

# A Shifted Power Method for Homogenous Polynomial Optimization over Unit Spheres

Shuquan Wang<sup>1</sup>

<sup>1</sup> School of Management Science, Qufu Normal University, China.

Correspondence: Shuquan Wang, School of Management Science, Qufu Normal University, Rizhao, Shandong, 276800, China. E-mail: msshuquan@163.com

Received: March 31, 2015 Accepted: April 21, 2015 Online Published: May 15, 2015

doi:10.5539/jmr.v7n2p175 URL: <http://dx.doi.org/10.5539/jmr.v7n2p175>

*This work was supported by the NSF of China (11171180).*

## Abstract

In this paper, we propose a shifted power method for a type of polynomial optimization problem over unit spheres. The global convergence of the proposed method is established and an easily implemented scope of the shifted parameter is provided.

**Keywords:** shifted power method, polynomial optimization problem, convergence

**AMS subject classification:** 65F99, 65K10, 15A69, 14M15, 90C53

## 1. Introduction

The polynomial optimization problem in which the objective function and constraints are polynomial functions received much attention recently due to their wide applications in such as signal processing (Ghosh 2008, Qi 2003), biomedical engineering (Kofidis 2002, Lasserre 2001), material science (Soare 2008), quantum mechanics (Dahl *et al.* 2008, Wang *et al.* 2009), and numerical linear algebra (Hof 2009, Qi 2005), see (Klerk 2008) for a survey on the various classes of polynomial optimization with simplex, hypercube, or sphere constraints. The polynomial optimization problem is a challenging task, as the simplest instances of polynomial optimization, such as maximizing a cubic polynomial over a sphere, is NP-hard (Nesterov 2003). However, researchers have made much contributions in this area from the theoretical side to numerical solution methods (Lathauwer 2000, He 2010, Lasserre 2001, Luo 2010, Qi 2004, 2009, Zhang 2012).

In this paper, we consider the following type of polynomial optimization over unit spheres

$$\begin{aligned} \min & f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s) \\ \text{s.t.} & \|\mathbf{x}_i\|^2 = 1, \mathbf{x}_i \in \mathbb{R}^{n_i} \quad i = 1, 2, \dots, s \end{aligned} \quad (1.1)$$

where  $f : \mathbf{R}^{n_1} \times \mathbf{R}^{n_2} \times \dots \times \mathbf{R}^{n_s} \rightarrow \mathbf{R}$  is a homogenous polynomial function whose each term is also  $d_i$ -order homogenous with respect to  $\mathbf{x}_i$  for  $i = 1, 2, \dots, s$ .

It is well known that tensor is a useful tool in polynomial optimization as a polynomial, especially a homogenous polynomial, has a very simple expression with the aid of tensor, and furthermore, the optimal condition of a homogenous polynomial optimization with a special structured feasible region can be deeply characterized (Qi 2005, Qi 2009).

A tensor is a multidimensional array of data whose elements are referred by using multiple indices, i.e.,

$$\mathcal{A} = (\mathcal{A}_{i_1 i_2 \dots i_d})_{n_1 \times n_2 \times \dots \times n_d},$$

where  $d$ , the number of indices, is called the order of the tensor, and the  $d$ -tuple  $(n_1, n_2, \dots, n_d)$  is called the dimension of the tensor. Tensor  $\mathcal{A} \in \mathbf{R}^{n_1 \times n_2 \times \dots \times n_d}$  is called super-symmetric if  $n_1 = n_2 = \dots = n_d$  and  $\mathcal{A}_{i_1 i_2 \dots i_d}$  is invariant under any permutation of indices  $(i_1 i_2 \dots i_d)$ , i.e.,  $\mathcal{A}_{i_1 i_2 \dots i_d} = \mathcal{A}_{\pi(i_1 i_2 \dots i_d)}$ , where  $\pi(i_1 i_2 \dots i_d)$  is a permutation

of indices  $(i_1 i_2 \cdots i_d)$ . Furthermore, a tensor  $\mathcal{A} \in \mathbf{R}^{n_1 \times n_2 \times \cdots \times n_d}$  is called partially symmetric if its indices can be partitioned into  $s$  index blocks and there exist positive integers  $d_1, d_2, \dots, d_s$  such that  $d_1 + d_2 + \cdots + d_s = d$ ,  $n_1 = n_2 = \cdots = n_{d_1}$ ,  $n_{d_1+1} = \cdots = n_{d_1+d_2}$ ,  $\dots$ ,  $n_{d_1+d_2+\cdots+d_{s-1}+1} = \cdots = n_d$ , and the element  $\mathcal{A}_{i_1 i_2 \cdots i_d}$  is invariant under any permutation of indices within each index block, i.e.,  $\mathcal{A}_{i_1 i_2 \cdots i_d} = \mathcal{A}_{\pi(i_1 i_2 \cdots i_{d_1}) \pi(i_{d_1+1} i_{d_1+2} \cdots i_{d_1+d_2}) \cdots \pi(i_{d_1+d_2+\cdots+d_{s-1}+1} \cdots i_d)}$ .

By virtue of partially symmetric tensor, problem (1.1) can be written as

$$\begin{aligned} \max \quad & \mathcal{A} \mathbf{x}_1^{d_1} \mathbf{x}_2^{d_2} \cdots \mathbf{x}_s^{d_s} \\ \text{s.t.} \quad & \|\mathbf{x}_i\|^2 = 1, \mathbf{x}_i \in \mathbb{R}^{n_i}, \quad i = 1, 2, \dots, s \end{aligned} \tag{1.2}$$

where  $\mathcal{A}$  is a partially symmetric tensor w.r.t. index blocks  $(i_1 i_2 \cdots i_{d_1}), (i_{d_1+1} i_{d_1+2}, \dots, i_{d_1+d_2}), \dots, (i_{d_1+d_2+\cdots+d_{s-1}+1} \cdots i_d)$ . This problem contains the problem of finding best rank-1 approximation or computing the largest eigenvalue in magnitude of a super-symmetric tensor (Kofidis 2002, Kolda 2011, Zhang 2012, Qi 2009) as special cases. If  $s = 2$  and  $d_1 = d_2 = 2$ , then this problem reduces to the problem considered in (Dah 2008, Wang 2009) arising from the nonlinear elastic materials analysis and entanglement studies in quantum physics. The problem is also a special case of the spherically constrained homogenous polynomial optimization problem considered in (Chen 2012).

Generally, there are three popular numerical solution methods for polynomial optimization. The first one is the SOS approach which is based on the decomposition of a multivariate polynomials into sum of squares (Lasserre 2001, Parrilo 2003), and the second one is the semidefinite programming relaxation technique (Luo 2010) where the concerned problem is approximated by a specially constructed semidefinite programming problem. These two kinds of methods can obtain a global solution of the problem in a sense, however, the computing quantity of these two methods is too large and they are efficient for small scale polynomial optimization problems. The third method is the power method. This method is initiated from the power method for computing the largest eigenvalue in magnitude of a square matrix (Golub 1996), and later was successfully extended to computing the largest singular eigenvalue of a higher-order tensor (Lathauwer 2000) and the largest Z-eigenvalue in magnitude of higher-order super-symmetric tensor (Kofidis 2002, Qi 2005), which in fact are a kind of homogeneously polynomial optimization problem with unit sphere constraints (Lathauwer 2000, kofidis 2002). It is also used to compute the largest eigenvalue of a nonnegative tensor. The distinct feature of this method is that it needs less computing cost at each iteration and its convergence can be guaranteed under convexity assumption. To remove the confine, a novel shifted technique is introduced into the objective function (Kolda 2011, Wang 2009).

In this paper, we extend the shifted power method to a more general type of polynomial optimization problem (1.1) and establish the convergence of the method. The main contribution of the paper is as follows.

- (1) We apply the shifted power method to a more general type of polynomial optimization problem defined on the unit spheres and establish its convergence.
- (2) We provide an easily implemented scope of the shifted parameters used in the designed iterative method.

The content of this paper is organized as follows. In the Section 2, we will give a short description on the tensor algebra and give some notations. In Section 3, we will give the design power method and establish its convergence.

### 2. Preliminaries and Notations

In this section, we will give a short description on the tensor algebra and some notations used throughout the paper.

For tensors  $\mathcal{A}, \mathcal{B} \in \mathbf{R}^{n_1 \times n_2 \times \cdots \times n_d}$ , their inner product is defined by

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_d=1}^{n_d} \mathcal{A}_{i_1 i_2 \cdots i_d} \mathcal{B}_{i_1 i_2 \cdots i_d}.$$

For vectors  $\mathbf{x}_i \in \mathbf{R}^{n_i}$ ,  $i = 1, 2, \dots, d$ , their outer product, denoted by  $\mathcal{X} = \mathbf{x}_1 \circ \mathbf{x}_2 \circ \cdots \circ \mathbf{x}_d$ , is a  $d$ -order  $(n_1, \dots, n_d)$ -dimensional tensor  $\mathcal{T}$  with the  $(i_1 i_2, \dots, i_d)$ -entry

$$\mathcal{T}_{i_1 i_2 \cdots i_d} = (\mathbf{x}_1)_{i_1} (\mathbf{x}_2)_{i_2} \cdots (\mathbf{x}_d)_{i_d},$$

where  $(\mathbf{x}_i)_{i_r}$  denotes the  $i_r$ -th entry of vector  $\mathbf{x}_i$  for  $r = 1, 2, \dots, n_i$ .

Based on the inner product of tensors with same order and same size, the inner product of  $\mathcal{X}$  with a general tensor  $\mathcal{A} \in \mathbf{R}^{n_1 \times n_2 \times \cdots \times n_d}$  can be expressed as

$$\langle \mathcal{A}, \mathcal{X} \rangle = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_d=1}^{n_d} \mathcal{A}_{i_1 i_2 \cdots i_d} (\mathbf{x}_1)_{i_1} (\mathbf{x}_2)_{i_2} \cdots (\mathbf{x}_d)_{i_d}.$$

For simplicity, we denote the product by  $\mathcal{A}\mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_d$ . If  $\mathcal{A}$  is super-symmetric of dimension  $n$  and  $\mathbf{x}_1 = \mathbf{x}_2 = \cdots = \mathbf{x}_d = \mathbf{x} \in \mathbf{R}^n$ , then the product above is denoted as  $\mathcal{A}\mathbf{x}^d$ . Certainly,  $\mathcal{A}\mathbf{x}^{d-1}$  is a  $n$ -dimensional vector with  $i$ -th entry

$$(\mathcal{A}\mathbf{x}^{d-1})_i = \sum_{i_2=1}^n \sum_{i_3=1}^n \cdots \sum_{i_d=1}^n \mathcal{A}_{ii_2i_3\cdots i_d} \mathbf{x}_{i_2} \mathbf{x}_{i_3} \cdots \mathbf{x}_{i_d},$$

and  $\mathcal{A}\mathbf{x}^{d-2}$  is an  $n \times n$  symmetric matrix.

Throughout this paper, we use  $\Gamma_n$  and  $\Sigma_n$  to denote the unit ball and sphere in  $\mathbf{R}^n$ , and the subscript  $n$  is omitted if no confusion is made.

### 3. Algorithm and Convergence

For problem (1.1), since the linear independence of the constraint qualification is always satisfied, thus for any optimal solution  $(\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_s^*)$ , there exists multipliers  $\lambda_i \in \mathbf{R}$ ,  $i = 1, 2, \dots, s$  such that

$$\begin{cases} \mathcal{A}(\mathbf{x}_1^*)^{d_1-1}(\mathbf{x}_2^*)^{d_2} \cdots (\mathbf{x}_s^*)^{d_s} = \frac{\lambda_1}{d_1} \mathbf{x}_1^*, \\ \mathcal{A}(\mathbf{x}_1^*)^{d_1}(\mathbf{x}_2^*)^{d_2-1} \cdots (\mathbf{x}_s^*)^{d_s} = \frac{\lambda_2}{d_2} \mathbf{x}_2^*, \\ \dots\dots\dots \\ \mathcal{A}(\mathbf{x}_1^*)^{d_1-1}(\mathbf{x}_2^*)^{d_2} \cdots (\mathbf{x}_s^*)^{d_s-1} = \frac{\lambda_s}{d_s} \mathbf{x}_s^*. \end{cases}$$

This system constitutes the KKT condition of problem (1.2), which corresponds to the stationary point condition of Lagrange function of the problem (Nocedal 1999).

As  $\|\mathbf{x}_i^*\|^2 = 1$  for  $i = 1, \dots, s$ , we conclude that

$$\frac{\lambda_1}{d_1} = \dots = \frac{\lambda_s}{d_s}.$$

Certainly, this is the optimal objective function value at point  $(\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_s^*)$ , i.e.,

$$\frac{\lambda_i}{d_i} = \mathcal{A}\mathbf{x}_1^{*d_1} \mathbf{x}_2^{*d_2} \cdots \mathbf{x}_s^{*d_s}, \quad i = 1, 2, \dots, s.$$

Based on the shifted power method for computing the largest eigenpairs of a super-symmetric tensor (Kolda 2011, Wang 2009), we have the following shifted power method for solving problem (1.1).

#### Algorithm 3.1.

**Input:** partially symmetric tensor  $\mathcal{A} \in \mathbf{R}^{n_1 \times \dots \times n_1 \times n_2 \times \dots \times n_2 \times \dots \times n_s \times \dots \times n_s}$ , positive shifted parameters  $\alpha_1, \alpha_2, \dots, \alpha_s$ , initial point  $\mathbf{x}_i^{(0)} \in \mathbf{R}^{n_i}$ ,  $i = 1, 2, \dots, s$ , set  $k = 0$  and take tolerance  $\varepsilon \geq 0$ .

**Output:** approximated solution  $\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_s^*$  to problem (1.1)

**Iterative Step:** FOR  $k = 0, 1, 2, \dots$ , DO

$$\begin{aligned} \hat{\mathbf{x}}_1^{(k+1)} &= \mathcal{A}(\mathbf{x}_1^{(k)})^{d_1-1}(\mathbf{x}_2^{(k)})^{d_2} \cdots (\mathbf{x}_s^{(k)})^{d_s} + \alpha_1 \mathbf{x}_1, & \mathbf{x}_1^{(k+1)} &= \frac{\hat{\mathbf{x}}_1^{(k+1)}}{\|\hat{\mathbf{x}}_1^{(k+1)}\|}, \\ \hat{\mathbf{x}}_2^{(k+1)} &= \mathcal{A}(\mathbf{x}_1^{(k+1)})^{d_1}(\mathbf{x}_2^{(k)})^{d_2-1} \cdots (\mathbf{x}_s^{(k)})^{d_s} + \alpha_2 \mathbf{x}_2, & \mathbf{x}_2^{(k+1)} &= \frac{\hat{\mathbf{x}}_2^{(k+1)}}{\|\hat{\mathbf{x}}_2^{(k+1)}\|}, \\ \dots & \dots \dots \dots, & \dots & \dots \\ \hat{\mathbf{x}}_s^{(k+1)} &= \mathcal{A}(\mathbf{x}_1^{(k+1)})^{d_1}(\mathbf{x}_2^{(k+1)})^{d_2} \cdots (\mathbf{x}_s^{(k)})^{d_s-1} + \alpha_s \mathbf{x}_s, & \mathbf{x}_s^{(k+1)} &= \frac{\hat{\mathbf{x}}_s^{(k+1)}}{\|\hat{\mathbf{x}}_s^{(k+1)}\|}. \\ \lambda^{(k+1)} &= \mathcal{A}(\mathbf{x}_1^{(k+1)})^{d_1}(\mathbf{x}_2^{(k+1)})^{d_2} \cdots (\mathbf{x}_s^{(k+1)})^{d_s}. \end{aligned}$$

IF  $|\lambda^{(k+1)} - \lambda^{(k)}| \leq \varepsilon$ , terminate; otherwise,  $k = k + 1$ .

END IF

END FOR

To see the well-definedness of the algorithm, we need to show that, at each iteration, the gradient of the shifted objective function,

$$g(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s) = f(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s) + \sum_{i=1}^s \alpha_i \mathbf{x}_i^\top \mathbf{x}_i = \mathcal{A} \mathbf{x}_1^{d_1} \mathbf{x}_2^{d_2} \dots \mathbf{x}_s^{d_s} + \sum_{i=1}^s \alpha_i \mathbf{x}_i^\top \mathbf{x}_i$$

w.r.t.  $\mathbf{x}_i$  on  $\Sigma_{n_1} \times \dots \times \Sigma_{n_s}$  is bounded away from zero for  $i = 1, 2, \dots, s$ .

**Lemma 3.1.** *There exist  $\alpha_i > 0, i = 1, 2, \dots, s$  such that the polynomial function*

$$g(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s) = \mathcal{A} \mathbf{x}_1^{d_1} \mathbf{x}_2^{d_2} \dots \mathbf{x}_s^{d_s} + \sum_{i=1}^s \alpha_i \mathbf{x}_i^\top \mathbf{x}_i$$

*is strictly convex with respect to  $\mathbf{x}_i$  on  $\Gamma_{n_1} \times \dots \times \Gamma_{n_s}$  for  $i = 1, 2, \dots, s$ , and its gradient w.r.t.  $\mathbf{x}_i$  on  $\Sigma_{n_1} \times \dots \times \Sigma_{n_s}$  is bounded away from zero for  $i = 1, 2, \dots, s$ .*

*Proof.* With the notation of the product of tensor, the function  $g$  can be expressed as

$$\begin{aligned} g(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s) &= \mathcal{A} \mathbf{x}_1^{d_1} \mathbf{x}_2^{d_2} \dots \mathbf{x}_s^{d_s} + \sum_{i=1}^s \alpha_i \mathbf{x}_i^\top \mathbf{x}_i \\ &= \mathbf{x}_i^\top (\mathbf{unfold}(\mathcal{A} \mathbf{x}_1^{d_1} \mathbf{x}_2^{d_2} \dots \mathbf{x}_i^{d_i-2} \dots \mathbf{x}_s^{d_s}) \mathbf{x}_i + \alpha I_{n_i}) \mathbf{x}_i + \sum_{i=1, j \neq i}^s \alpha_j \mathbf{x}_j^\top \mathbf{x}_j. \end{aligned}$$

where  $\mathbf{unfold}(\mathcal{A} \mathbf{x}_1^{d_1} \mathbf{x}_2^{d_2} \dots \mathbf{x}_i^{d_i-2} \dots \mathbf{x}_s^{d_s})$  is the unfolded matrix of 2-order tensor in  $\mathbf{R}^{n_i \times n_i}$ , and  $I_{n_i}$  is an identity matrix of dimension  $n_i$ .

To show the strict convexity of function  $g(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s)$  w.r.t.  $\mathbf{x}_i$  on  $\Gamma_{n_1} \times \dots \times \Gamma_{n_s}$  for  $i = 1, 2, \dots, s$ , it suffices to show that the unfolded matrix of 2-order tensor  $\mathcal{A} \mathbf{x}_1^{d_1} \mathbf{x}_2^{d_2} \dots \mathbf{x}_i^{d_i-2} \dots \mathbf{x}_s^{d_s} + \alpha I_{n_i}$  is positive definite on  $\Gamma_{n_1} \times \dots \times \Gamma_{n_s}$ . From the partial symmetry of tensor  $\mathcal{A}$ , this matrix is symmetric, and from the Geršgorin disc theorem (Golub 1996), the spectral radius of matrix  $\mathbf{unfold}(\mathcal{A} \mathbf{x}_1^{d_1} \mathbf{x}_2^{d_2} \dots \mathbf{x}_i^{d_i-2} \dots \mathbf{x}_s^{d_s})$  is less than or equal to

$$\sum_{i_1, i_2, \dots, i_{d_1}=1}^{n_1} \sum_{i_{d_1+1}, i_{d_1+2}, \dots, i_{d_1+d_2}=1}^{n_2} \dots \sum_{i_{d_1+\dots+d_{s-1}+1}, \dots, i_d=1}^{n_d} |\mathcal{A}_{i_1, i_2, \dots, i_d}|.$$

Denote the sum above as  $\alpha$ . Then, for any  $\alpha_i > \alpha$ , the matrix

$$\mathbf{unfold}(\mathcal{A} \mathbf{x}_1^{d_1} \mathbf{x}_2^{d_2} \dots \mathbf{x}_i^{d_i-2} \dots \mathbf{x}_s^{d_s}) + \alpha_i I_{n_i}$$

is positive definite w.r.t.  $\mathbf{x}_1, \dots, \mathbf{x}_s$  on  $\Gamma_{n_1} \times \dots \times \Gamma_{n_s}$ . From the continuity of the smallest eigenvalue and largest eigenvalue of matrix w.r.t. its elements and the boundedness of  $\Gamma_{n_1} \times \dots \times \Gamma_{n_s}$ , we conclude that the eigenvalue of the above matrix is bounded away from zero on  $\Gamma_{n_1} \times \dots \times \Gamma_{n_s}$ . This means that the matrix is nonsingular w.r.t.  $\mathbf{x}_1, \dots, \mathbf{x}_s$  on  $\Gamma_{n_1} \times \dots \times \Gamma_{n_s}$ , and thus the gradient of the polynomial function  $g(\mathbf{x}_1, \dots, \mathbf{x}_s)$  w.r.t.  $\mathbf{x}_j$  on  $\Sigma_{n_1} \times \dots \times \Sigma_{n_s}$  is bounded away from zero for  $j = 1, 2, \dots, s$ .  $\square$

**Theorem 3.1.** *For Algorithm 3.1 with  $\varepsilon = 0$ , if  $\alpha_i$  for  $i = 1, 2, \dots, k$  satisfies that*

$$g(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s) = \mathcal{A} \mathbf{x}_1^{d_1} \mathbf{x}_2^{d_2} \dots \mathbf{x}_s^{d_s} + \sum_{i=1}^s \alpha_i \mathbf{x}_i^\top \mathbf{x}_i$$

*is strictly convex with respect to  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s$ , then*

*(1) the generated sequence  $\{g(\mathbf{x}_1^{(k)}, \mathbf{x}_2^{(k)}, \dots, \mathbf{x}_s^{(k)})\}$  is strictly increasing;*

*(2) if the algorithm terminates after finite steps, then final point is a KKT point of problem (1.1), and if algorithm generate an infinite sequence  $\{(\mathbf{x}_1^{(k)}, \mathbf{x}_2^{(k)}, \dots, \mathbf{x}_s^{(k)})\}$ , then*

$$\lim_{k \rightarrow \infty} \|(\mathbf{x}_1^{(k+1)}, \mathbf{x}_2^{(k+1)}, \dots, \mathbf{x}_s^{(k+1)}) - (\mathbf{x}_1^{(k)}, \mathbf{x}_2^{(k)}, \dots, \mathbf{x}_s^{(k)})\| = 0$$

*and any accumulation point of  $\{(\mathbf{x}_1^{(k)}, \mathbf{x}_2^{(k)}, \dots, \mathbf{x}_s^{(k)})\}$  is a KKT point of the problem.*

*Proof.* To prove (1), we first show that for any  $k = 0, 1, 2, \dots$  and  $i = 1, 2, \dots, s$ , it holds that

$$g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_{i-1}^{(k+1)}, \mathbf{x}_i^{(k+1)}, \mathbf{x}_{i+1}^{(k)}, \dots, \mathbf{x}_s^{(k)}) \geq g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_{i-1}^{(k+1)}, \mathbf{x}_i^{(k)}, \mathbf{x}_{i+1}^{(k)}, \dots, \mathbf{x}_s^{(k)}). \tag{3.1}$$

In fact, since the function  $g(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_s)$  is convex w.r.t.  $\mathbf{x}_i$  for  $i = 1, 2, \dots, s$  on  $\Gamma_{n_1} \times \Gamma_{n_2} \times \dots \times \Gamma_{n_s}$ , thus for any  $i = 1, 2, \dots, s$ , it holds that

$$\begin{aligned} & g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_{i-1}^{(k+1)}, \mathbf{x}_i^{(k+1)}, \mathbf{x}_{i+1}^{(k)}, \dots, \mathbf{x}_s^{(k)}) - g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_{i-1}^{(k+1)}, \mathbf{x}_i^{(k)}, \mathbf{x}_{i+1}^{(k)}, \dots, \mathbf{x}_s^{(k)}) \\ & \geq \langle \nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_{i-1}^{(k+1)}, \mathbf{x}_i^{(k+1)}, \mathbf{x}_{i+1}^{(k)}, \dots, \mathbf{x}_s^{(k)}), \mathbf{x}_i^{(k+1)} - \mathbf{x}_i^{(k)} \rangle. \end{aligned} \tag{3.2}$$

Owing to

$$\mathbf{x}_i^{(k+1)} = \frac{\nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_{i-1}^{(k+1)}, \mathbf{x}_i^{(k)}, \dots, \mathbf{x}_s^{(k)})}{\|\nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_{i-1}^{(k+1)}, \mathbf{x}_i^{(k)}, \dots, \mathbf{x}_s^{(k)})\|},$$

one has from the Cauchy-Schwarz inequality that

$$\langle \nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k+1)}, \mathbf{x}_2^{(k+1)}, \dots, \mathbf{x}_i^{(k+1)}, \mathbf{x}_{i+1}^{(k)}, \dots, \mathbf{x}_s^{(k)}), \mathbf{x}_i^{(k)} - \mathbf{x}_i^{(k+1)} \rangle \geq 0. \tag{3.3}$$

This means that inequality (3.1) holds.

To show the assertion, we further need to analyze inequality (3.2). For this inequality, as

$$\nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_i^{(k+1)}, \mathbf{x}_{i+1}^{(k)}, \dots, \mathbf{x}_s^{(k)}) \neq 0, \quad k = 0, 1, 2, \dots$$

from Lemma 3.1 and the assumption, by the Cauchy-Schwarz inequality, one has

$$\langle \nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_i^{(k+1)}, \mathbf{x}_{i+1}^{(k)}, \dots, \mathbf{x}_s^{(k)}), \mathbf{x}_i^{(k+1)} - \mathbf{x}_i^{(k)} \rangle = 0$$

if and only if

$$\mathbf{x}_i^{(k+1)} = \mathbf{x}_i^{(k)} = \frac{\nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_i^{(k+1)}, \mathbf{x}_{i+1}^{(k)}, \dots, \mathbf{x}_s^{(k)})}{\|\nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_i^{(k+1)}, \mathbf{x}_{i+1}^{(k)}, \dots, \mathbf{x}_s^{(k)})\|}. \tag{3.4}$$

From the iterative procedure, (3.4) is also equivalent to

$$g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_{i-1}^{(k+1)}, \mathbf{x}_i^{(k)}, \dots, \mathbf{x}_s^{(k)}) = g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_{i-1}^{(k+1)}, \mathbf{x}_i^{(k+1)}, \mathbf{x}_{i+1}^{(k)}, \dots, \mathbf{x}_s^{(k)}).$$

Thus, the sequence  $\{g(\mathbf{x}_1^{(k+1)}, \mathbf{x}_2^{(k+1)}, \dots, \mathbf{x}_s^{(k+1)})\}$  is strictly increasing.

For (2), we first consider the case that the algorithm terminates after finite steps, say, stops at Step  $k_0$ . Then from Algorithm 3.1 and discussion for (1),

$$g(\mathbf{x}_1^{(k_0+1)}, \mathbf{x}_2^{(k_0+1)}, \dots, \mathbf{x}_s^{(k_0+1)}) = g(\mathbf{x}_1^{(k_0)}, \mathbf{x}_2^{(k_0)}, \dots, \mathbf{x}_s^{(k_0)})$$

and

$$\mathbf{x}_i^{(k_0+1)} = \mathbf{x}_i^{(k_0)}, \quad i = 1, 2, \dots, s.$$

Combining this with the iterative procedure of Algorithm 3.1 yields that

$$\nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k_0)}, \dots, \mathbf{x}_i^{(k_0)}, \dots, \mathbf{x}_s^{(k_0)}) = \|\nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k_0)}, \dots, \mathbf{x}_i^{(k_0)}, \dots, \mathbf{x}_s^{(k_0)})\| \mathbf{x}_i^{(k_0)},$$

i.e.,

$$\nabla_{\mathbf{x}_i} f(\mathbf{x}_1^{(k_0)}, \dots, \mathbf{x}_i^{(k_0)}, \dots, \mathbf{x}_s^{(k_0)}) = (\|\nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k_0)}, \dots, \mathbf{x}_i^{(k_0)}, \dots, \mathbf{x}_s^{(k_0)})\| - \alpha_i) \mathbf{x}_i^{(k_0)}.$$

This means that  $(\mathbf{x}_1^{(k_0)}, \mathbf{x}_2^{(k_0)}, \dots, \mathbf{x}_s^{(k_0)})$  is a KKT point of the problem (1.2).

Now, consider the case that the algorithm generates an infinite sequence. In this case, as the sequence  $g(\mathbf{x}_1^{(k)}, \dots, \mathbf{x}_s^{(k)})$  is increasing, we know that it converges to  $g^*$  as the sequence is defined over unit spheres and thus bounded from above. In particular, one has

$$\lim_{k \rightarrow \infty} (g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_s^{(k+1)}) - g(\mathbf{x}_1^{(k)}, \dots, \mathbf{x}_s^{(k)})) = 0.$$

Taking (3.2) and (3.3) into consideration, we obtain that

$$\lim_{k \rightarrow \infty} \langle \nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_i^{(k+1)}, \mathbf{x}_{i+1}^{(k)}, \dots, \mathbf{x}_s^{(k)}), \mathbf{x}_i^{(k+1)} - \mathbf{x}_i^{(k)} \rangle = 0, \quad i = 1, 2, \dots, s.$$

As

$$\mathbf{x}_i^{(k+1)} = \frac{\nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_{i-1}^{(k+1)}, \mathbf{x}_i^{(k)}, \dots, \mathbf{x}_s^{(k)})}{\|\nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_{i-1}^{(k+1)}, \mathbf{x}_i^{(k)}, \dots, \mathbf{x}_s^{(k)})\|}, \quad i = 1, 2, \dots, s, \quad (3.5)$$

and  $\nabla_{\mathbf{x}_i} g(\mathbf{x}_1^{(k+1)}, \dots, \mathbf{x}_{i-1}^{(k+1)}, \mathbf{x}_i^{(k)}, \dots, \mathbf{x}_s^{(k)})$  is bounded away from zero on  $\Gamma_{n_1} \times \dots \times \Gamma_{n_s}$ , we conclude that

$$\lim_{k \rightarrow \infty} \|\mathbf{x}_i^{(k+1)}\|^2 - \langle \mathbf{x}_i^{(k+1)}, \mathbf{x}_i^{(k)} \rangle = 0, \quad i = 1, 2, \dots, s.$$

Due to that  $\|\mathbf{x}_i^{(k)}\| = \|\mathbf{x}_i^{(k+1)}\| = 1$ , one in turn has

$$\lim_{k \rightarrow \infty} \langle \mathbf{x}_i^{(k+1)}, \mathbf{x}_i^{(k)} \rangle = 1, \quad i = 1, 2, \dots, s,$$

and

$$\lim_{k \rightarrow \infty} \|\mathbf{x}_i^{(k+1)} - \mathbf{x}_i^{(k)}\| = 0, \quad i = 1, 2, \dots, s. \quad (3.6)$$

Thus,

$$\lim_{k \rightarrow \infty} \|(\mathbf{x}_1^{(k+1)}, \mathbf{x}_2^{(k+1)}, \dots, \mathbf{x}_s^{(k+1)}) - (\mathbf{x}_1^{(k)}, \mathbf{x}_2^{(k)}, \dots, \mathbf{x}_s^{(k)})\| = 0.$$

Since the generated sequence  $\{(\mathbf{x}_1^{(k+1)}, \mathbf{x}_2^{(k+1)}, \dots, \mathbf{x}_s^{(k+1)})\}$  is bounded, by the Bolzano-Weierstrass theorem, the sequence must have at least one accumulation point and these accumulation points constitute a connected set on  $\Gamma_{n_1} \times \dots \times \Gamma_{n_s}$ . Now, we show that each such accumulation point is a KKT point of problem (1.1).

In fact, suppose  $\{(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_s)\}$  is an accumulation point of the generated sequence, then there exists a subsequence  $\{(\mathbf{x}_1^{(k_j)}, \mathbf{x}_2^{(k_j)}, \dots, \mathbf{x}_s^{(k_j)})\}$  converging to this point. From (3.6), we know that  $\{(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_s)\}$  is also a limit of subsequence  $\{(\mathbf{x}_1^{(k_j+1)}, \mathbf{x}_2^{(k_j+1)}, \dots, \mathbf{x}_s^{(k_j+1)})\}$ . Replacing  $k$  by  $k_j$  in (3.5) and letting  $j \rightarrow \infty$  yield that  $(\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_s)$  is a KKT point of problem (1.1) as discussed in the case of terminating within finite steps.  $\square$

#### 4. Conclusion

In this paper, we proposed a shifted power method for a type of polynomial optimization problem over unit spheres. Its validity is guaranteed theoretically. However, numerical experiments and numerical comparison are required to show the efficiency of the method and this will be done in our future research.

#### References

- Chen, B., He, S., Li, Z., & Zhang, S. (2012). Maximum Block Improvement and Polynomial Optimization, *SIAM J. Optim.*, 22, 87-107. <http://dx.doi.org/10.1137/110834524>
- Dahl, G., Leinaas, J. M., Myrheim, J., & Ovrum, E. (2007). A tensor product matrix approximation problem in quantum physics. *Linear Algebra Appl.*, 420, 711-725. <http://dx.doi.org/10.1016/j.laa.2006.08.026>
- de Klerk, E. (2008). The complexity of optimizing over a simplex, hypercube or sphere: A short survey. *Cent. Eur. J. Oper. Res.*, 16, 111-125. <http://dx.doi.org/10.1007/s10100-007-0052-9>
- de Lathauwer, L., De Moor, B., & Vandewalle, J. (2000). On the best rank-1 and rank- $(R_1, R_2, \dots, R_N)$  approximation of higher-order tensors. *SIAM Journal of Matrix Analysis and Applications*, 21, 1324-1342. <http://dx.doi.org/10.1137/S0895479898346995>
- Ghosh, A., Tsigaridas, E., Descoteaux, M., Comon, P., Mourrain, B., & Deriche, R. (2008). 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, in Proceedings of the Computational Diffusion MRI Workshop (CDMRI08), New York.
- Golub, G. H., & Van Loan, C. F. (1996). *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD.
- He, S., Li, Z., & Zhang, S. (2010). Approximation algorithms for homogeneous polynomial optimization with quadratic constraints. *Math. Program.*, 125, 353-383. <http://dx.doi.org/10.1007/s10107-010-0409-z>

- Hof, P. M. J., Scherer, C., & Heuberger, P. S. C. (2009). Model-Based Control: Bridging Rigorous Theory and Advanced Technology, Part 1, Springer, New York, 49-68. <http://dx.doi.org/10.1007/978-1-4419-0895-7>
- Kofidis, E., & Regalia, P. A. (2002). On the best rank-1 approximation of higher-order supersymmetric tensors. *SIAM Journal of Matrix Analysis and Applications*, 23, 863-884.
- Lasserre, J. B. (2001). Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11, 796-817. <http://dx.doi.org/10.1137/S1052623400366802>
- Ling, C., Zhang, X. Z., & Qi, L. Q. (2012). Semidefinite relaxation approximation for multivariate bi-quadratic optimization with quadratic constraints. *Numer Linear Algebra and Applications*, 19, 113-131.
- Luo, Z. Q., & Zhang, S. (2010). A semidefinite relaxation scheme for multivariate quartic polynomial optimization with quadratic constraints. *SIAM J. Optim.*, 20, 1716-1736.
- Nesterov, Y. (2003). Random Walk in a Simplex and Quadratic Optimization over Convex Polytopes, CORE Discussion Paper, UCL, Louvain-la-Neuve, Belgium.
- Nocedal, J., & Wright, S. J. (1999). Numerical Optimization, New York, Springer.
- Parrilo, P. A. (2003). Semidefinite programming relaxations for semialgebraic problems, Mathematical Programming, Series B, 96, 293-320.
- Qi, L., & Teo, K. L. (2003). Multivariate polynomial minimization and its applications in signal processing. *J. Global Optim.*, 26, 419-433.
- Qi, L. Q., Wan, Z., & Yang, Y. F. (2004). Global minimization of normal quartic polynomials based on global descent directions. *SIAM J Optim*, 15, 275-302.
- Qi, L. Q. (2005). Eigenvalues of a real supersymmetric tensor. *J Symbolic Computation*, 40, 1302-1324. <http://dx.doi.org/10.1016/j.jsc.2005.05.007>
- Qi, L., Wang, F., & Wang, Y. (2009). Z-eigenvalue methods for a global polynomial optimization problem. *Math Programming*, 118, 301-316.
- Soare, S., Yoon, J. W., & Cazacu, O. (2008). On the use of homogeneous polynomials to develop anisotropic yield functions with applications to sheet forming. *Int. J. Plasticity*, 24, 915-944.
- Tamara, G. K., & Jackson, R. M. (2012). Shifted power method for computing tensor eigenvalues. *SIAM J Matrix Analysis and Applications*, 32, 1095-1124.
- Wang, Y. J., Qi, L. Q., & Zhang, X. Z. (2009). A practical method for computing the largest M-eigenvalue of a fourth-order partially symmetric tensor. *Numerical Linear Algebra with Applications*, 16, 589-601. <http://dx.doi.org/10.1002/nla.633>
- Zhang, X., Ling, C., & Qi, L. Q. (2012). The best rank-1 approximation of a symmetric tensor and related spherical optimization problems. *SIAM J. Matrix Analysis and Applications*, 33, 806-821. <http://dx.doi.org/10.1137/110835335>
- Zhou, G. L., Caccetta, L., Teo, K. L., & Wu, S. Y. (2012). Nonnegative polynomial optimization over unit spheres and convex programming relaxations. *SIAM Journal on Optimization*, 22, 987-1008. <http://dx.doi.org/10.1137/110827910>

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).