded, and I_t and I_s are both positive definite. Denote by $\Lambda(x,y)$ the set of $\lambda \in \mathbb{R}$ satisfing (3.6) for any $(x,y) \in B_s \times B_t$. Since P^i's and Q^i's are given data after we decompose (3.1) using Theorem 2.7, and x and y are bounded, we have that the set of λ which is feasible for (3.6) is bounded from above. This, together with the continuity of λ in (3.6), implies that the set $\Lambda(x,y)$ contains its finite supremum. Hence, we essentially have $\Lambda(x,y) = (-\infty, \lambda(x,y)]$ for some finite $\lambda(x,y)$. That is to say, $\lambda(x,y)$ is defined to be $\max_{\lambda \in \Lambda(x,y)} \lambda$.

We now have the following result.

Theorem 3.2. *Suppose the optimal value of minimization problem (3.1) is λ_*. Then,*

$$\lambda_* = \bar{\lambda} := \min_{x \in B_s,\, y \in B_t} \lambda(x,y) = \min_{x \in B_s,\, y \in B_t} \max_{\lambda \in \Lambda(x,y)} \lambda. \tag{3.7}$$

Proof. Suppose (x_*, y_*) is an optimal solution pair of (3.1) with optimal value λ_*. It is obvious to see that (x_*, y_*, λ_*) is a feasible solution triple of (3.6) by Theorem 3.1. We also have $Bx_*^s \cdot y_*^t = \lambda_*$ by the hypothesis, this implies that $\lambda_* = \lambda(x_*, y_*)$. Otherwise, $\lambda_* < \lambda(x_*, y_*)$. This, together with (3.6) and $Bx_*^s \cdot y_*^t = \lambda_*$, yields a contradiction. Hence,

$$\bar{\lambda} = \min_{x \in B_s,\, y \in B_t} \lambda(x,y) \leq \lambda_*. \tag{3.8}$$

On the other hand, for any feasible $(x,y) \in B_s \times B_t$, note the followings

$$\begin{aligned} \left[\sum_{i=1}^{k} (P^i x^s) Q^i - \lambda_* I_t\right] y^t &= \sum_{i=1}^{k} (P^i x^s)(Q^i y^t) - \lambda_* I_t y^t \\ &= Bx^s \cdot y^t - \lambda_* \\ &\geq 0 \end{aligned}$$

hold for any y satisfing $\|y\|_t = 1$. Hence, $\sum_{i=1}^{k} (P^i x^s) Q^i - \lambda_* I_t \succeq \mathbf{0}$. Similarly, we can prove that $\sum_{i=1}^{k} (Q^i y^t) P^i - \lambda_* I_s \succeq \mathbf{0}$. Those imply $\lambda_* \in \Lambda(x,y)$. So,

$$\max_{\lambda \in \Lambda(x,y)} \lambda \geq \lambda_*$$

for any $(x,y) \in B_s \times B_t$. Taking minimization of both sides of the above inequality over the unit spheres, we get

$$\min_{x \in B_s,\, y \in B_t} \max_{\lambda \in \Lambda(x,y)} \lambda \geq \lambda_*.$$

This, together with (3.8), implies that $\lambda_* = \bar{\lambda}$. The proof is complete. □

We will call (3.7) the *primal* problem of the optimization problem (3.1). As expected, we define the *dual* problem of minimization problem (3.1) as the following problem:

$$\underline{\lambda} := \max_{\lambda \in \mathbb{R}} \min_{x \in B_s,\, y \in B_t} \delta(\lambda, x, y), \tag{3.9}$$

where δ is a function defined as

$$\delta(\lambda, x, y) := \begin{cases} \lambda & \text{if } \lambda \in \Lambda(x,y), \\ -\infty & \text{otherwise} \end{cases} \tag{3.10}$$

for $x \in B_s$ and $y \in B_t$.

Theorem 3.3. *There is no duality gap between the primal and the dual problems of (3.1), i.e., $\bar{\lambda} = \underline{\lambda}$.*

Proof. From the proof of Theorem 3.2, we see that $\lambda_* = \bar{\lambda} \in \Lambda(x,y)$ for every $(x,y) \in B_s \times B_t$. Hence, $\delta(\bar{\lambda}, x, y) = \bar{\lambda}$ for every (x,y) on the joint unit spheres. Thus,

$$\underline{\lambda} = \max_{\lambda \in \mathbb{R}} \min_{x \in B_s,\, y \in B_t} \delta(\lambda, x, y) \geq \bar{\lambda}.$$

On the other hand,

$$\underline{\lambda} = \max_{\lambda \in \mathbb{R}} \min_{x \in B_s,\, y \in B_t} \delta(\lambda, x, y) \leq \max_{\lambda \in \mathbb{R}} \min_{x \in B_s,\, y \in B_t} \max_{\lambda \in \Lambda(x,y)} \lambda = \max_{\lambda \in \mathbb{R}} \bar{\lambda} = \bar{\lambda}.$$

The proof is complete. □

In the following, we have a characterization on the dual problem of (3.1), i.e., (3.9).

Theorem 3.4. $\min_{x\in B_s,\, y\in B_t} \delta(\lambda, x, y)$ *is finite (i.e., not equal $-\infty$) if and only if*

$$\sum_{i=1}^{k}(P^i x^s)Q^i - \lambda I_t \succeq \mathbf{0} \ \ and \ \ \sum_{i=1}^{k}(Q^i y^t)P^i - \lambda I_s \succeq \mathbf{0} \tag{3.11}$$

for any x and y on the joint unit spheres.

Proof. The result follows from the definitions of δ and Λ immediately. □

Theorem 3.4 can be informative for solving optimization problem (3.9), and hence, the original optimization problem. First note that $\bar{\lambda} = \underline{\lambda} = \lambda_*$ are bounded, we can find in polynomial-time an upper bound $\bar{p}$ and a lower bound $\underline{p}$ for λ_*. Then, if we can solve the problem in Theorem 3.4 efficiently, we can use bisection method to $[\underline{p}, \bar{p}]$ to get an ε approximation solution for λ_* with any given $\varepsilon > 0$, i.e., finding λ such that $|\lambda - \lambda_*| < \varepsilon$. However, we have to be aware of the hardness of solving problem (3.11). With this notice and the NP-Hardness of (3.1), the two equivalent problems stated in Theorem 3.4 are essentially as hard as (3.1).

3.2. **General approaches and a class of trackable cases.** Since detecting conditions (3.11) for the general case is NP-Hard, *is there a framework to approach that*? *Is there a special case which is polynomial-time solvable*? We will give some answers to these questions in this subsection.

First, what does the condition (3.11) mean? The following result is easily proven.

Lemma 3.5.

$$\sum_{i=1}^{k}(P^i x^s)Q^i - \lambda I_t \succeq \mathbf{0}$$

for any $x \in B_s$ means the tensor polynomial (i.e., a tensor whose entries are polynomials)

$$K(x) := \sum_{i=1}^{k}(P^i x^s)Q^i - \lambda(\|x\|_s)^s I_t \tag{3.12}$$

is positive semidefinite for every $x \in \mathbb{R}^m$.

Since the positive semidefiniteness of (3.12) for every $x \in \mathbb{R}^m$ is equivalent to

$$\begin{aligned} K(x)y^t &= \sum_{i=1}^{k}(P^i x^s)(Q^i y^t) - \lambda(\|x\|_s)^s I_t y^t \\ &= \sum_{i=1}^{k}(P^i x^s)(Q^i y^t) - \lambda(\|x\|_s)^s(\|y\|_t^t) \\ &\geq 0 \end{aligned}$$

for any $x \in \mathbb{R}^m$ and $y \in \mathbb{R}^n$, we have that $\sum_{i=1}^{k}(P^i x^s)Q^i - \lambda I_t \succeq \mathbf{0}$ for any $x \in B_s$ is equivalent to $\sum_{i=1}^{k}(Q^i y^t)P^i - \lambda I_s \succeq \mathbf{0}$ for any $y \in B_t$. Hence, considering one of the conditions in Theorem 3.4 is sufficient.

Proposition 3.6. *If*

$$K(x) = \sum_{k=1}^{w}(h_k(x))^t \tag{3.13}$$

with $h_k(x) \in \mathbb{R}^n[x]$ for $k \in \{1,\dots,w\}$, then $K(x)$ is positive semidefinite for all $x \in \mathbb{R}^m$. Actually, all the K-eigenvalues of $K(x)$ are nonnegative. We will call the representation (3.13) of $K(x)$ sums of powers*, abbreviated as* s.o.p.

The notion s.o.p has an intimate relationship with symmetric tensor decomposition [3]. Note that when $t = 2$, s.o.p reduces to the well known *sums of squares*, abbreviated as *s.o.s*, in the literature [20]. We will also define the s.o.s representation of $K(x)$ as

$$K(x) = \sum_{k=1}^{w} H_k(x) \otimes H_k(x) \tag{3.14}$$

with $H_k(x) \in \mathbb{R}^{n^{t/2}}[x]$ for $k \in \{1,\dots,w\}$. Note that if $K(x)$ has an s.o.p representation, then it has an s.o.s representation since t is even. Moreover, Proposition 3.6 is also true with (3.13) being replaced by (3.14).

In the following, we will derive some characterizations of the conditions in Theorem 3.4. Hence, by the preceding discussion, we get some heuristic numerical approaching schemes for the optimization problem (3.1).

Theorem 3.7. *If $t = 2$ and $K(x) \succ \mathbf{0}$ for any $x \in \mathbb{R}^m \setminus \{0\}$, then there exists $N \in \mathscr{N}_+$ such that $(\|x\|_2)^{2N} K(x)$ has an s.o.s representation. Moreover, if we define*

$$\kappa(K) := \max_{\|z\|_2=1} \frac{\max_{\|x\|_2=1} z^T \overline{K(x)} z}{\min_{\|x\|_2=1} z^T \overline{K(x)} z},$$

then $N \geq \frac{m\frac{s}{2}(s-1)}{2log2}\kappa(K) - \frac{m+s}{2}$ is sufficient.

Proof. From the proof of [20, Lemma 3.5], especially [20, (3.7)], we can get that there exists $N \in \mathscr{N}_+$ such that $(\|x\|_2)^{2N} K(x)$ has an s.o.s representation in the sense of (3.14). □

Theorem 3.7 is also a corollary of Theorem 3.9 below since $\mathscr{P}_{2,n}$ is self-dual. Nevertheless, the general case for $t \geq 2$ is much more involved. Based on the definition of K-eigenvalues and Theorem 3.7, we have the following result.

Theorem 3.8. *For general even $t \geq 2$, if all K-eigenvalues of $K(x)$ are positive for any $x \in \mathbb{R}^m \setminus \{0\}$, then there exists $N \in \mathscr{N}_+$ such that $(\|x\|_2)^{2N} K(x)$ has an s.o.s representation.*

Proof. Since $t \geq 2$ is even, we can write the matrix representation of $K(x)$ as $\overline{K(x)}$, see Section 2 for details. Now, since all K-eigenvalues of $K(x)$ are positive for any $x \in \mathbb{R}^m \setminus \{0\}$, $\overline{K(x)}$ is positive definite for any $x \in \mathbb{R}^m \setminus \{0\}$. So, by Theorem 3.7, there exists $N \in \mathscr{N}_+$ such that $(\|x\|_2)^{2N}\overline{K(x)}$ has an s.o.s representation. However, the corresponding tensor of matrix $(\|x\|_2)^{2N}\overline{K(x)}$ is exactly $(\|x\|_2)^{2N} K(x)$. Thus, by (3.14) and isomorphism of space of tensors and space of matrices established in Section 2, we get the desired result. □

Under what condition(s), $K(x)$ or $h(x)^2 K(x)$ with some $h(x) \in \mathbb{R}[x]$ has an s.o.p representation? We give a result in the following theorem.

Theorem 3.9. *For general even $t \geq 2$, if $K(x) \in int(\mathscr{P}^*)$ for any $x \in \mathbb{R}^m \setminus \{0\}$, then there exists $N \in \mathscr{N}_+$ such that $(\|x\|_2)^{2N} K(x)$ has an s.o.p representation. Define*

$$\sigma(K) := \max_{T \in \mathscr{P} \setminus \{\mathbf{0}\}} \frac{\max_{\|x\|_2=1} K(x) \bullet T}{\max_{\|x\|_2=1} K(x) \bullet T},$$

then $N \geq N_* := \left\lceil \frac{2^{\frac{m\frac{s}{2}(s-1)}{2log_2}\sigma(K)-\frac{m+s}{2}-s}}{t} \right\rceil t+s$ *is sufficient.*

Proof. Write $K(x) = \sum_i f_i(x)K_i$, where $f_i(x) \in \mathbb{R}[x]$ are homogeneous and K_i's are symmetric tensors. Let $G(x) := \|x\|_2^2$. For any polynomial $p(x)$, we define the differential operator $p(\partial)$ as x_j being replaced by $\frac{\partial}{\partial x_j}$. Hence, $G(\partial) = \Delta$ is the Laplacian operator. $K(\partial)$ is defined to be $\sum_i f_i(\partial)K_i$.

Let d be an arbitrary given positive integer, and $H_d(\mathbb{R}^m) \subset \mathbb{R}[x]$ be the set of homogeneous forms of degree d. For any polynomial $h \in H_d(\mathbb{R}^m)$, from the results in [26, Pages 267-268], we have

$$h(\partial)G^N = \Phi_N(h)G^{N-d}, \text{ where } \Phi_N(h) = \sum_{j\geq 0} \frac{(N)_{d-j}}{2^{2j-d}j!}\Delta^j(h)G^j.$$

Here, $(N)_w := N(N-1)\cdots(N-(w-1))$ and Φ_N is a linear map from $H_d(\mathbb{R}^m)$ to itself. When $N > d$, Φ_N is invertible with explicit formula for its inverse. Please refer to [26] for details. With the *Hilbert Identity*:

$$G(x)^N = \sum_{i=1}^{q} \mu_i(\alpha_{i1}x_1 + \cdots + \alpha_{im}x_m)^{2N},$$

where $\mu_i > 0$ and $\alpha_{ij} \in \mathbb{R}$, we can have the followings:

$$\begin{aligned} h(\partial)G^N &= h(\partial)\left(\textstyle\sum_{i=1}^q \mu_i(\alpha_{i1}x_1+\cdots+\alpha_{im}x_m)^{2N}\right), \\ \Phi_N(h)G^{N-d} &= (2N)_d \textstyle\sum_{i=1}^q \mu_i h(\alpha_{i1},\ldots,\alpha_{im})(\alpha_{i1}x_1+\cdots+\alpha_{im}x_m)^{2N-d}. \end{aligned}$$

Let $N > s$ and $d = s$, then Φ_N is a bilinear map between $H_s(\mathbb{R}^m)$ and itself. Now, let $h = \Phi_N^{-1}(f_j)$, we get

$$f_j(x)G^{N-s}(x) = (2N)_s \sum_{i=1}^{q} \mu_i \Phi_N^{-1}(f_j)(\alpha_{i1},\ldots,\alpha_{im})(\alpha_{i1}x_1+\cdots+\alpha_{im}x_m)^{2N-s}.$$

Hence,

$$K(x)G^{N-s} = (2N)_s \sum_{i=1}^{q} \mu_i \sum_j K_j \Phi_N^{-1}(f_j)(\alpha_{i1},\ldots,\alpha_{im})(\alpha_{i1}x_1+\cdots+\alpha_{im}x_m)^{2N-s}. \tag{3.15}$$

From the general formula for Φ_N^{-1} in [26], and since f_j's are homogeneous polynomial of degree s, we have the following formula:

$$\begin{aligned} \Phi_N^{-1}(f_j) &= \frac{1}{(N)_s 2^s} \sum_{i=0}^{s/2+1} \frac{(-1)^i \Delta^i(f_j)G^i}{2^{2i} i! (\frac{m}{2}+N-1)_i} \\ &= \frac{1}{(N)_s 2^s} f_j + \frac{1}{(N)_s 2^s} \sum_{i=1}^{s/2+1} \frac{(-1)^i \Delta^i(f_j)G^i}{2^{2i} i! (\frac{m}{2}+N-1)_i}. \end{aligned}$$

So,

$$\lim_{N\to\infty} (N)_s 2^s \Phi_N^{-1}(f_j) = f_j.$$

Thus,

$$\lim_{N\to\infty}(N)_s 2^s \sum_i K_i \Phi_N^{-1}(f_i)(x) = K(x).$$

This, together with (3.15), yields

$$\begin{aligned}
&\lim_{N\to\infty} K(x)G^{N-s}\\
=\ &\lim_{N\to\infty}(2N)_s \sum_{i=1}^{q}\mu_i \sum_j K_j \Phi_N^{-1}(f_j)(\alpha_{i1},\dots,\alpha_{im})(\alpha_{i1}x_1+\cdots+\alpha_{im}x_m)^{2N-s}\\
=\ &\lim_{N\to\infty}\frac{(2N)_s}{(N)_s 2^s}\sum_{i=1}^{q}\mu_i (N)_s 2^s \sum_j K_j \Phi_N^{-1}(f_j)(\alpha_{i1},\dots,\alpha_{im})(\alpha_{i1}x_1+\cdots+\alpha_{im}x_m)^{2N-s}.
\end{aligned}$$

Now, for any $N \geq N_0$ with $N_0 > s$ sufficiently large, we can get

$$(N)_s 2^s \sum_j K_j \Phi_N^{-1}(f_j)(\alpha_{i1},\dots,\alpha_{im}) \in \mathrm{int}(\mathscr{P}^*)$$

since $K(\alpha_{i1},\dots,\alpha_{im}) \in \mathrm{int}(\mathscr{P}^*)$ and Φ_N is linear [26]. Similar to the discussion in [20, Page 1297] and [26, Section 7], we can conclude that

$$N_0 \geq \frac{m^{\frac{s}{2}}(s-1)}{2\log 2}\sigma(K) - \frac{m+s}{2}$$

is sufficient. Hence, with each $N \geq N_0$, we have by [25] that

$$(N)_s 2^s \sum_i K_i \Phi_N^{-1}(f_i)(\alpha_{i1},\dots,\alpha_{im}) \in \mathrm{int}\mathscr{P}^* = \sum_{j=1}^{r} h_j^t$$

for some $h_j \in \mathbb{R}^n$ with $j \in \{1\dots,r\}$. Since t is even, we can conclude that the defined $N_* \geq N_0$ such that t divides $2N_* - s$. Now, for this N_*, we can write $K(x)G(x)^{N_*-s}$ as

$$K(x)G(x)^{N_*-s} = \frac{(2N_*)_s}{(N_*)_s 2^s}\sum_{i=1}^{q}\mu_i \sum_{j=1}^{r}\left(h_j(\alpha_{i1}x_1+\cdots+\alpha_{im}x_m)^{\frac{2N_*-s}{t}}\right)^t.$$

Since $\mu_i > 0$, we obviously arrive at an s.o.p representation of $K(x)G(x)^{N_*-s}$. The proof is complete. $\square$

Theorem 3.9 actually presents a Positivstellensatz for the considered problem. We refer to [18, 26] for some classical results on this topic.

We note that another reason for investigating Theorem 3.9 is that it may give some light to tensor conic optimization problems [25], and also symmetric tensor decomposition [3].

Theorem 3.10. *If $K(x) \in \mathbb{R}^{n^t}[x]$ defined in (3.12) is positive semidefinite for every $x \in \mathbb{R}^m$, then there exists a homogeneous polynomial $h(x,y) \in \mathbb{R}[x,y]$ such that $h^2(x,y)K(x)y^t$ is a sum of squares of polynomials in $\mathbb{R}[x,y]$.*

Proof. Due to Artin's theorem for Hilbert's 17th problem [2], we only need to prove that $K(x)y^t$ is positive semidefinite in $\mathbb{R}[x,y]$. This is obvious. The proof is complete. $\square$

At the end of this section, we point out that one class of problems (3.1) can be solved in polynomial-time completely. At first, we prove the following result.

Theorem 3.11. *If $K(x) \in \mathbb{R}^{n^t}[x]$ defined in (3.12) is positive semidefinite for every $x \in \mathbb{R}^m$, $t = 2$ and $m = 2$, then $K(x)y^2$ is a sum of squares of polynomials in $\mathbb{R}[x, y]$. In particular, we have*

$$K(x)y^2 = \sum_{i=1}^{r} (H_i x^{s/2} y)^2 \tag{3.16}$$

for a set of tensors $H_i \in \mathbb{R}^{m^{s/2} \times n}$.

Proof. We only need to prove that $K(x)y^2$ is a positive semidefinite polynomial in $\mathbb{R}[x, y]$ [6, Theorem 7.1]. This is easy, since $K(x)$ is positive semidefinite for every $x \in \mathbb{R}^m$. The proof is complete. □

We note that, when $n = 2$, we can get a result similar to Theorem 3.11. Due to Theorem 3.11, when $m = 2$ and $t = 2$, we can characterize $K(x) \succeq \mathbf{0}$ through SDP based on the representation matrix developed in Sections 2 and 3 of the tensors in (3.16). So, we can solve the underlying bi-form optimization problem (3.1) completely in polynomial-time. When $s = 2$, the underlying problems are a class of bi-quadratic optimization problems. Nevertheless, to our best knowledge, this class of problems which can be solve in polynomial-time was not pointed out in the literature of bi-quadratic optimization.

4. Bi-Form Optimization with Nonpositive Coefficients

In this section, we consider the following special case of (1.1) with nonnegative tensor T.

$$\begin{aligned} \min \quad & -Tx^p \cdot y^q := \textstyle\sum_{i_1,\dots,i_p=1}^m \sum_{j_1,\dots,j_q=1}^n B_{i_1\dots i_p j_1 \dots j_q} x_{i_1} \cdots x_{i_p} y_{j_1} \cdots y_{j_q} \\ \text{s.t.} \quad & \|x\|_{p+q} = 1, \quad \|y\|_{p+q} = 1, \end{aligned}$$

which is equivalent to

$$\begin{aligned} \max \quad & Tx^p \cdot y^q := -\textstyle\sum_{i_1,\dots i_p=1}^m \sum_{j_1\dots,j_q=1}^n B_{i_1\dots i_p j_1 \dots j_q} x_{i_1} \cdots x_{i_p} y_{j_1} \cdots y_{j_q} \\ \text{s.t.} \quad & \|x\|_{p+q} = 1, \quad \|y\|_{p+q} = 1 \end{aligned} \tag{4.1}$$

in the sense that they have the same optimal solution set and the optimal value is the minus of the other. Without loss of generality, T is assumed to be partially symmetric. We refer to [4, 7, 8, 10] and references therein for more on nonnegative tensors.

It is easy to check that the first order necessary optimality condition for (4.1) is the following KKT system [1]:

$$\begin{cases} Tx^{p-1}y^q &= \lambda x^{[p+q-1]}, \\ Tx^p y^{q-1} &= \lambda y^{[p+q-1]}. \end{cases} \tag{4.2}$$

If $(\lambda, x, y) \in \mathbb{C} \times \mathbb{C}^m \setminus \{0\} \times \mathbb{C}^n \setminus \{0\}$ satisfies (4.2), then it is named a singular triple of the tensor T. Denote by $\sigma(T)$ the set of singular values of T and $\rho(T)$ the largest absolute value of elements in $\sigma(T)$. We call $\sigma(T)$ and $\rho(T)$ the spectra and the spectral radius of the tensor T, respectively [4]. It is easy to see that $\rho(T)$ and its corresponding singular vectors are the global optimal value and its corresponding optimal solution, respectively.

Denote by $\{e_i\}_1^m$ and $\{f_j\}_1^n$ the bases of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively, and let $e_i^p := \underbrace{e_i \otimes \cdots \otimes e_i}_{p \text{ times}}$ and $f_j^q := \underbrace{f_j \otimes \cdots \otimes f_j}_{q \text{ times}}$, where $\otimes$ denotes the tensor outer product. Given T, for any $j \in$

$\{1,\ldots,n\}$, let $T^j := (T_{i_1\ldots i_p j\ldots j})$ be the p-th order m dimensional square sub-tensor, and for any $i\in\{1,\ldots,m\}$, let $T_i := (T_{i\ldots i j_1\ldots j_q})$ be the q-th order n dimensional square sub-tensor.

Definition 4.1. Square tensor T of order m and dimension n is *reducible* if there exists a nonempty proper index subset $I\subset\{1,\ldots,n\}$ such that

$$T_{i_1 i_2\ldots i_m} = 0, \quad \forall i_1\in I, \quad \forall i_2,\ldots,i_m\notin I. \tag{4.3}$$

If T is not reducible, then T is called *irreducible*.

Definition 4.2. A nonnegative rectangular tensor T is called irreducible if all the square tensors T_i with $i\in\{1,\ldots,m\}$ and T^j with $j\in\{1,\ldots,n\}$ are irreducible.

Definition 4.3. For any nonnegative rectangular tensor T, we associate it an $(m+n)\times(m+n)$ matrix $R(T)$ called the *representation* of T as:

$$R(T)_{ij} := \begin{cases} \sum_{j\in\{i_2,\ldots,i_p\}} T_{ii_2\ldots i_p j_1\ldots j_q} & \text{if } i,j\in\{1,\ldots,m\}; \\ \sum_{j\in\{j_1,\ldots,j_q\}} T_{ii_2\ldots i_p j_1\ldots j_q} & \text{if } i\in\{1,\ldots,m\}, \text{ and } j\in\{m+1,\ldots,m+n\}; \\ \sum_{j\in\{j_2,\ldots,j_p\}} T_{i_1 i_2\ldots i_p i j_2\ldots j_q} & \text{if } i,j\in\{m+1,\ldots,m+n\}; \\ \sum_{j\in\{i_2,\ldots,i_p\}} T_{i_1 i_2\ldots i_p i j_2\ldots j_q} & \text{if } i\in\{m+1,\ldots,m+n\}, \text{ and } j\in\{1,\ldots,m\}. \end{cases} \tag{4.4}$$

We call T *weakly irreducible* if $R(T)$ is irreducible and *weakly primitive* if $R(T)$ is primitive.

Theorem 4.4. [4, Theorem 4] *Assume that the nonnegative tensor T is irreducible, then there exists a solution (λ,x,y) of system (4.2), satisfying $\lambda>0$ and $(x,y)\in\mathbb{R}^m_{++}\times\mathbb{R}^n_{++}$. Moreover, if λ_0 is a singular value with positive[1] left and right singular vectors, then $\lambda_0=\lambda$. The positive left and right singular vectors are unique up to a multiplicative constant.*

Theorem 4.5. [4, Theorem 5] *Assume that T is an irreducible nonnegative rectangular tensor of order (p,q) and dimension $m\times n$, then*

$$\begin{aligned} & \min_{(x,y)\in\mathbb{R}^n_+\setminus\{0\}\times\mathbb{R}^n_+\setminus\{0\}} \max_{i,j} \left(\frac{(Tx^{p-1}y^q)_i}{x_i^{p+q-1}}, \frac{(Tx^p y^{q-1})_j}{y_j^{p+q-1}} \right) \\ = \ & \lambda \\ = \ & \max_{(x,y)\in\mathbb{R}^n_+\setminus\{0\}\times\mathbb{R}^n_+\setminus\{0\}} \min_{i,j} \left(\frac{(Tx^{p-1}y^q)_i}{x_i^{p+q-1}}, \frac{(Tx^p y^{q-1})_j}{y_j^{p+q-1}} \right) \end{aligned} \tag{4.5}$$

where λ is the unique positive singular value corresponding to positive left and right singular vectors.

Theorem 4.6. [4, Theorem 6] *Assume that T is an irreducible nonnegative rectangular tensor, and λ is the positive singular value with positive left and right singular vectors. Then $|\mu|\le\lambda$ for all $\mu\in\sigma(T)$. Hence, $\lambda=\rho(T)$.*

Actually, we have the following result, see [10] and references herein.

Theorem 4.7. *Assume that the nonnegative tensor T is weakly irreducible, then*

[1] Actually, the result holds with "positive" being replaced by "nonnegative".

- *there exists a solution (λ, x, y) of system (4.2), satisfying $\lambda > 0$ and $(x,y) \in \mathbb{R}^m_{++} \times \mathbb{R}^n_{++}$. Moreover, when T is weakly primitive, if λ_0 is a singular value with nonnegative left and right singular vectors, then $\lambda_0 = \lambda$. The strict positive left and right singular vectors are unique up to a multiplicative constant;*
- *we have*

$$\begin{aligned} & \min_{(x,y)\in\mathbb{R}^n_+\setminus\{0\}\times\mathbb{R}^n_+\setminus\{0\}} \max_{i,j} \left(\frac{(Tx^{p-1}y^q)_i}{x_i^{p+q-1}}, \frac{(Tx^py^{q-1})_j}{y_j^{p+q-1}} \right) \\ = \ & \lambda \\ = \ & \max_{(x,y)\in\mathbb{R}^n_+\setminus\{0\}\times\mathbb{R}^n_+\setminus\{0\}} \min_{i,j} \left(\frac{(Tx^{p-1}y^q)_i}{x_i^{p+q-1}}, \frac{(Tx^py^{q-1})_j}{y_j^{p+q-1}} \right) \end{aligned} \tag{4.6}$$

  where λ is the unique positive singular value corresponding to strict positive left and right singular vectors;
- *$|\mu| \leq \lambda$ for all $\mu \in \sigma(T)$. Hence, $\lambda = \rho(T)$.*

Remark 4.8. For weakly irreducible nonnegative tensor T, we have

$$\begin{aligned} p^* := \lambda &= \min_{(x,y)\in\mathbb{R}^n_{++}\times\mathbb{R}^m_{++}} \max_{i,j} \left(\frac{(Tx^{p-1}y^q)_i}{x_i^{p+q-1}}, \frac{(Tx^py^{q-1})_j}{y_j^{p+q-1}} \right) \\ &= \min_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^m} f(x,y) := \max_{i,j} \left(\frac{(T\exp(x)^{p-1}\exp(y)^q)_i}{\exp(x)_i^{p+q-1}}, \frac{(T\exp(x)^p\exp(y)^{q-1})_j}{\exp(x)_j^{p+q-1}} \right). \end{aligned}$$

Note that $f(x,y)$ is a piecewise convex and piecewise smooth function on $\mathbb{R}^n \times \mathbb{R}^m$. So, optimization problem (1.1) with $k = l = s+t$ is actually an unconstrained convex but nonsmooth optimization for nonnegative tensors.

Algorithm 1 A Higher Order Power Method (HOPM)

Step 0: Initialization: choose $x^0 \in \mathbb{R}^n_{++}$ and $y^0 \in \mathbb{R}^m_{++}$. Let $k := 0$.

Step 1: Compute

$$\bar{x}^{k+1} := T(x^k)^{p-1}(y^k)^q, \quad ,\bar{y}^{k+1} := T(x^k)^p(y^k)^{q-1},$$

$$x^{k+1} := \frac{\left(\bar{x}^{k+1}\right)^{\left[\frac{1}{p+q-1}\right]}}{\left\| \left(\bar{x}^{k+1}\right)^{\left[\frac{1}{p+q-1}\right]}, \left(\bar{y}^{k+1}\right)^{\left[\frac{1}{p+q-1}\right]} \right\|_{p+q}}, \quad y^{k+1} := \frac{\left(\bar{y}^{k+1}\right)^{\left[\frac{1}{p+q-1}\right]}}{\left\| \left(\bar{x}^{k+1}\right)^{\left[\frac{1}{p+q-1}\right]}, \left(\bar{y}^{k+1}\right)^{\left[\frac{1}{p+q-1}\right]} \right\|_{p+q}}$$

$$\alpha^{k+1} := \max\left\{ \frac{\left(\bar{x}^{k+1}\right)_i}{(x^k)_i^{p+q-1}}, \frac{\left(\bar{y}^{k+1}\right)_j}{(y^k)_j^{p+q-1}} \right\} \quad \text{and} \quad \beta^{k+1} := \min\left\{ \frac{\left(\bar{x}^{k+1}\right)_i}{(x^k)_i^{p+q-1}}, \frac{\left(\bar{y}^{k+1}\right)_j}{(y^k)_j^{p+q-1}} \right\}.$$

Step 2: If $\alpha^{k+1} = \beta^{k+1}$, stop. Otherwise, let $k := k+1$, go to Step 1.

Theorem 4.9. *Suppose that T is a weakly primitive nonnegative tensor, and the sequence $\{x^k, y^k\}$ is generated by Algorithm 1. Then, $\{x^k, y^k\}$ converges to the unique optimal solution vectors $x^* \in \mathbb{R}^n_{++}$ and $y^* \in \mathbb{R}^m_{++}$, and there exist a constant $\theta \in (0,1)$ and a positive integer M such that*

$$d((x^k,y^k),(x^*,y^*)) \leq \theta^{\frac{k}{M}} \frac{d((x^0,y^0),(x^*,y^*))}{\theta}, \tag{4.7}$$

for all $k \geq 1$.

Proof. We can prove this theorem following from those in [7, 8]. $\square$

Lemma 4.10. *For any nonnegative rectangular tensor* T*, if* (λ, x, y) *is a nonnegative singular triple of* T*, then*

$$\|x\|_{p+q} = \|y\|_{p+q}. \tag{4.8}$$

Proof. From (4.2) and the nonnegativeness of (x, y), we can get the desired result whether $\lambda = 0$ or not. $\square$

Lemma 4.11. *For any nonnegative rectangular tensors* A *and* T*, if* $T \geq A$ *and* T *is irreducible, then* $\rho(A) \leq \rho(T)$.

Proof. Suppose $(\mu.x, y)$ is a singular triple of A such that $|\mu| = \rho(A)$. Then, we must have

$$\rho(A)|x|^{p+q-1} = \left|\mu x^{[p+q-1]}\right| = \left|Ax^{p-1}y^q\right| \leq A|x|^{p-1}|y|^q \leq T|x|^{p-1}|y|^q,$$

and

$$\rho(A)|y|^{p+q-1} = \left|\mu y^{[p+q-1]}\right| = \left|Ax^p y^{q-1}\right| \leq A|x|^p|y|^{q-1} \leq T|x|^p|y|^{q-1}.$$

Hence, by Theorem 4.5, we get $\rho(A) \leq \rho(T)$. $\square$

Theorem 4.12. *For any nonnegative rectangular tensor* T*,* $\rho(T)$ *is a singular value of* T *with a pair of nonnegative singular vectors.*

Proof. Denote by E the tensor of the same size of T with its elements being 1. Now, tensor $T(k) := T + \frac{1}{k}E$ is a positive tensor for any $k \geq 1$. Hence, $T(k)$ is irreducible. By Theorem 4.4, $T(k)$ has a positive singular triple (λ_k, x_k, y_k). By Lemma 4.10 and the homogeneity of (4.2), we can choose x_k, y_k such that $\|x_k\|_{p+q} = \|y_k\|_{p+q} = 1$. Hence, sequences $\{x_k\}$ and $\{y_k\}$ are bounded. Suppose, without loss of generality, that both $\{x_k\}$ and $\{y_k\}$ converge and to $x_* \geq 0$ and $y_* \geq 0$, respectively. Since $T(k) > T(k+1)$, by Lemma 4.11 and Theorem 4.12, $\{\lambda_k\}$ is nonincreasing and bounded below by $\rho(T)$ since $T(k) > T$ for any $k \geq 1$. Suppose $\lambda_k \to \lambda \geq 0$, then, by continuity of (4.2), we have

$$0 = \lim_{k\to\infty} T(k)x_k^{p-1}y_k^q - \lambda_k x_k^{[p+q-1]} = Tx_*^{p-1}y_*^q - \lambda x_*^{[p+q-1]},$$

and

$$0 = \lim_{k\to\infty} T(k)x_k^p y_k^{q-1} - \lambda_k y_k^{[p+q-1]} = Tx_*^p y_*^{q-1} - \lambda y_*^{[p+q-1]}.$$

Since $\|x_*\|_{p+q} = \|y_*\|_{p+q} = 1$, we have that (λ, x_*, y_*) is a singular triple of T. So, $\lambda \leq \rho(T)$. This, together with $\lambda \geq \rho(T)$, yields that $\rho(T) = \lambda$ and (x_*, y_*) is one corresponding nonnegative singular pair. $\square$

By the homogeneity of the objective function and the results in this section, we arrive at a result for the general case.

Theorem 4.13. *For general bi-form optimization problem* (1.1) *with* $k, l \geq s+t$*, if the tensor* B *in objective function is nonpositive, then, there is a globally linearly convergent algorithm which can find a feasible solution pair* (x, y) *for* (1.1)*, with its corresponding objective value within* $m^{s\frac{s+t}{k}-s}n^{t\frac{s+t}{k}-t}$ *the optimal value of* (1.1).

Proof. Let the objective function of (1.1) be f.

In this case, we will solve the problem when $k=l=s+t$ by the preceding algorithm for nonnegative tensors. Denote the corresponding optimal solution found by the higher order power method as $(\bar{x},\bar{y})$. It follows that

$$\|\bar{x}\|_{s+t}=\|\bar{y}\|_{s+t}=1.$$

Let $x:=\frac{\bar{x}}{\|\bar{x}\|_k}$ and $y:=\frac{\bar{y}}{\|\bar{y}\|_l}$. Then, (x,y) is a feasible solution for (1.1). Let the corresponding objective function value being p. Let the optimal objective function value of (1.1) be p^* with optimal solution (x^*,y^*). Then, we must have

$$p^*\leq p\leq 0.$$

While, $(\frac{x^*}{\|x^*\|_{s+t}},\frac{y^*}{\|y^*\|_{s+t}})$ is a feasible solution for the problem when $k=l=s+t$. It follows from the homogeneity of the objective function that

$$p\leq f(\frac{x^*}{\|x^*\|_{s+t}},\frac{y^*}{\|y^*\|_{s+t}})=\frac{1}{\|x^*\|^s_{s+t}\|y^*\|^t_{s+t}}f(x^*,y^*)=\frac{1}{\|x^*\|^s_{s+t}\|y^*\|^t_{s+t}}p^*.$$

On the other hand, it is a standard analysis to show that

$$\|x^*\|_k\leq\|x^*\|_{s+t}\leq m^{1-\frac{s+t}{k}}\|x^*\|_k.$$

Thus, by the nonpositivity of p^*, we have

$$p\leq m^{s\frac{s+t}{k}-s}n^{t\frac{s+t}{k}-t}p^*.$$

The result follows. □

5. Conclusions

In this paper, we recalled the concepts of K-eigenvalues and K-eigenvectors for even order tensors. It has some consequences on homogeneous polynomials of even degrees and the bi-form optimization problems. Some natural further questions arose from those discussions. For example, theoretical or numerical refinements of bi-quadratic optimization problems using various techniques from quadratic optimization [16, 28, 32, 33, 34]. Some further investigations should be put on bi-form optimization problems proposed in this paper as well, such as further properties on sums of powers of polynomials.

Acknowledgements

The author was partially supported by the Natural Science Foundation of Zhejiang Province, China (Grant No. LY22A010022) and the National Science Foundation of China (Grant No. 12171128).

References

[1] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1995.

[2] J. Bochnak, M. Coste, M.-F. Roy, *Real Algebraic Geometry*. Vol. 36. Springer Science & Business Media, Berlin, 1998.

[3] J. Brachat, P. Comon, B. Mourrain, E.P. Tsigaridas, Symmetric tensors decomposition. *Linear Algebra and its Applications* 433 (2010) 1851–1872.

[4] K. Chang, L. Qi and G. Zhou, Singular values of a real rectangular tensor, *Journal of Mathematical Analysis and Applications* 370 (2010) 284–294.

[5] K.C. Chang, T. Zhang, Multiplicity of singular values for tensors, *Communications in Mathematical Sciences* 7 (2009) 611–625.

[6] M.D. Choi, T.Y. Lam, B. Reznick, Real zeros of Positive semidefinite forms. I, *Mathematische Zeitschrift* 171 (1980) 1–26.

[7] S. Friedland, S. Gaubert, L. Han, Perron-Frobenius theorem for nonnegative multilinear forms and extensions, *Linear Algebra and Its Applications* 438 (2013) 738–749.

[8] S. Hu, Z.H. Huang, L. Qi, Strictly nonnegative tensors and nonnegative tensor partition, *Science China Mathematics* 57 (2014) 181–195.

[9] S. Hu, Certifying the global optimality of quartic minimization over the sphere, *Journal of the Operations Research Society of China* 10 (2022) 241–287.

[10] S. Hu, L. Qi, A necessary and sufficient condition for existence of a positive Perron vector, *SIAM Journal on Matrix Analysis and Applications* 37 (2016) 1747–1770.

[11] S. Hu, G. Li, Convergence rate analysis for the higher order power method in best rank one approximations of tensors, *Numerische Mathematik* 140 (2018) 993–1031.

[12] S. Hu, Nondegeneracy of eigenvectors and singular vector triples of tensors, *Science China Mathematics* 65 (2022) 2483–2492.

[13] S. Hu, K. Ye, Linear convergence of an alternating polar decomposition method for low rank orthogonal tensor approximations, *Mathematical Programming* https://doi.org/10.1007/s10107-022-01867-8, 2022

[14] G. H. Golub, C. F. Van Loan, *Matrix Computations*. Johns Hopkins Univ. Press, 1996.

[15] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1990.

[16] S. Kim, M. Kojima, Second order cone programming relaxation of nonconvex quadratic optimization problems, *Optimization Methods and Software* 15 (2001) 201–224.

[17] J. Landsberg, *Tensors: Geometry and Applications*, Graduate Studies in Mathematics vol 128, AMS, 2012.

[18] M. Laurent, Sums of squares, moment matrices and optimization over polynomials, in Emerging Applications of Algebraic Geometry, IMA Vol. Math. Appl., 149, M. Putinar and S. Sullivant, eds., pp. 157–270, Springer, New York, 2009.

[19] L.-H. Lim, Singular values and eigenvalues of tensors: a variational approach. Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, CAMSAP '05, 2005, 1: 129–132.

[20] C. Ling, J.W. Nie, L. Qi, Y. Ye, Bi-quadratic optimization over unit spheres and semidefinite programming relaxations, *SIAM Journal on Optimization* 20 (2009) 1286–1310.

[21] L. Qi, Eigenvalues of a real supersymmetric tensor, *Journal of Symbolic Computation* 40 (2005) 1302–1324.

[22] L. Qi, Rank and eigenvalues of a supersymmetric tensor, a multivariate homogeneous polynomial and an algebraic surface defined by them. *Journal of Symbolic Computation* 41 (2006) 1309–1327.

[23] L. Qi, Eigenvalues and invariants of tensors. *Journal of Mathematical Analysis and Applications* 325 (2007) 1363–1377.

[24] L. Qi, F. Wang, Y.J. Wang, Z-eigenvalue methods for a global polynomial optimization problem, *Mathematical Programming* 118 (2009) 301–316.

[25] L. Qi, Y. Ye, Space tensor conic programming, *Computational Optimization and Applications* 59 (2014) 307–319.

[26] B. Reznick, Some concrete aspects of Hilbert's 17th problem, *Contemporary Mathematics* 253, American Mathematical Society, Providence, RI, (2000) 251–272.

[27] S. Rota Buló. A game-theoretic framework for similarity-based data clustering. PhD thesis, University of Venice, 2009.

[28] J.F. Sturm, S. Zhang, On cones of nonnegative quadratic functions, *Mathematics of Operations Research* 28 (2003) 246-267.

[29] N.Z. Shor, *Nondifferentiable Optimization and Polynomial Problems*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.

[30] W. (Lord Kelvin) Thomson, Elements of a mathematical theory of elasticity, *Philosophical Transactions of the Royal Society of London*, 146 (1856) 481-498.

[31] W. (Lord Kelvin) Thomson, Elasticity, *Encyclopedia Briannica* Vol. 7, nineth ed., Adam and Charles Black, London, Edingburgh, 1878, pp. 796-825.

[32] Y. Ye, Approximating quadratic programming with bound and quadratic constraints, *Mathematical Programming* 84 (1999) 219-226.
[33] Y. Ye, S. Zhang, New results on quadratic minimization, *SIAM Journal on Optimization* 14 (2003) 245-267.
[34] S. Zhang, Quadratic maximization and semidefinite relaxation, *Mathematical Programming* 87 (2000) 453-465.