

## THE BEST RANK-1 APPROXIMATION OF A SYMMETRIC TENSOR AND RELATED SPHERICAL OPTIMIZATION PROBLEMS\*

XINZHEN ZHANG<sup>†</sup>, CHEN LING<sup>‡</sup>, AND LIQUN QI<sup>§</sup>

**Abstract.** In this paper, we show that for a symmetric tensor, its best symmetric rank-1 approximation is its best rank-1 approximation. Based on this result, a positive lower bound for the best rank-1 approximation ratio of a symmetric tensor is given. Furthermore, a higher order polynomial spherical optimization problem can be reformulated as a multilinear spherical optimization problem. Then, we present a modified power algorithm for solving the homogeneous polynomial spherical optimization problem. Numerical results are presented, illustrating the effectiveness of the proposed algorithm.

**Key words.** symmetric tensor, the best rank-1 approximation, the best symmetric rank-1 approximation, power algorithm

**AMS subject classifications.** 15A18, 15A69, 90C30

**DOI.** 10.1137/110835335

**1. Introduction.** A real  $m$ -order  $n$ -dimensional square tensor  $\mathcal{A}$  is a multidimensional array consisting of  $n^m$  real entries  $a_{i_1 \dots i_m} \in \mathfrak{R}$ , where  $i_k = 1, \dots, n$  for  $k = 1, \dots, m$ . Tensor  $\mathcal{A}$  is said to be symmetric if its element  $a_{i_1 i_2 \dots i_m}$  is invariant under any permutation of indices  $(i_1, i_2, \dots, i_m)$ . Symmetric tensors arise in higher order derivatives of smooth functions, moments, and cumulants of random vectors and have wide applications in signal and image processing, blind source separation (BSS), statistics, investment science, and so on; see [1, 8, 17, 18, 26] and references therein.

A tensor is said to be rank-1 if it can be expressed as an outer product of a number of vectors. Specifically, if these vectors are all equal, then the tensor is called a symmetric rank-1 tensor. Given an  $m$ -order  $n$ -dimensional square tensor  $\mathcal{A}$ , rank-1 tensor  $\mathcal{B} = \lambda x^{(1)} \dots x^{(m)}$  is said to be its best rank-1 approximation if it minimizes the least-squares cost function  $\|\mathcal{A} - \mathcal{B}\|_F$  over the manifold of rank-1 tensors. Similarly, symmetric rank-1 tensor  $\mathcal{C} = \mu x^m$  is said to be the best symmetric rank-1 approximation if it minimizes the least-squares cost function  $\|\mathcal{A} - \mathcal{C}\|_F$  over the manifold of symmetric rank-1 tensors. From optimization theory,  $\mathcal{B}$  and  $\mathcal{C}$  can be obtained by solving the optimization problems

$$(1) \quad \begin{aligned} \lambda := & \max_{x^{(1)}, \dots, x^{(m)} \in \mathfrak{R}^n} |\mathcal{A}x^{(1)}x^{(2)} \dots x^{(m)}| \\ \text{s.t.} & \|x^{(i)}\| = 1 \text{ for } i = 1, 2, \dots, m, \end{aligned}$$

---

\*Received by the editors May 24, 2011; accepted for publication (in revised form) by P. Comon May 21, 2012; published electronically July 26, 2012.

<http://www.siam.org/journals/simax/33-3/83533.html>

<sup>†</sup>Department of Mathematics, School of Science, Tianjin University, Tianjin, 300072, China (xzzhang@tju.edu.cn). This author's work was supported by the National Natural Science Foundation of China (grants 11101303 and 11171180) and the Independent Innovation Fund of Tianjin University.

<sup>‡</sup>Corresponding Author. School of Science, Hangzhou Dianzi University, Hangzhou, 310018, China (macling@hdu.edu.cn). This author's work was supported by the National Natural Science Foundation of China (grants 10871168, 11171083, and 10971187) and the Zhejiang Provincial Natural Science Foundation of China (grant Y6100366).

<sup>§</sup>Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong (maqilq@polyu.edu.hk). This author's work was supported by the Hong Kong Research Grant Council.

and

$$(2) \quad \mu \equiv \mu(\mathcal{A}) := \max_{x \in \mathbb{R}^n} |\mathcal{A}x^m| \\ \text{s.t.} \quad \|x\| = 1,$$

respectively.

The best (symmetric) rank-1 approximation to higher order (symmetric) square tensors plays important roles in theory and arises from many applications such as analytical BSS and deflation procedure; see, e.g., [2, 4, 6, 12, 33, 35].

Throughout this paper, and unless otherwise specified,  $\mathcal{A}$  is symmetric. For the case that  $m = 2$ ,  $\mathcal{A}$  is a symmetric matrix so  $\lambda$  will be the largest singular value of  $\mathcal{A}$  and  $\mu$  will be the largest eigenvalue in magnitude of  $\mathcal{A}$ , which implies that  $\lambda = |\mu|$ . Therefore, the best symmetric rank-1 approximation to  $\mathcal{A}$  is its best rank-1 approximation when  $m = 2$ . For the case that  $m \geq 3$ , there is a remarkable difference between matrices and tensors. In [11], the author gave a counterexample, Example 3, to show that the best symmetric rank-1 approximation to a symmetric tensor is not necessarily its best rank-1 approximation. However, tensor  $\mathcal{A} = xyz + zxy + yzx$  related in Example 3 is not symmetric. This assertion had been cited in [21], etc. Thus, it was still unknown if the best symmetric rank-1 approximation to a symmetric tensor is its best rank-1 approximation or not.

Recently, in [21], the authors showed that when  $m = 4$  and  $n = 2$ , the best symmetric rank-1 approximation to a symmetric tensor is its best rank-1 approximation. In [36], it was shown that the conclusion is also true for  $m = 3$  and any  $n$ .

More recently, in [27], the best rank-1 approximation ratio of a tensor space was introduced. A positive lower bound of the best rank-1 approximation ratio of a tensor space gives a convergence rate for the greedy rank-1 update algorithm. A conjecture (Conjecture 1) was given in [27] that the best symmetric rank-1 approximation to a symmetric tensor is its best rank-1 approximation for  $m \geq 4$  and any  $n$ . If this conjecture is true, then a positive lower bound can be given for the best rank-1 approximation ratio of a symmetric tensor space of order  $m$  for  $m \geq 2$ .

In this paper, we will prove that this conjecture is true for general cases.

On the other hand, the spherical homogeneous polynomial optimization problem

$$(3) \quad \mu_1 := \min_{x \in \mathbb{R}^n} \mathcal{A}x^m \\ \text{s.t.} \quad \|x\| = 1$$

is a fundamental model in optimization, closely related to problem (2), and  $\mu_1$  is the smallest  $Z$ -eigenvalue of tensor  $\mathcal{A}$ ; see [1, 26, 29]. As such, it is also widely used in practice, for example, in signal and image processing, investment science, and material sciences; see [3, 16, 22, 32]. The polynomial optimization problem (3) is NP-hard when  $m \geq 3$ ; see [5, 15, 20, 36]. Some polynomial time approximation methods for solving it were proposed; see [5, 15, 31, 36] for details. Therefore, the search for efficient algorithms for the polynomial optimization problem has been a priority for many mathematical programming researchers. Many solution methods based on nonlinear programming and global optimization have been studied and tested; see, e.g., [25, 28, 29]. The power algorithm is one of the important methods and was successfully extended to compute the best rank-1 approximations to higher-order tensors; see, e.g., [6, 7, 12]. Another different approach based upon the so-called sum of squares (SOS) was also proposed by Lasserre [9, 10], Nie [23], and Parrilo [24]. For more details, we refer to the excellent survey by Laurent [13].

Our paper is organized as follows. In section 2, we prove that the problem (1) has always a global optimal solution  $x^{(1)} = \dots = x^{(m)}$  by induction. That is, the best symmetric rank-1 approximation of symmetric tensor is always its best rank-1 approximation. We also give a positive lower bound for the best rank-1 approximation ratio of a symmetric tensor space of order  $m$  for  $m \geq 2$ . In section 3, by applying the obtained result, (3) is reformulated equivalently as a multilinear optimization problem over unit spheres. Based on this reformulation, we propose a modified power method for solving (3). Some numerical results are presented in section 4.

Some words about the notation. Throughout this paper, we denote the space of symmetric  $m$ -order  $n$ -dimensional tensors by  $\text{Sym}^m(\mathfrak{R}^n)$  and

$$\mathcal{A} = (a_{i_1 \dots i_m})_{1 \leq i_1, \dots, i_m \leq n} \in \text{Sym}^m(\mathfrak{R}^n)$$

is a nonzero tensor. And for any positive integer  $k$  with  $k \leq m$  and any  $x^{(s)} \in \mathfrak{R}^n$ ,  $s = 1, \dots, k$ , we denote  $\mathcal{A}x^{(1)} \dots x^{(k)} \in \text{Sym}^{(m-k)}(\mathfrak{R}^n)$  whose  $(i_{k+1}, \dots, i_m)$ th entry is

$$\left( \mathcal{A}x^{(1)} \dots x^{(k)} \right)_{i_{k+1} \dots i_m} = \sum_{1 \leq i_1, \dots, i_k \leq n} a_{i_1 \dots i_k i_{k+1} \dots i_m} x_{i_1}^{(1)} \dots x_{i_k}^{(k)}.$$

Specially, if  $x^{(1)} = \dots = x^{(k)} = x$ , we denote  $\mathcal{A}x^k = \underbrace{\mathcal{A}x \dots x}_k$ .

**2. The best rank-1 approximation of a symmetric tensor.** In this section, we consider the relationship between the best rank-1 approximation and the best symmetric rank-1 approximation of a symmetric tensor. We show that the best symmetric rank-1 approximation to a symmetric tensor is its best rank-1 approximation, which can be stated as follows.

**THEOREM 2.1.** *Suppose that  $\mathcal{A} \in \text{Sym}^m(\mathfrak{R}^n)$ . Then the problem (1) has always a global optimal solution satisfying  $x^{(1)} = \dots = x^{(m)}$ , i.e., the best symmetric rank-1 approximation to  $\mathcal{A}$  is its best rank-1 approximation. This implies that the optimal objective function values of (1) and (2) are the same.*

This theorem shows that Conjecture 1 of [27] is true.

We prove this theorem by induction on  $m$ . For the case that  $m = 2$ , Theorem 2.1 is true from the well-known Eckart–Young theorem. For the case that  $m = 3$ , Theorem 2.1 holds; see [36]. In subsection 2.1, we show that if  $m$  is even and such a conclusion holds for all  $l$  with  $2 \leq l \leq m - 1$ , then such a conclusion holds for  $m$ . For the case that  $m$  is odd, we in subsection 2.2 show that if the conclusion is true for all  $l$  with  $2 \leq l \leq m - 1$ , then such a conclusion is also true for  $m$ . Combining these results together, we have Theorem 2.1.

**2.1. Proof when  $m$  is even.** In this subsection, we discuss the result for which the related symmetric tensor  $\mathcal{A}$  is of even order.

**THEOREM 2.2.** *Suppose that  $m$  is even. Assume that Theorem 2.1 is true for all  $l$ -order symmetric tensor  $\mathcal{A}$  with  $2 \leq l \leq m - 1$ . Then, Theorem 2.1 is also true for  $m$ -order symmetric tensor  $\mathcal{A}$ .*

*Proof.* Let  $m = 2k$ . Assume that  $(x^{(1)}, x^{(2)}, \dots, x^{(m)})$  is an optimal solution of (1) with optimal value  $\lambda > 0$ . Let  $\mathcal{B} = \mathcal{A}x^{(1)}x^{(2)} \dots x^{(k)}$ . Then  $\mathcal{B} \in \text{Sym}^{(m-k)}(\mathfrak{R}^n)$ . It is easy to see that  $(x^{(k+1)}, x^{(k+2)}, \dots, x^{(m)})$  is an optimal solution of

$$(4) \quad \begin{aligned} \max \quad & |\mathcal{B}x^{(k+1)}x^{(k+2)} \dots x^{(m)}| \\ \text{s.t.} \quad & \|x^{(i)}\| = 1, \quad \text{for } i = k+1, k+2, \dots, m, \end{aligned}$$

whose optimal value is still  $\lambda$ . From the assumption, we know that there exists a unit vector  $y \in \mathbb{R}^n$  such that  $\underbrace{(y, y, \dots, y)}_k$  is an optimal solution of (4). That is, we have  $\lambda = |\mathcal{B}y^k|$ .

For the fixed  $y$ , let  $\mathcal{C} = \mathcal{A}y^k$ . Then there exists an optimal solution  $\underbrace{(x, x, \dots, x)}_k$  of

$$(5) \quad \begin{aligned} \max \quad & |\mathcal{C}x^{(1)}x^{(2)} \dots x^{(k)}| \\ \text{s.t.} \quad & \|x^{(i)}\| = 1, \quad \text{for } i = 1, 2, \dots, k. \end{aligned}$$

It is easy to see that  $|(\mathcal{A}y^k)x^k| = \lambda$ , which means that  $\underbrace{(x, x, \dots, x)}_k, \underbrace{(y, y, \dots, y)}_k$  is also an optimal solution of (1). If  $x = -y$ , then  $\underbrace{(x, x, \dots, x)}_m$  is an optimal solution of (1), and the assertion holds. We now assume  $x \neq -y$  and discuss the conclusion in two cases.

*Case 1.*  $\mathcal{A}x^k y^k > 0$ . Then  $\mathcal{A}x^k y^k = \lambda$ . Moreover, by the optimality condition of (1), we have

$$\begin{cases} \mathcal{A}x^{k-1}y^k = \lambda x, \\ \mathcal{A}x^k y^{k-1} = \lambda y, \end{cases}$$

which implies that

$$\begin{cases} M_0 x = \lambda y, \\ M_0 y = \lambda x, \end{cases}$$

where  $M_0 = \mathcal{A}x^{k-1}y^{k-1}$ . By this we know that  $x + y \neq 0$  is an eigenvector of  $M_0$ , associated with the eigenvalue  $\lambda$ . Hence, it holds that  $(x + y)^T M_0 (x + y) / \|x + y\|^2 = \lambda$ , which implies  $\mathcal{A}x^{k-1}y^{k-1}(\frac{x+y}{\|x+y\|})^2 = \lambda$ . Consequently, we know that  $(\frac{x+y}{\|x+y\|}, \underbrace{x, \dots, x}_{k-1}, \frac{x+y}{\|x+y\|}, \underbrace{y, \dots, y}_{k-1})$  is also an optimal solution of (1). Furthermore,

by letting  $M_1 = \mathcal{A}x^{k-2}y^{k-2}(\frac{x+y}{\|x+y\|})^2$  and considering the optimality condition of (1) again, we know that  $x + y$  is also an eigenvector of  $M_1$  associated with the eigenvalue  $\lambda$ . Hence, we have  $(x + y)^T M_1 (x + y) / \|x + y\|^2 = \lambda$ , which means that

$$\mathcal{A}x^{k-2}y^{k-2} \left( \frac{x + y}{\|x + y\|} \right)^4 = \lambda.$$

By repeating this procedure, we have  $\mathcal{A}(\frac{x+y}{\|x+y\|})^m = \lambda$  and know that the assertion holds.

*Case 2.*  $\mathcal{A}x^k y^k < 0$ . Then  $\mathcal{A}x^k y^k = -\lambda$ . By considering the optimality condition of (1), it holds that

$$\begin{cases} \mathcal{A}x^{k-1}y^k = -\lambda x, \\ \mathcal{A}x^k y^{k-1} = -\lambda y, \end{cases}$$

which can be rewritten as

$$\begin{cases} M_0 x = -\lambda y, \\ M_0 y = -\lambda x, \end{cases}$$

where  $M_0 = \mathcal{A}x^{k-1}y^{k-1}$ . Moreover, by a similar way to that used above, we can show that  $\mathcal{A}(\frac{x+y}{\|x+y\|})^m = -\lambda$  and know that the assertion holds.

By combining the two cases above, we complete the proof.  $\square$

**2.2. Proof when  $m$  is odd.** In this subsection, we discuss the case that  $\mathcal{A}$  is an odd-order symmetric tensor. The procedure of the proof will be different from that of Theorem 2.2. In [36], it was shown that Theorem 2.1 holds when  $m = 3$ . In this subsection, we generalize the proof for  $m = 3$  to the general odd-order case.

Now we are ready to propose an algorithm to obtain an optimal solution of (2) from an optimal solution of (1) for general odd  $m$ . This algorithm is a generalization of Algorithm 2.1 in [36]. For convenience of notation, let  $m = 2k + 1$  with integer  $k > 0$ .

ALGORITHM 2.1.

- *Initial Step: Input symmetric tensor  $\mathcal{A}$  and an optimal solution*

$$\underbrace{(x^{(0)}, \dots, x^{(0)})}_k, \underbrace{(z^{(0)}, \dots, z^{(0)})}_{k+1}$$

*of (1) with  $(x^{(0)})^\top z^{(0)} \geq 0$ . Let  $p = 0$  and  $q_p = k$ .*

- *Repeat Step: If  $x^{(p)} = z^{(p)}$ , stop; Otherwise, let*

$$\begin{cases} x^{(p+1)} = \frac{x^{(p)} + z^{(p)}}{\|x^{(p)} + z^{(p)}\|}, & z^{(p+1)} = z^{(p)}, & \text{if } 2q_p < (m - 2q_p), \\ x^{(p+1)} = z^{(p)}, & z^{(p+1)} = \frac{x^{(p)} + z^{(p)}}{\|x^{(p)} + z^{(p)}\|}, & \text{otherwise.} \end{cases}$$

*Let  $q_{p+1} := \min\{2q_p, m - 2q_p\}$  and  $p := p + 1$ .*

*Remark 2.1.* From the proof of Theorem 2.2,  $\underbrace{(x^{(p)}, \dots, x^{(p)})}_{q_p}, \underbrace{(z^{(p)}, \dots, z^{(p)})}_{m-q_p}$  is

always an optimal solution of problem (1) for all  $p$ . Furthermore, from the above procedure together with the parallelogram law, it holds that

$$\widehat{(x^{(p+1)}, z^{(p+1)})} = \frac{1}{2} \widehat{(x^{(p)}, z^{(p)})},$$

where  $\widehat{(a, b)}$  denotes the angle between vectors  $a$  and  $b$ .

Now we are ready to show Theorem 2.1 with a tensor of odd order.

**THEOREM 2.3.** *Suppose that  $m = 2k + 1$  and Theorem 2.1 holds for all  $l$ -order symmetric tensors  $\mathcal{A}$  with  $2 \leq l \leq m - 1$ . Then there exists  $x^* \in \mathbb{R}^n$  such that  $\underbrace{(x^*, \dots, x^*)}_m$  is an optimal solution of (1).*

*Proof.* We prove the result in two cases. If Algorithm 2.1 terminates in finitely many steps, then there exists  $p$  such that  $x^{(p)} = z^{(p)}$ . Since  $\underbrace{(x^{(p)}, \dots, x^{(p)})}_{q_p}, \underbrace{(z^{(p)}, \dots, z^{(p)})}_{m-q_p}$

is always an optimal solution of (1) from Remark 2.1, we know that the conclusion holds by letting  $x^* = x^{(p)}$ .

Suppose that  $\{(x^{(p)}, z^{(p)})\}$  is the infinite sequence generated by Algorithm 2.1. Let  $(x^*, z^*)$  be an accumulation point of  $\{(x^{(p)}, z^{(p)})\}_{p=1}^\infty$ . As a consequence of Remark 2.1,  $x^* = z^*$ . Hence, we assert that  $\underbrace{(x^*, \dots, x^*)}_m$  is an optimal solution of (1). So we have

the desired result and complete the proof.  $\square$

**2.3. An application: A positive lower bound for the best rank-1 approximation ratio of a symmetric tensor space.** As in [27], for  $\mathcal{A} = (a_{i_1 \dots i_m}) \in \text{Sym}^m(\mathbb{R}^n)$ ,  $\|\mathcal{A}\|$  is defined by

$$\|\mathcal{A}\| = \sqrt{\sum_{i_1, \dots, i_m=1}^n a_{i_1 \dots i_m}^2}.$$

The best rank-1 approximation ratio of  $\mathcal{T}^m(\mathfrak{R}^n)$  is defined as

$$(6) \quad App(\text{Sym}^m(\mathfrak{R}^n)) = \max \left\{ \sigma : \sigma \leq \frac{\mu(\mathcal{A})}{\|\mathcal{A}\|} \forall \mathcal{A} \in \text{Sym}^m(\mathfrak{R}^n), \mathcal{A} \neq \mathcal{O} \right\}.$$

As stated in the introduction, a positive lower bound of  $App(\text{Sym}^m(\mathfrak{R}^n))$  gives a convergence rate for the greedy rank-1 update algorithm in  $\text{Sym}^m(\mathfrak{R}^n)$ . For more discussion on this, see [27].

By Theorem 2.1 of this paper and Theorem 3.1, (2.4), and (2.5) of [27], we have the following theorem.

THEOREM 2.4. *For any integer  $m \geq 2$ , we have*

$$App(\text{Sym}^m(\mathfrak{R}^n)) \geq \frac{1}{n^{m-1}}.$$

This shows that Conjecture 2 of [27] is also true.

**3. The related polynomial optimization problem over the unit sphere.**

In this section, we apply Theorem 2.1 to the homogeneous polynomial optimization problem, which is closely related with (2) and has the following form:

$$(7) \quad \begin{aligned} f_{\min} &:= \min_{x \in \mathfrak{R}^n} f(x) := \mathcal{A}x^m \\ \text{s.t.} \quad &\|x\| = 1, \end{aligned}$$

where  $0 \neq \mathcal{A} \in \text{Sym}^m(\mathfrak{R}^n)$ . It is clear that if  $m$  is odd, then  $|f_{\min}| = |f_{\max}| = \max_{\|x\|=1} |f(x)|$ , where  $f_{\max}$  is the maximum value of (7). Therefore, by Theorem 2.1, it is easy to see that when  $m = 2k + 1$ , the problem (7) and the following multilinear optimization problem

$$(8) \quad \begin{aligned} \min_{x^{(1)}, \dots, x^{(m)} \in \mathfrak{R}^n} \quad &g(x^{(1)}, \dots, x^{(m)}) := \mathcal{A}x^{(1)} \dots x^{(m)} \\ \text{s.t.} \quad &\|x^{(i)}\| = 1 \text{ for } i = 1, 2, \dots, m \end{aligned}$$

have the same optimal value and  $f_{\min} \leq 0$ .

However, the equivalence described above does not hold when  $m = 2k$ . For example, if the tensor  $\mathcal{A}$  in (7) is positive semidefinite (see [26]) and  $\mathcal{A} \neq 0$ , then  $f_{\min} \geq 0$  whereas the optimal value of (8) is negative. For this case, we consider

$$(9) \quad \begin{aligned} h_{\min} &= \min_{x \in \mathfrak{R}^n} h(x) := f(x) - \alpha(x^\top x)^k \\ \text{s.t.} \quad &\|x\| = 1, \end{aligned}$$

where  $\alpha$  is a large real number such that  $f(x) \leq \alpha(x^\top x)^k$  for any  $x \in \mathfrak{R}^n$  with  $\|x\| = 1$ . It is clear that  $h_{\min} = f_{\min} - \alpha$  and  $|h_{\min}| = \alpha - f_{\min}$ . Consequently, it holds that

$$(10) \quad |h_{\min}| = \max_{\|x\|=1} (\alpha - f(x)) = \max_{\|x\|=1} |h(x)|.$$

Let  $\mathcal{B}$  be the  $2k$ -order  $n$ -dimensional identity tensor such that  $\mathcal{B}x^m = (x^\top x)^k$ , which appeared first in Property 2.4 of [7]. Therefore, problem (7) can be reformulated as follows

$$\begin{aligned} \min_{x \in \mathfrak{R}^n} \quad &h(x) = (\mathcal{A} - \alpha\mathcal{B})x^m \\ \text{s.t.} \quad &\|x\| = 1 \end{aligned}$$

when  $m$  is even. Consequently, by (10) and Theorem 2.1, we know that if  $m = 2k$ , the optimal value of problem (7) can be obtained by solving the following multilinear optimization problem

$$h_{\min} = \min_{x^{(1)}, \dots, x^{(m)} \in \mathbb{R}^n} (\mathcal{A} - \alpha\mathcal{B})x^{(1)} \dots x^{(m)} \\ \text{s.t.} \quad \|x^{(i)}\| = 1 \text{ for } i = 1, 2, \dots, m.$$

*Remark 3.1.* Here,  $\alpha$  is adopted to ensure the equivalence between above two spherical optimizations so that the problem can be solved via solving the multilinear program. However,  $\alpha$  adopted in [7] is used to force  $\max_{\|x\|=1} (\mathcal{A} + \alpha\mathcal{B})x^m$  concavity and consequently guarantee convergence.

From the discussion above, we have the following theorem.

**THEOREM 3.1.** *The optimal value of the spherical polynomial optimization problem (7) can be obtained by solving a related multilinear spherical optimization problem.*

*Remark 3.2.* How can one choose a suitable  $\alpha$  such that  $|h_{\min}| = \max_{\|x\|=1} |h(x)|$ ? In fact, this can be obtained based on the estimation of the largest value of  $f(x)$  over the unit sphere. It is clear that  $f(x) \leq \sum_{1 \leq i_1, \dots, i_m \leq n} |\mathcal{A}_{i_1 i_2 \dots i_m}|$  for any unit vector  $x$ . Therefore, if we take  $\alpha = \sum_{1 \leq i_1, \dots, i_m \leq n} |\mathcal{A}_{i_1 i_2 \dots i_m}|$ , then Theorem 3.1 holds.

In the rest of this section, we continue to study how to solve the homogeneous polynomial optimization problem (7) based upon the multilinear program (8). As mentioned above, if  $m = 2k$ , we can replace  $\mathcal{A}$  by  $\mathcal{A} - \alpha\mathcal{B}$  with a suitable positive number  $\alpha$ . Therefore, without loss of generality, we assume that (7) and multilinear program (8) have the same optimal value. In spite of this, they do not have the same optimal solution set. Motivated by this, we consider the following optimization problem:

$$(11) \quad q_{\min} = \min_{x^{(1)}, \dots, x^{(m)} \in \mathbb{R}^n} q(x^{(1)}, \dots, x^{(m)}) := \mathcal{A}x^{(1)} \dots x^{(m)} - \sum_{i=2}^m (x^{(1)})^\top x^{(i)} \\ \text{s.t.} \quad \|x^{(i)}\| = 1 \text{ for } i = 1, 2, \dots, m.$$

For the relationship between (7) and (11), we have the following result.

**PROPOSITION 3.1.** *It holds that  $q_{\min} = f_{\min} - (m - 1)$  and the optimal value  $q_{\min}$  of (11) is attained at  $x^{(1)} = \dots = x^{(m)}$ .*

*Proof.* It is easy to see that

$$q_{\min} \geq \min_{\|x^{(1)}\|=1, \dots, \|x^{(m)}\|=1} \mathcal{A}x^{(1)} \dots x^{(m)} - (m - 1) \\ = \min_{\|x\|=1} \mathcal{A}x^m - (m - 1) \\ = f_{\min} - (m - 1).$$

On the other hand, since any  $\underbrace{(x, x, \dots, x)}_m \in \mathbb{R}^n \times \dots \times \mathbb{R}^n$  with  $\|x\| = 1$  is feasible for (11), it holds that

$$q_{\min} \leq \min_{\|x\|=1} [\mathcal{A}x^m - (m - 1)x^\top x] = f_{\min} - (m - 1).$$

Hence, the desired conclusion holds.  $\square$

To establish the monotone convergence in the sense that the objective function value of the iterative sequence are monotone when the iterative points are different, the global line search approach has been developed and used recently; see [14, 30, 34]. Now we are ready to propose a modified power method for (7) by solving the optimization problem (11). Furthermore, as presented in Theorem 3.2, the proposed algorithm is also monotonic convergence.

ALGORITHM 3.1.

- Step 0: Let  $(x^{(1,0)}, x^{(2,0)}, \dots, x^{(m,0)}) \in \mathbb{R}^n \times \dots \times \mathbb{R}^n$  with  $\|x^{(i,0)}\| = 1$  for  $i = 1, \dots, m$  be an initial point such that  $q(x^{(1,0)}, x^{(2,0)}, \dots, x^{(m,0)}) < 0$ . Let  $l = 0$ .
- Step 1: Let

$$x^{(1,l+1)} = - \frac{\mathcal{A}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m x^{(i,l)}}{\left\| \mathcal{A}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m x^{(i,l)} \right\|}.$$

For  $i = 2, \dots, m$ , let

$$x^{(i,l+1)} = - \frac{\mathcal{A}x^{(1,l+1)} \dots x^{(i-1,l+1)} x^{(i+1,l)} \dots x^{(m,l)} - x^{(1,l+1)}}{\left\| \mathcal{A}x^{(1,l+1)} \dots x^{(i-1,l+1)} x^{(i+1,l)} \dots x^{(m,l)} - x^{(1,l+1)} \right\|}$$

if  $\mathcal{A}x^{(1,l+1)} \dots x^{(i-1,l+1)} x^{(i+1,l)} \dots x^{(m,l)} - x^{(1,l+1)} \neq 0$  and  $x^{(i,l+1)} = x^{(i,l)}$  otherwise.

- Step 2: If a certain convergence criterion is satisfied for  $(x^{(1,l+1)}, \dots, x^{(m,l+1)})$ , then stop. Otherwise, let  $l := l + 1$ , and go to Step 1.

Remark 3.3. From  $q(x^{(1,0)}, \dots, x^{(m,0)}) < 0$  and the decreasing of  $\{q(x^{(1,l)}, \dots, x^{(m,l)})\}$  which will be seen in the proof of the next theorem, it holds that

$$\left( \mathcal{A}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m x^{(i,l)} \right)^\top x^{(1,l)} < 0$$

for any  $l = 1, 2, \dots$ . Hence,  $\mathcal{A}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m x^{(i,l)} \neq 0$ , which indicates that Algorithm 3.1 is well defined.

It is clear that the sequence  $\{(x^{(1,l)}, x^{(2,l)}, \dots, x^{(m,l)})\}_{l=1}^\infty$  generated by Algorithm 3.1 is bounded, hence there exists an accumulation point. Now we are ready to state and prove the convergence of Algorithm 3.1, which shows that the whole generated sequence converges to a KKT point of (11) under some conditions. To this end, we need the following proposition which has already been shown by Moré and Sorensen [19].

PROPOSITION 3.2. Assume that  $w^* \in \mathbb{R}^s$  is an isolated accumulation point of a sequence  $\{w^{(k)}\} \subset \mathbb{R}^s$  such that for every subsequence  $\{w^{(k)}\}_K$  converging to  $w^*$  there is an infinite subset  $\bar{K} \subseteq K$  such that  $\{\|w^{(k+1)} - w^{(k)}\|\}_{\bar{K}} \rightarrow 0$ . Then the whole sequence  $\{w^{(k)}\}$  converges to  $w^*$ .

THEOREM 3.2. Let  $\{(x^{(1,l)}, x^{(2,l)}, \dots, x^{(m,l)})\}_{l=1}^\infty$  be a sequence generated by Algorithm 3.1. Then the sequence  $\{q(x^{(1,l)}, x^{(2,l)}, \dots, x^{(m,l)})\}_{l=1}^\infty$  is monotone convergence. Furthermore, suppose that the sequence  $\{(x^{(1,l)}, x^{(2,l)}, \dots, x^{(m,l)})\}_{l=1}^\infty$  has an isolated accumulation point. Then the generated sequence converges to a KKT point of (11).

Proof. For convenience of notation, let

$$q_l := q(x^{(1,l)}, \dots, x^{(m,l)}) = \mathcal{A}x^{(1,l)} \dots x^{(m,l)} - \sum_{i=2}^m (x^{(1,l)})^\top x^{(i,l)}.$$



Then there holds

$$\begin{aligned}
 q_l &= (x^{(1,l)})^\top \left( \mathcal{A}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m x^{(i,l)} \right) \\
 &\geq - \left\| \mathcal{A}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m x^{(i,l)} \right\| \\
 (12) \quad &= (x^{(1,l+1)})^\top \left( \mathcal{A}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m x^{(i,l)} \right) \\
 &= \mathcal{A}x^{(1,l+1)}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m \left( x^{(1,l+1)} \right)^\top x^{(i,l)} \\
 &= q(x^{(1,l+1)}, x^{(2,l)}, \dots, x^{(m,l)}).
 \end{aligned}$$

Furthermore, for  $i = 2, \dots, m$ , we consider two cases:

- (i)  $\mathcal{A}x^{(1,l+1)} \dots x^{(i-1,l+1)}x^{(i+1,l)} \dots x^{(m,l)} - x^{(1,l+1)} \neq 0$ ,
- (ii)  $\mathcal{A}x^{(1,l+1)} \dots x^{(i-1,l+1)}x^{(i+1,l)} \dots x^{(m,l)} - x^{(1,l+1)} = 0$ .

For case (i), it holds that

$$\begin{aligned}
 &\mathcal{A}x^{(1,l+1)} \dots x^{(i-1,l+1)}x^{(i,l+1)}x^{(i+1,l)} \dots x^{(m,l)} - (x^{(1,l+1)})^\top x^{(i,l+1)} \\
 &= - \left\| \mathcal{A}x^{(1,l+1)} \dots x^{(i-1,l+1)}x^{(i+1,l)} \dots x^{(m,l)} - x^{(1,l+1)} \right\| \\
 &\leq (x^{(i,l)})^\top \left( \mathcal{A}x^{(1,l+1)} \dots x^{(i-1,l+1)}x^{(i+1,l)} \dots x^{(m,l)} - x^{(1,l+1)} \right) \\
 &= \mathcal{A}x^{(1,l+1)} \dots x^{(i-1,l+1)}x^{(i,l)}x^{(i+1,l)} \dots x^{(m,l)} - (x^{(1,l+1)})^\top x^{(i,l)}.
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 &q(x^{(1,l+1)}, \dots, x^{(i-1,l+1)}, x^{(i,l+1)}, x^{(i+1,l)}, \dots, x^{(m,l)}) \\
 &= \mathcal{A}x^{(1,l+1)} \dots x^{(i-1,l+1)}x^{(i,l+1)}x^{(i+1,l)} \dots x^{(m,l)} \\
 &\quad - \sum_{j=2}^{i-1} \left( x^{(1,l+1)} \right)^\top \left( x^{(j,l+1)} \right) - \left( x^{(1,l+1)} \right)^\top x^{(i,l+1)} - \sum_{j=i+1}^m \left( x^{(1,l+1)} \right)^\top \left( x^{(j,l)} \right) \\
 &\leq \mathcal{A}x^{(1,l+1)} \dots x^{(i-1,l+1)}x^{(i,l)}x^{(i+1,l)} \dots x^{(m,l)} \\
 &\quad - \left( x^{(1,l+1)} \right)^\top x^{(i,l)} - \sum_{j=2}^{i-1} \left( x^{(1,l+1)} \right)^\top \left( x^{(j,l+1)} \right) - \sum_{j=i+1}^m \left( x^{(1,l+1)} \right)^\top \left( x^{(j,l)} \right) \\
 &= q(x^{(1,l+1)}, \dots, x^{(i-1,l+1)}, x^{(i,l)}, x^{(i+1,l)}, \dots, x^{(m,l)}),
 \end{aligned}$$

which indicates, together with (12), that for any  $l$  and  $i$ ,

$$\begin{aligned}
 (13) \quad &q_{l+1} \leq q(x^{(1,l+1)}, \dots, x^{(i-1,l+1)}, x^{(i,l+1)}, x^{(i+1,l)}, \dots, x^{(m,l)}) \\
 &\leq q(x^{(1,l+1)}, \dots, x^{(i-1,l+1)}, x^{(i,l)}, x^{(i+1,l)}, \dots, x^{(m,l)}) \leq q_l.
 \end{aligned}$$

For case (ii), it is clear that (13) holds, since  $x^{(i,l+1)} = x^{(i,l)}$ .

So we can assert that  $\{q_l\}$  is decreasing, which implies, together with the fact that  $\{q_l\}$  is bounded, that  $\{q_l\}$  converges. The monotone convergence of  $\{q_l\}$  is proved.

Without loss of generality, we assume that  $\lim_{l \rightarrow \infty} q_l = q_*$ .

Since  $0 \geq q_l \geq - \left\| \mathcal{A}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m x^{(i,l)} \right\| \geq q_{l+1}$ , we have

$$\lim_{l \rightarrow \infty} \left\| \mathcal{A}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m x^{(i,l)} \right\| = -q_*.$$

Furthermore, it holds that

$$\begin{aligned} \left(x^{(1,l)}\right)^\top \left(x^{(1,l+1)}\right) &= -\frac{\mathcal{A}x^{(1,l)}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m \left(x^{(1,l)}\right)^\top \left(x^{(i,l)}\right)}{\left\|\mathcal{A}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m x^{(i,l)}\right\|} \\ &= \frac{q_l}{-\left\|\mathcal{A}x^{(2,l)} \dots x^{(m,l)} - \sum_{i=2}^m x^{(i,l)}\right\|} \rightarrow 1 \end{aligned}$$

when  $l \rightarrow \infty$ , which means that  $\lim_{l \rightarrow \infty} \|x^{(1,l)} - x^{(1,l+1)}\| \rightarrow 0$ .

Let  $(x^{(1,*)}, \dots, x^{(m,*)})$  be an isolated accumulate point of  $\{(x^{(1,l)}, \dots, x^{(m,l)})\}_{l=1}^\infty$ . For every subsequence  $\{(x^{(1,l)})\}_{l \in K}$  converging to  $x^{(1,*)}$ , the condition in Proposition 3.2 holds. Hence, by Proposition 3.2 we know that the whole sequence  $\{x^{(1,l)}\}_{l=1}^\infty$  converges to  $x^{(1,*)}$  as  $l \rightarrow \infty$ .

Now we are ready to show the convergence of  $\{x^{(2,l)}\}_{l=1}^\infty$ . Without loss of generality, we assume that  $x^{(2,l+1)} \neq x^{(2,l)}$ . By Step 1 in Algorithm 3.1, we have

$$\begin{aligned} &\mathcal{A}x^{(1,l+1)}x^{(2,l+1)}x^{(3,l)} \dots x^{(m,l)} \\ &= (x^{(2,l+1)})^\top x^{(1,l+1)} - \left\|\mathcal{A}x^{(1,l+1)}x^{(3,l)} \dots x^{(m,l)} - x^{(1,l+1)}\right\|, \end{aligned}$$

which implies that

$$\begin{aligned} &q\left(x^{(1,l+1)}, x^{(2,l+1)}, x^{(3,l)}, \dots, x^{(m,l)}\right) \\ &= -\left\|\mathcal{A}x^{(1,l+1)}x^{(3,l)} \dots x^{(m,l)} - x^{(1,l+1)}\right\| - \sum_{i=3}^m \left(x^{(1,l+1)}\right)^\top x^{(i,l)}. \end{aligned}$$

Consequently, it holds that

$$\begin{aligned} (14) \quad &\left(x^{(2,l+1)}\right)^\top x^{(2,l)} = -\frac{\mathcal{A}x^{(1,l+1)}x^{(2,l)}x^{(3,l)} \dots x^{(m,l)} - \left(x^{(1,l+1)}\right)^\top x^{(2,l)}}{\left\|\mathcal{A}x^{(1,l+1)}x^{(3,l)} \dots x^{(m,l)} - x^{(1,l+1)}\right\|} \\ &= -\frac{q\left(x^{(1,l+1)}, x^{(2,l)}, x^{(3,l)}, \dots, x^{(m,l)}\right) + \sum_{i=3}^m \left(x^{(1,l+1)}\right)^\top x^{(i,l)}}{\left\|\mathcal{A}x^{(1,l+1)}x^{(3,l)} \dots x^{(m,l)} - x^{(1,l+1)}\right\|} \\ &= \frac{q\left(x^{(1,l+1)}, x^{(2,l)}, x^{(3,l)}, \dots, x^{(m,l)}\right) + \sum_{i=3}^m \left(x^{(1,l+1)}\right)^\top x^{(i,l)}}{q\left(x^{(1,l+1)}, x^{(2,l+1)}, x^{(3,l)}, \dots, x^{(m,l)}\right) + \sum_{i=3}^m \left(x^{(1,l+1)}\right)^\top x^{(i,l)}}. \end{aligned}$$

For every subsequence  $\{x^{(2,l)}\}_{l \in K}$  converging to  $x^{(2,*)}$ , there exists an infinite subset  $\bar{K} \subseteq K$  such that  $\{x^{(i,l)}\}_{l \in \bar{K}}$  converges for every  $i = 3, \dots, m$ , since  $\{x^{(i,l)}\}_{l \in K}$  is bounded. Furthermore, by (13) it holds that

$$\lim_{l \rightarrow \infty} q\left(x^{(1,l+1)}, x^{(2,l)}, \dots, x^{(m,l)}\right) = \lim_{l \rightarrow \infty} q\left(x^{(1,l+1)}, x^{(2,l+1)}, x^{(3,l)}, \dots, x^{(m,l)}\right) = q_*.$$

Thus, by (14) we know

$$(15) \quad \left\{ \left(x^{(2,l+1)}\right)^\top x^{(2,l)} \right\}_{l \in \bar{K}} \rightarrow 1,$$

and hence  $\{\|x^{(2,l+1)} - x^{(2,l)}\|\}_{l \in \bar{K}} \rightarrow 0$ . Consequently, by Proposition 3.2 we assert that the whole sequence  $\{x^{(2,l)}\}_{l=1}^\infty$  converges to  $x^{(2,*)}$ .

Similarly, we can prove that  $\lim_{l \rightarrow \infty} x^{(i,l)} = x^{(i,*)}$  for every  $i = 3, \dots, m$ . Consequently, by Step 1 in Algorithm 3.1 we have

$$(16) \quad \begin{cases} \mathcal{A}x^{(2,*)} \dots x^{(m,*)} - \sum_{i=2}^m x^{(i,*)} + \lambda_1^* x^{(1,*)} = 0, \\ \mathcal{A}x^{(1,*)} \dots x^{(i-1,*)} x^{(i+1,*)} \dots x^{(m,*)} - x^{(1,*)} + \lambda_i^* x^{(i,*)} = 0 \\ \quad \text{for } i = 2, 3, \dots, m, \\ \|x^{(i,*)}\| = 1 \text{ for } i = 1, 2, \dots, m, \end{cases}$$

where  $\lambda_1^* = \|\mathcal{A}x^{(2,*)} \dots x^{(m,*)} - \sum_{i=2}^m x^{(i,*)}\|$  and

$$\lambda_i^* = \left\| \mathcal{A}x^{(1,*)} \dots x^{(i-1,*)} x^{(i+1,*)} \dots x^{(m,*)} - x^{(1,*)} \right\| \quad \text{for } i = 2, \dots, m.$$

By (16), we assert that  $(x^{(1,*)}, \dots, x^{(m,*)})$  is a KKT point of problem (11).  $\square$

Although the global optimal solution cannot be guaranteed by Algorithm 3.1, some good solution can be obtained if a good initial point is chosen.

**4. Preliminary numerical results.** In this section, we report some numerical results to illustrate the algorithm for solving problem (7) based upon the problem (11). In our numerical experiments, we take

$$\left| \mathcal{A}x^{(1,l+1)} x^{(2,l+1)} \dots x^{(m,l+1)} - \mathcal{A}x^{(1,l)} x^{(2,l)} \dots x^{(m,l)} \right| < 10^{-6}$$

as the stopping criterion for Algorithm 3.1. And for Example 4.1 and Example 4.2, the vector  $(x^{(1,0)}, \dots, x^{(m,0)})$  in Step 0 of Algorithm 3.1 is taken by the scheme HOSVD proposed in [12]. Throughout this section, for cases in which  $m$  is even, we take  $\alpha = \sum_{i_1, i_2, \dots, i_m=1}^n |\mathcal{A}_{i_1 i_2 \dots i_m}|$  in (9) and the symmetric tensor  $\mathcal{B}$  satisfying  $\mathcal{B}x^m = \|x\|^{\frac{m}{2}}$ .

*Example 4.1.* We consider the problem  $\max_{\|x\|=1} \mathcal{A}x^4$ , where  $\mathcal{A}$  is a 4-order 3-dimensional symmetric tensor  $\mathcal{A}$  with entries

$$\begin{aligned} \mathcal{A}_{1111} &= 0.2883, & \mathcal{A}_{1112} &= -0.0031, & \mathcal{A}_{1113} &= 0.1973, & \mathcal{A}_{1122} &= -0.2485, \\ \mathcal{A}_{1123} &= -0.2939, & \mathcal{A}_{1133} &= 0.3847, & \mathcal{A}_{1222} &= 0.2972, & \mathcal{A}_{1223} &= 0.1862, \\ \mathcal{A}_{1233} &= 0.0919, & \mathcal{A}_{1333} &= -0.3619, & \mathcal{A}_{2222} &= 0.1241, & \mathcal{A}_{2223} &= -0.3420, \\ \mathcal{A}_{2233} &= 0.2127, & \mathcal{A}_{2333} &= 0.2727, & \mathcal{A}_{3333} &= -0.3054. \end{aligned}$$

This example comes from [6], and its optimal value  $f_{\max} = 0.8893$  was obtained in [7]. Let  $\mathcal{C} = \mathcal{A} + \alpha\mathcal{B}$  with  $\alpha = 18.5540$ . Then by Theorem 3.1, the original problem can be converted into the following multilinear optimization problem

$$\begin{aligned} h_{\max} &= \max_{x^{(1)}, \dots, x^{(4)} \in \mathbb{R}^3} \mathcal{C}x^{(1)} \dots x^{(4)} \\ \text{s.t.} & \quad \|x^{(i)}\| = 1 \text{ for } i = 1, \dots, 4. \end{aligned}$$

Since  $f_{\max} = 0.8893$ , y to know that  $h_{\max} = 0.8893 + \alpha = 19.4433$ . We apply HOPM proposed in [6] to the multilinear optimization problems with  $\mathcal{A}$  and  $\mathcal{C}$ , respectively.

We apply SHOPM proposed in [6] to the original problem and  $\max_{\|x\|=1} \mathcal{C}x^4$ , respectively. Notice that SHOPM for  $\max_{\|x\|=1} \mathcal{C}x^4$  is equivalent to the SSHOPM in [7].

We also apply Algorithm 3.1 to the following optimization problems:

$$\begin{aligned} \max_{x^{(1)}, \dots, x^{(4)} \in \mathbb{R}^3} & \quad \mathcal{A}x^{(1)} \dots x^{(4)} + \sum_{i=2}^4 (x^{(1)})^\top x^{(i)} \\ \text{s.t.} & \quad \|x^{(i)}\| = 1 \text{ for } i = 1, \dots, 4 \end{aligned}$$

and

$$\begin{aligned} \max_{x^{(1)}, \dots, x^{(4)} \in \mathbb{R}^3} & \quad \mathcal{C}x^{(1)} \dots x^{(4)} + \sum_{i=2}^4 (x^{(1)})^\top x^{(i)} \\ \text{s.t.} & \quad \|x^{(i)}\| = 1 \text{ for } i = 1, \dots, 4. \end{aligned}$$

The numerical results are presented in Figure 4.1 and Figure 4.2, respectively. The curves in Figure 4.1 depict the values of the function  $\mathcal{A}x^{(1)} \dots x^{(4)}$  at every iterative point generated by HOPM, SHOPM, and Algorithm 3.1, and the curves in Figure 4.2 depict the values of the functions  $\mathcal{A}x^{(1)} \dots x^{(4)}$  at every iterative point generated by applying three algorithms to solve the optimization problem with  $\mathcal{C}$ .

From Figure 4.1, we see that for the optimization problem with  $\mathcal{A}$ , the objective value obtained by HOPM is larger than the objective value obtained by Algorithm 3.1. However  $x^{(1)} = \dots = x^{(4)}$  does not hold for the solution obtained by HOPM. In fact, from Table 3.1 in [7], we can see that the value obtained by HOPM is the absolute value of the minimum value for the original problem, which is  $\max_{\|x\|=1} |\mathcal{A}x^4|$ . Moreover, from these figures, we can see that objective value obtained by Algorithm 3.1 always improves the value obtained by SHOPM within less steps.

By comparing Figure 4.1 with Figure 4.2, we can see that three algorithms are more suitable for solving the optimization problem with tensor  $\mathcal{C}$ , which can be seen from Table 4.1 too, where  $T$  denotes the related tensor,  $(\bar{x}^{(1)}, \bar{x}^{(2)}, \bar{x}^{(3)}, \bar{x}^{(4)})^\top$  denotes the obtained final iteration,  $\bar{f}$  denotes the value  $\mathcal{A}\bar{x}^{(1)} \dots \bar{x}^{(4)}$  of the objective function at the final iteration, and Nit denotes the total number of iterations for solving this problem.

*Example 4.2.* We consider the problem  $\min_{\|x\|=1} \mathcal{A}x^3$ , in which the 3-order 3-dimensional symmetric tensor  $\mathcal{A}$  is defined by

$$\begin{aligned} \mathcal{A}_{111} &= -0.1281, & \mathcal{A}_{112} &= 0.0516, & \mathcal{A}_{113} &= -0.0954, & \mathcal{A}_{122} &= -0.1958, \\ \mathcal{A}_{123} &= -0.1790, & \mathcal{A}_{133} &= -0.2676, & \mathcal{A}_{222} &= 0.3251, & \mathcal{A}_{223} &= 0.2513, \\ \mathcal{A}_{233} &= 0.1773, & \mathcal{A}_{333} &= 0.0338. \end{aligned}$$

From [7], we know that  $f_{\min} = -0.8730$ . We applied SHOPM, HOPM, and Algorithm 3.1 to solve the original problem or the corresponding multilinear optimization problems and ran 500 trials with initial points generated randomly. The test results are listed in Table 4.2, where SR denotes the success ratio in the tested problems, i.e., the occurrences ratio that the optimal value  $f_{\min}$  is arrived, and AIN denotes the average iterations numbers in methods for the successful cases.

From Table 4.2, we see that SR obtained by Algorithm 3.1 is clearly larger than those obtained by the other two algorithms.

*Example 4.3.* Consider the problem  $\min_{\|x\|=1} \mathcal{A}x^3$  with the 3-order  $n$ -dimensional tensor  $\mathcal{A}$  whose entries are uniformly distributed in  $(0, 1)$ .

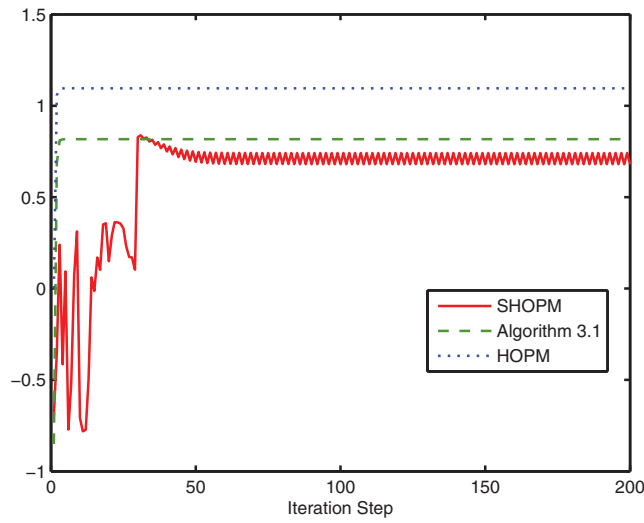


FIG. 4.1. Numerical results of Example 4.1 for tensor  $\mathcal{A}$ .

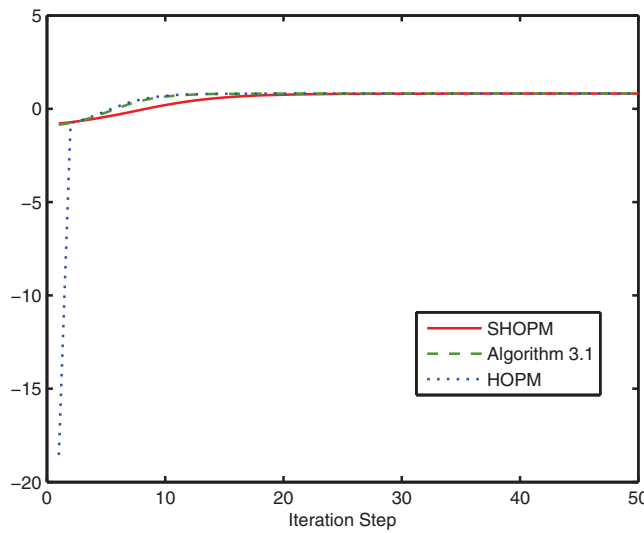


FIG. 4.2. Numerical results of Example 4.1 for tensor  $\mathcal{C}$ .

TABLE 4.1  
The test results of Example 4.1 by SHOPM, HOPH, and Algorithm 3.1.

T	Algorithm	$\bar{f}$	$\bar{x}^{(1)}$	$\bar{x}^{(2)}$	$\bar{x}^{(3)}$	$\bar{x}^{(4)}$	Nit
$\mathcal{A}$	SHOPM	0.7420	(0.8412, 0.2635, 0.4722)	$\bar{x}^{(1)}$	$\bar{x}^{(1)}$	$\bar{x}^{(1)}$	201
	HOPM	1.0954	(0.5915, -0.7467, -0.3044)	$-\bar{x}^{(1)}$	$-\bar{x}^{(1)}$	$-\bar{x}^{(1)}$	12
	Algorithm 3.1	0.8169	(-0.8412, 0.2635, -0.4722)	$\bar{x}^{(1)}$	$\bar{x}^{(1)}$	$\bar{x}^{(1)}$	10
$\mathcal{C}$	SHOPM	0.8893	(0.6672, 0.2472, -0.7027)	$\bar{x}^{(1)}$	$\bar{x}^{(1)}$	$\bar{x}^{(1)}$	201
	HOPM	0.8893	(-0.6672, -0.2472, 0.7027)	$\bar{x}^{(1)}$	$\bar{x}^{(1)}$	$\bar{x}^{(1)}$	97
	Algorithm 3.1	0.8893	(-0.6672, -0.2472, 0.7027)	$\bar{x}^{(1)}$	$\bar{x}^{(1)}$	$\bar{x}^{(1)}$	104

TABLE 4.2  
The test results of Example 4.2.

	Algorithm		
	SHOPM	HOPM	Algorithm 3.1
SR	0.4480	0.5940	0.8060
AIN	11.2723	6.9798	11.9305

TABLE 4.3  
The test results of Example 4.3 by Algorithm 3.1.

DIM	INVAL	LVAL	ALVAL	SR(%)	ACPU
3	-1.8301	-3.1699	-2.7520	100	0.0544
10	-5.6626	-17.9066	-15.9230	100	1.5278
30	-16.030	-87.799	-82.653	100	35.049
100	-51.192	-511.916	-499.885	100	910.111

We test 100 problems for both cases  $n = 3$  and  $n = 10$  and test 10 problems for both cases  $n = 30$  and  $n = 100$ . The numerical results are listed in the following table, where DIM denotes the dimension of the tensor  $\mathcal{A}$ , INVAL, and LVAL denote the average initial value and average lower bound provided by Theorem 7.1 in [36], respectively, ALVAL denotes the average value obtained by Algorithm 3.1, SR denotes the success rate in the tested problems, and ACPU denotes the average CPU time (in seconds) used by the trials. Here, a case is said to be successful if the solution  $(x^{(1)}, x^{(2)}, x^{(3)})^\top$  obtained by Algorithm 3.1 satisfies  $x^{(1)} = x^{(2)} = x^{(3)}$ .

From Table 4.3, we can see that a Z-eigenvalue of the tensor  $\mathcal{A}$  generated randomly in Example 4.3 can always be obtained by Algorithm 3.1, since the obtained solution always satisfies  $x^{(1)} = x^{(2)} = x^{(3)}$  for all test cases.

**Final remark.** Let  $\mathcal{A}$  be an  $m$ th-order  $(n_1 \times n_2 \times \cdots \times n_m)$ -dimensional tensor.  $\mathcal{A}$  can be decomposed as

$$\mathcal{A} = \sum_{i=1}^r \alpha_i x^{(i,1)} \cdots x^{(i,m)},$$

where  $x^{(i,j)}$  are  $n_j$ -dimensional vectors and  $\alpha_i$  are numbers. Then the minimum value of  $r$  is called the rank of  $\mathcal{A}$ . When  $\mathcal{A}$  is symmetric,  $n_1 = \cdots = n_m$  and  $\mathcal{A}$  can be decomposed as

$$\mathcal{A} = \sum_{i=1}^r \alpha_i (x^{(i)})^m,$$

where  $x^{(i)}$  are  $n$ -dimensional vectors and  $\alpha_i$  are numbers. The minimum value of  $r$  is called the symmetric rank of  $\mathcal{A}$ . It is conjectured by Comon et al. [2] that for a symmetric tensor, its symmetric rank is the same as its rank. We call this conjecture the symmetric rank conjecture.

**Acknowledgments.** The authors would like to thank Prof. Pierre Comon, the associate editor, and the anonymous referees for their constructive comments and suggestions which lead to a significantly improved version of the paper. One referee of this paper pointed out that Theorem 2.1 can be regarded as the first step to prove the symmetric rank conjecture.

## REFERENCES

- [1] P. COMON, *Blind channel identification and extraction of more sources than sensors*, in Proceedings of the Conference on Advanced Signal Processing, Architectures, and Implementations VIII, San Diego, 1998, pp. 2–13.
- [2] P. COMON, G. GOLUB, L.H. LIM, AND B. MOURRAIN, *Symmetric tensors and symmetric tensor rank*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1254–1279.
- [3] A. GHOSH, E. TSIGARIDAS, M. DESCOTEAUX, P. COMON, B. MOURRAIN, AND R. DERICHE, *A polynomial based approach to extract the maxima of an antipodally symmetric spherical function and its application to extract fiber directions from the orientation distribution function in diffusion MRI*, Computational Diffusion MRI Workshop, New York, 2008.
- [4] O. GRELLIER AND P. COMON, *Analytical blind discrete source separation*, in Proceedings of the the European Signal Processing Conference, Tampere, Finland, 2000.
- [5] S. HE, Z. LI, AND S. ZHANG, *Approximation algorithms for homogeneous polynomial optimization with quadratic constraints*, Math. Program., 125 (2011), pp. 353–383.
- [6] E. KOFIDIS AND P.A. REGALIA, *On the best rank-1 approximation of higher-order supersymmetric tensors*, SIAM J. Matrix Anal. Appl., 23 (2002), pp. 863–884.
- [7] T.G. KOLDA AND J.R. MAYO, *Shifted power method for computing tensor eigenpairs*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 1095–1124.
- [8] S. LANG, *Real and functional analysis*, Grad. Texts in Math. 142, Springer-Verlag, New York, 1993.
- [9] J.B. LASSERRE, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., 11 (2001), pp. 796–817.
- [10] J.B. LASSERRE, *Polynomials nonnegative on a grid and discrete representations*, Trans. Amer. Math. Soc., 354 (2001), pp. 631–649.
- [11] L. DE LATHAUWER, *First-order perturbation analysis of the best rank- $(R_1, R_2, R_3)$  approximation in multilinear algebra*, J. Chemometrics, 18 (2004), pp. 2–11.
- [12] L. DE LATHAUWER, B.D. MOOR, AND J. VANDEWALLE, *On the best rank-1 and rank- $(R_1, R_2, \dots, R_N)$  approximation of higher-order tensor*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1324–1342.
- [13] M. LAURENT, *Sums of squares, moment matrices and optimization over polynomials*, in Emerging Applications of Algebraic Geometry, IMA Volumes in Mathematics and its Applications 149, M. Putinar and S. Sullivant, eds., Springer, New York, 2009, pp. 157–270.
- [14] X.L. LI, *A new gradient search interpretation of super exponential algorithm*, IEEE Signal Process. Lett., 13 (2006), pp. 173–176.
- [15] Z.Q. LUO AND S. ZHANG, *A semidefinite relaxation scheme for multivariate quartic polynomial optimization with quadratic constraints*, SIAM J. Optim., 20 (2010), pp. 1716–1736.
- [16] H.M. MARKOWITZ, *Portfolio selection*, J. Finance, 7 (1952), pp. 77–91.
- [17] P. McCULLAGH, *Tensor Methods in Statistics*, Chapman and Hall, London, 1987.
- [18] J.G. McWHIRTER AND I.K. PROUDLER, *Mathematics in Signal Processing V*, Clarendon Press, Oxford, UK, 2002.
- [19] J.J. MORÉ AND D.C. SORESENSEN, *Computing a trust region step*, SIAM J. Sci. Statist. Comput., 4 (1983), pp. 553–572.
- [20] Y. NESTEROV, *Random Walk in a Simplex and Quadratic Optimization over Convex Polytopes*, Catholic University of Louvain, Louvain-la-Neuve, Belgium, 2003.
- [21] G. NI AND Y. WANG, *On the best rank-1 approximation to higher-order symmetric tensor*, Math. Comput. Modelling, 46 (2007), pp. 1345–1352.
- [22] Q. NI, L. QI, AND F. WANG, *An eigenvalue method for testing positive definiteness of a multivariate form*, IEEE Trans. Automat. Control, 53 (2008), pp. 1096–1107.
- [23] J. NIE, *Sum of squares methods for minimizing polynomial forms over spheres and hypersurfaces*, Frontiers of Mathematics in China, 7 (2012), pp. 321–346.
- [24] P.A. PARRILO, *Semidefinite programming relaxations for semialgebraic problems*, Math. Program., 96 (2003), pp. 293–320.
- [25] L. QI, *Extrema of a real polynomial*, J. Global Optim., 30 (2004), pp. 405–433.
- [26] L. QI, *Eigenvalues of a real supersymmetric tensor*, J. Symbolic Comput., 40 (2005), pp. 1302–1324.
- [27] L. QI, *The best rank-one approximation ratio of a tensor space*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 430–442.
- [28] L. QI, Z. WAN, AND Y.F. YANG, *Global minimization of normal quadratic polynomials based on global descent directions*, SIAM J. Optim., 15 (2004), pp. 275–302.
- [29] L. QI, F. WANG, AND Y. WANG, *Z-eigenvalue methods for a global polynomial optimization problem*, Math. Program., 118 (2009), pp. 301–316.

- [30] P.A. REGALIA AND E. KOFIDES, *Monotonic convergence of fixed-point algorithms for ICA*, IEEE Trans. Neural Networks, 14 (2003), pp. 943–949.
- [31] A. M-C. SO, *Deterministic approximation algorithms for sphere constrained homogeneous polynomial optimization problems*, Math. Program., 129 (2011), pp. 357–382.
- [32] S. SOARE, J.W. YOON, AND O. CAZACU, *On the use of homogeneous polynomials to develop anisotropic yield functions with applications to sheet forming*, Internat. J. Plasticity, 24 (2008), pp. 915–944.
- [33] A.J. VAN DER VEEN AND A. PAULRAJ, *An analytical constant modulus algorithm*, IEEE Trans. Signal Process., 44 (1996), pp. 1136–1155.
- [34] V. ZARZOSO AND P. COMON, *Robust independent component analysis by iterative maximization of kurtosis contrast with algebraic optimal step size*, IEEE Trans. Neural Networks, 21 (2010), pp. 534–550.
- [35] T. ZHANG AND G.H. GOLUB, *Rank-1 approximation of higher-order tensors*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 534–550.
- [36] X.Z. ZHANG, L. QI, AND Y. YE, *The cubic spherical optimization problems*, Math. Comput., 81 (2012), pp. 1513–1525.



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.