

## NONNEGATIVE POLYNOMIAL OPTIMIZATION OVER UNIT SPHERES AND CONVEX PROGRAMMING RELAXATIONS\*

GUANGLU ZHOU<sup>†</sup>, LOUIS CACCETTA<sup>‡</sup>, KOK LAY TEO<sup>‡</sup>, AND SOON-YI WU<sup>§</sup>

**Abstract.** We consider approximation algorithms for nonnegative polynomial optimization over unit spheres. Such optimization models have wide applications, e.g., in signal and image processing, high order statistics, and computer vision. Since polynomial functions are nonconvex, the problems under consideration are all NP-hard. In this paper, based on convex polynomial optimization relaxations, we propose polynomial-time approximation algorithms with new approximation bounds. Numerical results are reported to show the effectiveness of the proposed approximation algorithms.

**Key words.** nonnegative polynomial optimization, convex optimization relaxation, approximation algorithm

**AMS subject classifications.** 90C33, 90C30, 65H10

**DOI.** 10.1137/110827910

**1. Introduction.** Maximizing (or minimizing) a polynomial function, subject to some suitable polynomial constraints, is a fundamental model in optimization. Such an optimization model has wide applications, e.g., in signal processing, speech recognition, biomedical engineering, material science, investment science, quantum mechanics, and numerical linear algebra; see [10, 14, 17, 18, 20, 22, 24, 32, 33, 34] for details. In this paper, we consider nonnegative polynomial optimization over unit spheres which is a special case of the optimization problem considered in [14, 22].

Let  $\Re$  be the real field and let  $\Re_+^n$  be the nonnegative cone in  $\Re^n$ , that is, the subset of vectors with nonnegative coordinates. The interior of  $\Re_+^n$ , consisting of vectors with positive coordinates, will be denoted by  $\Re_{++}^n$ . A *generalized polynomial* is a function  $P : \Re_{++}^n \rightarrow \Re$  of the form

$$(1.1) \quad P(z) = \sum_{\alpha} c_{\alpha} z^{\alpha},$$

where  $\alpha$  ranges over a finite set of  $\Re_+^n$ , and let  $z^{\alpha}$  stand for  $z_1^{\alpha_1} \dots z_n^{\alpha_n}$  with  $\alpha = (\alpha_1, \dots, \alpha_n)$ . The degree of  $P$  is  $h = \max_{\alpha} h_{\alpha}$ , where  $h_{\alpha} = \sum_i \alpha_i$ . We say that  $P$  has nonnegative coefficients if  $c_{\alpha} \geq 0$  for all  $\alpha$ . If  $P$  has nonnegative coefficients,  $P$  is said to be a nonnegative polynomial.

A  $d$ th order *tensor*  $\mathcal{A}$  is defined as

$$(1.2) \quad \mathcal{A} = (a_{i_1 i_2 \dots i_d}), \quad a_{i_1 i_2 \dots i_d} \in \Re, 1 \leq i_1 \leq n_1, 1 \leq i_2 \leq n_2, \dots, 1 \leq i_d \leq n_d.$$

\*Received by the editors March 18, 2011; accepted for publication (in revised form) April 11, 2012; published electronically August 15, 2012.

<http://www.siam.org/journals/siopt/22-3/82791.html>

<sup>†</sup>Corresponding author. Department of Mathematics and Statistics, Curtin University, Perth, Australia (G.Zhou@curtin.edu.au). This author's work was supported by the Australian Research Council and the National Natural Science Foundation of China (grant 11171180).

<sup>‡</sup>Department of Mathematics and Statistics, Curtin University, Perth, Australia (caccetta@maths.curtin.edu.au, K.L.Teo@curtin.edu.au).

<sup>§</sup>Institute of Applied Mathematics, National Cheng-Kung University, Tainan, Taiwan (soonyi@mail.ncku.edu.tw).

$\mathcal{A}$  is called nonnegative (or, respectively, positive) if  $a_{i_1 i_2 \dots i_d} \geq 0$  (or, respectively,  $a_{i_1 i_2 \dots i_d} > 0$ ). Let  $F$  be the following multilinear function defined by tensor  $\mathcal{A}$ :

$$(1.3) \quad F_{\mathcal{A}}(x^1, x^2, \dots, x^d) := \sum_{1 \leq i_1 \leq n_1, 1 \leq i_2 \leq n_2, \dots, 1 \leq i_d \leq n_d} a_{i_1 i_2 \dots i_d} x_{i_1}^1 x_{i_2}^2 \cdots x_{i_d}^d,$$

where  $x^i \in \mathbb{R}^{n_i}, 1 \leq i \leq d$ . A  $d$ th order  $n$ -dimensional square tensor  $\mathcal{B}$  consists of  $n^d$  entries in  $\mathbb{R}$ , which is defined as

$$(1.4) \quad \mathcal{B} = (b_{i_1 i_2 \dots i_d}), \quad b_{i_1 i_2 \dots i_d} \in \mathbb{R}, \quad 1 \leq i_1, i_2, \dots, i_d \leq n.$$

Tensor  $\mathcal{B}$  is called symmetric if its entries  $a_{i_1 i_2 \dots i_d}$  are invariant under any permutation of their indices  $\{i_1, i_2, \dots, i_d\}$  [30]. Let  $f(x), x \in \mathbb{R}^n$ , a  $d$ th degree homogeneous polynomial form of  $n$  variables, be defined by

$$(1.5) \quad f_{\mathcal{B}}(x) := \sum_{i_1, i_2, \dots, i_d=1}^n b_{i_1 i_2 \dots i_d} x_{i_1} x_{i_2} \cdots x_{i_d}.$$

Assume that  $p, q, m$ , and  $n$  are positive integers,  $m, n \geq 2$ , and  $d = p + q$ . A  $(p, q)$ th order  $(m, n)$ -dimensional rectangular tensor is defined as

$$(1.6) \quad \mathcal{C} = (c_{i_1 \dots i_p j_1 \dots j_q}), \quad c_{i_1 \dots i_p j_1 \dots j_q} \in \mathbb{R}, \\ i_k = 1, \dots, n, \quad k = 1, \dots, p, \quad \text{and} \quad j_l = 1, \dots, m, \quad l = 1, \dots, q.$$

We say that  $\mathcal{C}$  is a partially symmetric rectangular tensor [8] if  $c_{i_1 \dots i_p j_1 \dots j_q}$  is invariant under any permutation of indices among  $i_1, \dots, i_p$ , and any permutation of indices among  $j_1, \dots, j_q$ , i.e.,

$$c_{\pi(i_1 \dots i_p) \sigma(j_1 \dots j_q)} = c_{i_1 \dots i_p j_1 \dots j_q}, \quad \pi \in S_p, \quad \sigma \in S_q,$$

where  $S_r$  is the permutation group of  $r$  indices. Let

$$(1.7) \quad G_{\mathcal{C}}(x, y) := \sum_{i_1, \dots, i_p=1}^n \sum_{j_1, \dots, j_q=1}^m c_{i_1 \dots i_p j_1 \dots j_q} x_{i_1} \cdots x_{i_p} y_{j_1} \cdots y_{j_q}, \quad x \in \mathbb{R}^n, \quad y \in \mathbb{R}^m.$$

When  $p = q = 1$ , this is simply a bilinear form of  $x$  and  $y$ .

Suppose  $\mathcal{A}, \mathcal{B}$ , and  $\mathcal{C}$  are all nonnegative tensors. In this paper, we shall study optimization of a nonnegative polynomial function subject to spherical constraints. To be specific, we consider the following models:

$$(P^1) \quad \begin{aligned} & \max f_{\mathcal{B}}(x) \\ & \text{s.t. } \|x\| = 1, x \in \mathbb{R}^n; \end{aligned}$$

$$(P^2) \quad \begin{aligned} & \max G_{\mathcal{C}}(x, y) \\ & \text{s.t. } \|x\| = 1, x \in \mathbb{R}^n, \\ & \quad \|y\| = 1, y \in \mathbb{R}^m; \end{aligned}$$

$$(P^d) \quad \begin{aligned} & \max F_{\mathcal{A}}(x^1, x^2, \dots, x^d) \\ & \text{s.t. } \|x^i\| = 1, x^i \in \mathbb{R}^{n_i}, \quad i = 1, 2, \dots, d. \end{aligned}$$

These models arise from the best rank-one approximation problem for nonnegative tensors [2, 9, 35] which has wide applications in signal and image processing, statistics,

TABLE 1

New approximation results. Note that the approximation bounds in the third column are obtained in [14, 22] for the general cases of  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$ .

Model	New approximation bound	Approximation bound [14, 22]
$(P^1)$	$n^{-\frac{d-2}{2}}$	$d!d^{-d}n^{-\frac{d-2}{2}}$ [14]
$(P^2)$	$(n^p m^q)^{-\frac{p+q-2}{2(p+q)}}$	
$(P^2)$	$\frac{1}{\sqrt{mn}}$ ( $p = q = 2$ )	$\frac{1}{2^{\max\{m,n\}^2}}$ ( $p = q = 2$ ) [22]
$(P^d)$	$(n_1 n_2 \cdots n_d)^{-\frac{d-2}{2d}}$	$(n_1 n_2 \cdots n_{d-2})^{-\frac{1}{2}}$ [14]

and computer vision. These models also have links with higher-order Markov chains [29] and spectral hypergraph theory [5, 15].

In this paper we shall focus on polynomial-time approximation algorithms for the NP-hard problems  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$ . A quality measure of approximation is defined as follows.

DEFINITION 1.1 (see [22]). *Suppose the optimization problem*

$$(P) \quad \begin{aligned} &\max g(x) \\ &\text{s.t. } x \in \Omega \subseteq \mathbb{R}^n \end{aligned}$$

is NP-hard. Let  $\mathfrak{S}$  be a polynomial-time approximation algorithm to solve  $(P)$ .  $\mathfrak{S}$  is said to have a relative approximation bound  $C \in (0, 1]$  if, for any instance of  $(P)$ , the algorithm  $\mathfrak{S}$  can find a lower bound  $g$  for  $(P)$  such that

$$g_{\max} \geq g \geq Cg_{\max},$$

where  $g_{\max}$  is the maximum value of the instance of  $(P)$ .

In this definition, the closer  $C$  is to 1, the better the approximation algorithm would be. Recently, it has been proved in [14, 22] that the general cases of  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$  are NP-hard when  $d > 2$ . Furthermore, by using semidefinite programming (SDP) relaxations, some approximation methods have been proposed in [14, 22], and approximation bounds for these approximation methods have been derived.

**Contributions.** In section 2, we show that  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$  are NP-hard. Furthermore, we show that these NP-hard optimization problems can be solved approximately by using convex optimization relaxations and their approximation bounds are analyzed. Table 1 summarizes our new approximation bounds obtained in this paper. In addition, unlike the SDP relaxations in [14, 22], the convex optimization relaxations used in this paper have the same size as the original problems. This indicates that our proposed approximation algorithms may be used to solve large size problems.

In section 3, we propose some practical computational methods for solving the convex optimization relaxations. In particular, we show that these convex optimization relaxations can be reformulated as geometric programming (GP) problems which are extensively studied in [3, 4]. The standard barrier-based interior-point method for convex optimization can be applied to GP with a worst-case polynomial-time complexity; see [4]. Additionally, we present a power method (PM) and a smoothing Newton method (SNM) for these convex optimization relaxations.

In section 4, we report our numerical results on the testing of the efficiency of the proposed approximation methods. Tables 4 and 5 show that our numerical approximation ratios for randomly generated test problems are close to 1, indicating that a

high quality solution can be obtained by our proposed approximation methods. We conclude the paper with some remarks in section 5.

We conclude this section with some notation. For  $x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$ , the 2-norm is denoted by  $\|x\|$  and the  $l^d$ -norm is denoted by  $\|x\|_d$ , i.e.,  $\|x\|_d = (\sum_{i=1}^n |x_i|^d)^{\frac{1}{d}}$ . We use  $|x|$  to denote the vector  $[|x_1|, |x_2|, \dots, |x_n|]^T$ , and for any  $\alpha \in \mathbb{R}$ ,  $x^{[\alpha]} = [x_1^\alpha, x_2^\alpha, \dots, x_n^\alpha]^T$ .

**2. Convex optimization relaxations.** In this section, we show that  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$  are NP-hard problems. Furthermore, we show that these NP-hard problems can be relaxed by some *convex polynomial optimization* problems, and their approximation bounds are analyzed.

The NP-hardness of  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$  when  $d > 2$  can be obtained by similar arguments as in [14, 22, 40]. For the completeness of this paper, in this section, we still give the proof for the NP-hardness of  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$  when  $d > 2$ . Let  $G = (V, E)$  be a graph with the set of nodes  $V = \{1, 2, \dots, n\}$  and the set of undirected arcs  $E = \{(i_k, j_k), k = 1, 2, \dots, m\}$ . Denote by  $A$  its adjacency matrix. In [25], Motzkin and Straus established a link between the problem of finding the clique number of  $G$  and the problem of optimizing the Lagrangian of  $G$  over the simplex  $\Delta_n = \{x \in \mathbb{R}_+^n : \sum_{i=1}^n x_i = 1\}$ . Define this optimization problem as follows:

$$(2.1) \quad \text{Find } f^* = \max \left\{ \sum_{(i,j) \in E} x_i x_j : x \in \Delta_n \right\}.$$

It is proved [25] that  $f^* = \frac{1}{2}[1 - \frac{1}{\omega(G)}]$ , where  $\omega(G)$  is the clique number of the graph  $G$ . It is well known that (2.1) is NP-hard since the problem of finding the clique number of  $G$  is NP-hard. In order to show that  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$  are NP-hard, we first give some lemmas in the following.

LEMMA 2.1 (see [27]). *After an appropriate change of variables, problem (2.1) can be posed in any of the following settings:*

1. *Quartic maximization over the Euclidean ball:*

$$(2.2) \quad \max_{u \in \mathbb{R}^n} \left\{ \sum_{k=1}^m (u^T A_k u)^2 : \|u\| = 1 \right\},$$

where  $A_k = \frac{e_{i_k}^T e_{i_k} + e_{j_k}^T e_{j_k}}{\sqrt{2}}$ ,  $(i_k, j_k) \in E, 1 \leq k \leq m$ , and  $e_i, 1 \leq i \leq n$ , are the basis vectors of  $\mathbb{R}^n$ .

2. *Cubic maximization over Euclidean balls:*

$$(2.3) \quad \max_{u \in \mathbb{R}^n, w \in \mathbb{R}^m} \left\{ \sum_{k=1}^m w_k (u^T A_k u) : \|u\| = 1, \|w\| = 1 \right\}.$$

LEMMA 2.2. *Let  $(u^*, w^*)$  be a global solution of (2.3). Then,  $(u^*, u^*, w^*)$  is a global solution of the following optimization problem:*

$$(2.4) \quad \max_{u, v \in \mathbb{R}^n, w \in \mathbb{R}^m} \left\{ \sum_{k=1}^m w_k (u^T A_k v) : \|u\| = 1, \|v\| = 1, \|w\| = 1 \right\}.$$

*Proof.* Let  $f(u, v, w) = \sum_{k=1}^m w_k (u^T A_k v)$ . Since  $A_k, 1 \leq k \leq m$ , are symmetric matrices, there exists a nonnegative third order  $(n, n, m)$ -dimensional tensor

$\mathcal{E} = (e_{rst}), 1 \leq r, s \leq n, 1 \leq t \leq m$ , such that  $f(u, v, w) = \sum_{k=1}^m w_k(u^T A_k v) = \sum_{1 \leq r, s \leq n, 1 \leq t \leq m} e_{rst} u_r v_s w_t$ , and tensor  $\mathcal{E}$  is partially symmetric with respect to the first two indices. By Theorem 2.1 of [40], this lemma is satisfied.  $\square$

It follows from the NP-hardness of (2.1) that (2.2), (2.3), and (2.4) are NP-hard. Since (2.2), (2.3), and (2.4) are a special case of  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$ , respectively, we have the following result.

**THEOREM 2.1.** *Suppose  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{C}$  are all nonnegative tensors and  $d > 2$ . Then,  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$  are NP-hard problems.*

Since  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$  are NP-hard when  $d > 2$  and the objective functions of these problems are nonconvex, it is difficult to find a global solution for these problems. In the following, we will show that these NP-hard problems can be relaxed by some convex polynomial optimization problems. We first give some lemmas which will be used later.

**LEMMA 2.3.** *Suppose that  $x^* \in \Re^n$  is a global solution of  $(P^1)$ . Then,  $|x^*| = [|x_1^*|, |x_2^*|, \dots, |x_n^*|]^T$  is a global solution of  $(P^1)$ .*

*Proof.* Since  $x^* \in \Re^n$  is a global solution of  $(P^1)$ , we have  $\|x^*\| = 1$  and  $f_{\mathcal{B}}(x^*) \geq f_{\mathcal{B}}(x)$  for any  $x \in \Re^n$  satisfying  $\|x\| = 1$ . Because the  $l^2$ -norm of  $|x^*|$  is 1, we have  $f_{\mathcal{B}}(|x^*|) \leq f_{\mathcal{B}}(x^*)$ . Since  $|x_i^*| \geq x_i^*, i = 1, 2, \dots, n$ , and  $\mathcal{B}$  is a nonnegative tensor, we obtain  $f_{\mathcal{B}}(|x^*|) \geq f_{\mathcal{B}}(x^*)$ . Hence,  $f_{\mathcal{B}}(|x^*|) = f_{\mathcal{B}}(x^*)$ , which implies that  $|x^*|$  is a global solution of  $(P^1)$ .  $\square$

From Lemma 2.3, we have the following result.

**LEMMA 2.4.** *Suppose that  $x^* \in \Re^n$  is a global solution of the following problem:*

$$(P_+^1) \quad \max f_{\mathcal{B}}(x) \\ \text{s.t. } \|x\| \leq 1, x \geq 0, x \in \Re^n.$$

*Then,  $x^*$  is a global solution of  $(P^1)$ .*

*Proof.* By Lemma 2.3, if  $x^* \in \Re^n$  is a global solution of the following problem:

$$(P_{\geq}^1) \quad \max f_{\mathcal{B}}(x) \\ \text{s.t. } \|x\| = 1, x \geq 0, x \in \Re^n,$$

then  $x^*$  is a global solution of  $(P^1)$ . Since  $\mathcal{B}$  is a nonnegative tensor, we can readily prove that  $x^* \in \Re^n$  is a global solution of  $(P_{\geq}^1)$  if and only if  $x^* \in \Re^n$  is a global solution of  $(P_+^1)$ . Hence, this lemma holds.  $\square$

Similarly, we have the following results for  $(P^2)$  and  $(P^d)$ .

**LEMMA 2.5.** *Suppose that  $(x^*, y^*)$  is a global solution of the following problem:*

$$(P_+^2) \quad \max G_{\mathcal{C}}(x, y) \\ \text{s.t. } \|x\| \leq 1, x \geq 0, x \in \Re^n, \\ \|y\| \leq 1, y \geq 0, y \in \Re^m.$$

*Then,  $(x^*, y^*)$  is a global solution of  $(P^2)$ .*

**LEMMA 2.6.** *Suppose that  $((x^1)^*, \dots, (x^d)^*)$  is a global solution of the following problem:*

$$(P_+^d) \quad \max F_{\mathcal{A}}(x^1, x^2, \dots, x^d) \\ \text{s.t. } \|x^i\| \leq 1, x^i \geq 0, x^i \in \Re^{n_i}, \quad i = 1, 2, \dots, d.$$

*Here,  $d > 2$ . Then,  $((x^1)^*, \dots, (x^d)^*)$  is a global solution of  $(P^d)$ .*

Let  $\|x\|_d$  be the  $l^d$ -norm of  $x \in \mathfrak{R}^l$ , i.e.,  $\|x\|_d = (\sum_{i=1}^l |x_i|^d)^{\frac{1}{d}}$ . We first consider a relaxation of  $(P_+^d)$ . By relaxing the constraints  $\|x^i\| \leq 1, i = 1, 2, \dots, d$ , to  $\|x^i\|_d \leq 1, i = 1, 2, \dots, d$ , the problem  $(P_+^d)$  can be relaxed to

$$(2.5) \quad \begin{aligned} (\bar{P}_+^d) \quad & \max F_{\mathcal{A}}(x^1, x^2, \dots, x^d) \\ & \text{s.t. } \|x^i\|_d \leq 1, x^i \geq 0, x^i \in \mathfrak{R}^{n_i}, \quad i = 1, 2, \dots, d. \end{aligned}$$

In order to analyze the approximation bounds, we give the following lemma which is crucial for our analysis.

LEMMA 2.7. *Suppose that  $x \in \mathfrak{R}^l$  satisfying  $x \geq 0$  and  $\|x\|_d = 1$ . Then,*

$$\|x\| \leq l^{\frac{d-2}{2d}}.$$

*Proof.* Consider the following optimization problem:

$$\begin{aligned} & \max \|x\| \\ & \text{s.t. } \|x\|_d = 1, x \geq 0, x \in \mathfrak{R}^l. \end{aligned}$$

By simple computation, the optimal solution of the above problem is  $x^* = [l^{-\frac{1}{d}}, \dots, l^{-\frac{1}{d}}]^T$  and  $\|x^*\| = l^{\frac{d-2}{2d}}$ . Hence this lemma holds.  $\square$

Let

$$\begin{aligned} \Omega^d &= \{(x^1, \dots, x^d) : \|x^i\| \leq 1, x^i \geq 0, x^i \in \mathfrak{R}^{n_i}, i = 1, 2, \dots, d\}, \\ \bar{\Omega}^d &= \{(x^1, \dots, x^d) : \|x^i\|_d \leq 1, x^i \geq 0, x^i \in \mathfrak{R}^{n_i}, i = 1, 2, \dots, d\}. \end{aligned}$$

Clearly,  $\Omega^d \subseteq \bar{\Omega}^d$ . We have the following result about  $(P_+^d)$  and its relaxation  $(\bar{P}_+^d)$ .

THEOREM 2.2. *Suppose that  $(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^d)$  is a global solution of  $(P_+^d)$  and  $(\hat{x}^1, \hat{x}^2, \dots, \hat{x}^d)$  is a global solution of  $(\bar{P}_+^d)$ , respectively. Let  $((x^1)^*, \dots, (x^d)^*) = (\frac{\hat{x}^1}{\|\hat{x}^1\|}, \dots, \frac{\hat{x}^d}{\|\hat{x}^d\|})$ . Then,*

$$(2.6) \quad F_{\mathcal{A}}(\hat{x}^1, \hat{x}^2, \dots, \hat{x}^d) \geq F_{\mathcal{A}}(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^d),$$

$$(2.7) \quad F_{\mathcal{A}}(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^d) \geq F_{\mathcal{A}}((x^1)^*, \dots, (x^d)^*) \geq \frac{F_{\mathcal{A}}(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^d)}{(n_1 n_2 \dots n_d)^{\frac{d-2}{2d}}}.$$

*Proof.* Since  $\Omega^d \subseteq \bar{\Omega}^d$ , (2.6) is satisfied. Clearly,  $((x^1)^*, \dots, (x^d)^*) \in \Omega^d$ . Hence,  $F_{\mathcal{A}}(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^d) \geq F_{\mathcal{A}}((x^1)^*, \dots, (x^d)^*)$ . By (2.6) and Lemma 2.7,

$$\begin{aligned} F_{\mathcal{A}}((x^1)^*, \dots, (x^d)^*) &= \frac{F_{\mathcal{A}}(\hat{x}^1, \hat{x}^2, \dots, \hat{x}^d)}{\|\hat{x}^1\| \|\hat{x}^2\| \dots \|\hat{x}^d\|} \\ &\geq (n_1 n_2 \dots n_d)^{-\frac{d-2}{2d}} F_{\mathcal{A}}(\hat{x}^1, \hat{x}^2, \dots, \hat{x}^d) \\ &\geq (n_1 n_2 \dots n_d)^{-\frac{d-2}{2d}} F_{\mathcal{A}}(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^d). \end{aligned}$$

Therefore, this theorem holds.  $\square$

From Theorem 2.2, we note that the optimal value of  $(\bar{P}_+^d)$  is an upper bound of the optimal value of  $(P_+^d)$ .  $((x^1)^*, \dots, (x^d)^*)$  defined in Theorem 2.2 is an approximation solution of  $(P_+^d)$  with the approximation bound  $C = (n_1 n_2 \dots n_d)^{-\frac{d-2}{2d}}$ . In section 3, we will show that a global solution of  $(\bar{P}_+^d)$  can be obtained by a polynomial-time algorithm. Hence, by Lemma 2.6 and Theorem 2.2, we have the following theorem.

**THEOREM 2.3.** *Suppose  $\mathcal{A}$  is a nonnegative tensor. Then, problem  $(P^d)$  can be solved by a polynomial-time approximation algorithm with a relative approximation bound  $C_d = (n_1 n_2 \cdots n_d)^{-\frac{d-2}{2d}}$ . When  $d = 2$ ,  $C_d = 1$ .*

In the following, we will show that solving  $(\bar{P}_+^d)$  is equivalent to solving a convex optimization problem. To this end, we give the following theorem for the generalized polynomial  $P$  defined in (1.1).

**THEOREM 2.4.** *A generalized polynomial  $P$  with nonnegative coefficients of degree at most 1 is concave on  $\mathfrak{R}_+^n$ .*

*Proof.* See Theorem 5.2 of [1] for the proof.  $\square$

Let  $y^i = [(x_1^i)^d, (x_2^i)^d, \dots, (x_{n_i}^i)^d]^T$ ,  $x^i = (y^i)^{[1/d]} = [(y_1^i)^{1/d}, (y_2^i)^{1/d}, \dots, (y_{n_i}^i)^{1/d}]^T$ ,  $i = 1, 2, \dots, d$ . Then,  $(\bar{P}_+^d)$  can be formulated equivalently as the following optimization problem:

$$(2.8) \quad \begin{aligned} (\hat{P}_+^d) \quad & \min -F_{\mathcal{A}}((y^1)^{[1/d]}, (y^2)^{[1/d]}, \dots, (y^d)^{[1/d]}) \\ & \text{s.t. } \sum_{j=1}^{n_i} y_j^i \leq 1, y^i \geq 0, y^i \in \mathfrak{R}^{n_i}, \quad i = 1, 2, \dots, d. \end{aligned}$$

**THEOREM 2.5.** *Suppose that  $\mathcal{A}$  is a nonnegative tensor. Then,  $(\hat{P}_+^d)$  is a convex optimization problem. Moreover, if  $(y^1, y^2, \dots, y^d)$  is a global solution of  $(\hat{P}_+^d)$ , then  $((y^1)^{[1/d]}, (y^2)^{[1/d]}, \dots, (y^d)^{[1/d]})$  is a global solution of  $(\bar{P}_+^d)$ .*

*Proof.* Since  $\mathcal{A}$  is a nonnegative tensor,  $F_{\mathcal{A}}((y^1)^{[1/d]}, (y^2)^{[1/d]}, \dots, (y^d)^{[1/d]})$  is a polynomial function with nonnegative coefficients of degree 1. By Theorem 2.4,  $F_{\mathcal{A}}((y^1)^{[1/d]}, (y^2)^{[1/d]}, \dots, (y^d)^{[1/d]})$  is concave on  $\mathfrak{R}^{n_1 + \dots + n_d}$ . Hence,  $-F_{\mathcal{A}}((y^1)^{[1/d]}, (y^2)^{[1/d]}, \dots, (y^d)^{[1/d]})$  is convex on  $\mathfrak{R}^{n_1 + \dots + n_d}$ . Therefore, this theorem holds.  $\square$

We now move on to consider the relaxations of  $(P_+^1)$  and  $(P_+^2)$ . Like the relaxation  $(\bar{P}_+^d)$  of  $(P_+^d)$ ,  $(P_+^1)$  can be relaxed to

$$(2.9) \quad \begin{aligned} (\bar{P}_+^1) \quad & \max f_{\mathcal{B}}(x) \\ & \text{s.t. } \|x\|_d \leq 1, x \geq 0, x \in \mathfrak{R}^n, \end{aligned}$$

and  $(P_+^2)$  can be relaxed to

$$(2.10) \quad \begin{aligned} (\bar{P}_+^2) \quad & \max G_{\mathcal{C}}(x, y) \\ & \text{s.t. } \|x\|_d \leq 1, x \geq 0, x \in \mathfrak{R}^n, \\ & \quad \|y\|_d \leq 1, y \geq 0, y \in \mathfrak{R}^m, \end{aligned}$$

where  $d = p + q$ . Since  $f_{\mathcal{B}}(x) = F_{\mathcal{B}}(x, x, \dots, x)$  and  $G_{\mathcal{C}}(x, y) = F_{\mathcal{C}}(x, \dots, x, y, \dots, y)$ , we have the following two theorems by using arguments given for Theorems 2.2, 2.3, and 2.5.

**THEOREM 2.6.** *Suppose that  $\mathcal{B}$  is a nonnegative tensor. Then, we have the following results:*

(i) *A global solution of  $(\bar{P}_+^1)$  can be obtained by solving the following convex polynomial optimization problem:*

$$(2.11) \quad \begin{aligned} (\hat{P}_+^1) \quad & \min -f_{\mathcal{B}}(y^{[1/d]}) \\ & \text{s.t. } \sum_{i=1}^n y_i \leq 1, y \geq 0, y \in \mathfrak{R}^n, \end{aligned}$$

where  $y^{[1/d]} = [y_1^{1/d}, y_2^{1/d}, \dots, y_n^{1/d}]^T$ . If  $y^*$  is a global solution of  $(\hat{P}_+^1)$ , then  $x^* = (y^*)^{[1/d]}$  is a global solution of  $(\bar{P}_+^1)$ .

(ii) Suppose that  $\bar{x}$  is a global solution of  $(P_+^1)$  and  $\hat{x}$  is a global solution of  $(\bar{P}_+^1)$ , respectively. Let  $x^* = \frac{\hat{x}}{\|\hat{x}\|}$ . Then,

$$(2.12) \quad f_{\mathcal{B}}(\hat{x}) \geq f_{\mathcal{B}}(\bar{x}) \geq f_{\mathcal{B}}(x^*) \geq n^{-\frac{d-2}{2}} f_{\mathcal{B}}(\bar{x}).$$

(iii) Problem  $(P^1)$  can be solved by a polynomial-time approximation algorithm with a relative approximation bound  $C_1 = n^{-\frac{d-2}{2}}$ . When  $d = 2$ ,  $C_1 = 1$ .

**THEOREM 2.7.** Suppose that  $\mathcal{C}$  is a nonnegative tensor. Then, the following results are valid:

(i) A global solution of  $(\bar{P}_+^2)$  can be obtained by solving the following convex polynomial optimization problem:

$$(2.13) \quad \begin{aligned} (\hat{P}_+^2) \quad & \min -G_{\mathcal{C}}(r^{[1/d]}, s^{[1/d]}) \\ & \text{s.t. } \sum_{i=1}^n r_i \leq 1, r \geq 0, r \in \mathbb{R}^n, \\ & \sum_{j=1}^m s_j \leq 1, s \geq 0, s \in \mathbb{R}^m, \end{aligned}$$

where  $r^{[1/d]} = [r_1^{1/d}, r_2^{1/d}, \dots, r_n^{1/d}]^T$  and  $s^{[1/d]} = [s_1^{1/d}, s_2^{1/d}, \dots, s_m^{1/d}]^T$ . If  $(r^*, s^*)$  is a global solution of  $(\hat{P}_+^2)$ , then  $(x^*, y^*) = ((r^*)^{[1/d]}, (s^*)^{[1/d]})$  is a global solution of  $(\bar{P}_+^2)$ .

(ii) Suppose that  $(\bar{x}, \bar{y})$  is a global solution of  $(P_+^2)$  and  $(\hat{x}, \hat{y})$  is a global solution of  $(\bar{P}_+^2)$ , respectively. Let  $(x^*, y^*) = (\frac{\hat{x}}{\|\hat{x}\|}, \frac{\hat{y}}{\|\hat{y}\|})$ . Then,

$$(2.14) \quad G_{\mathcal{C}}(\hat{x}, \hat{y}) \geq G_{\mathcal{C}}(\bar{x}, \bar{y}) \geq G_{\mathcal{C}}(x^*, y^*) \geq (n^p m^q)^{-\frac{(p+q-2)}{2(p+q)}} G_{\mathcal{C}}(\bar{x}, \bar{y}).$$

(iii) Problem  $(P^2)$  can be solved by a polynomial-time approximation algorithm with a relative approximation bound  $C_2 = (n^p m^q)^{-\frac{(p+q-2)}{2(p+q)}}$ . When  $p = q = 1$ ,  $C_2 = 1$ .

**3. Algorithms for the convex optimization relaxations.** In this section, we will present some algorithms for solving the relaxations  $(\bar{P}_+^1)$ ,  $(\bar{P}_+^2)$ , and  $(\bar{P}_+^d)$  which are defined in (2.9), (2.10), and (2.5), respectively. In particular, in section 3.1, we will give some polynomial-time algorithms for  $(\bar{P}_+^1)$ ,  $(\bar{P}_+^2)$ , and  $(\bar{P}_+^d)$  by reformulating them into GP problems. These GP reformulations have been studied recently in [36]. In section 3.2 we propose some PMs for  $(\bar{P}_+^1)$ ,  $(\bar{P}_+^2)$ , and  $(\bar{P}_+^d)$ , and we present SNMs for these relaxations in section 3.3.

**3.1. Polynomial-time algorithms.** We first give a polynomial-time algorithm for  $(\bar{P}_+^1)$  by reformulating it into a GP problem. If  $\mathcal{B}$  is not a symmetric tensor, we can find a symmetric tensor  $\bar{\mathcal{B}}$  such that  $f_{\mathcal{B}}(x) = f_{\bar{\mathcal{B}}}(x)$ ; see [30]. Hence, in this paper we always assume that  $\mathcal{B}$  is a symmetric tensor. By simple computation, we have

$$(3.1) \quad \nabla f_{\mathcal{B}}(x) = d(\mathcal{B}x^{d-1}),$$

where

$$(3.2) \quad \mathcal{B}x^{d-1} = \left( \sum_{i_2, \dots, i_d=1}^n b_{i \ i_2 \dots i_d} x_{i_2} \cdots x_{i_d} \right)_{1 \leq i \leq n},$$



and

$$(3.3) \quad f_{\mathcal{B}}(x) = x^T (\mathcal{B}x^{d-1}).$$

For nonnegative tensor  $\mathcal{B}$ , we have the following definitions and results.

DEFINITION 3.1 (see [6, 21, 30]). *If there exist a complex number  $\lambda$  and a nonzero complex vector  $x$  such that*

$$(3.4) \quad \mathcal{B}x^{d-1} := \lambda x^{[d-1]},$$

where  $x^{[\alpha]} = [x_1^\alpha, x_2^\alpha, \dots, x_n^\alpha]^T$ , then  $(\lambda, x)$  is called an eigenvalue-eigenvector of  $\mathcal{B}$ .

The spectral radius of  $\mathcal{B}$  is defined as  $\rho(\mathcal{B}) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } \mathcal{B}\}$ .

DEFINITION 3.2 (see [6]). *Tensor  $\mathcal{B}$  is called reducible if there exists a nonempty proper index subset  $I \subset \{1, 2, \dots, n\}$  such that*

$$b_{i_1 i_2 \dots i_d} = 0 \quad \forall i_1 \in I, \quad \forall i_2, \dots, i_d \notin I.$$

If  $\mathcal{B}$  is not reducible, then  $\mathcal{B}$  is called irreducible.

THEOREM 3.1 (see [6]). *If  $\mathcal{B}$  is an irreducible nonnegative tensor, then there exist  $\lambda_0 > 0$  and  $x_0 \in \mathfrak{R}_{++}^n$  such that*

$$(3.5) \quad \mathcal{B}x_0^{d-1} = \lambda_0 x_0^{[d-1]}.$$

Moreover, if  $\lambda$  is an eigenvalue with nonnegative eigenvector, then  $\lambda = \lambda_0$ . If  $\lambda$  is an eigenvalue of  $\mathcal{B}$ , then  $|\lambda| \leq \lambda_0$ . Clearly,  $\rho(\mathcal{B}) = \lambda_0$ .

THEOREM 3.2 (see [6]). *Assume that  $\mathcal{B}$  is an irreducible nonnegative tensor. Then*

$$(3.6) \quad \lambda_0 = \min_{x \in \mathfrak{R}_{++}^n} \max_{1 \leq i \leq n} \frac{(\mathcal{B}x^{d-1})_i}{x_i^{d-1}},$$

where  $\lambda_0$  is the unique positive eigenvalue corresponding to the positive eigenvector.

THEOREM 3.3. *Let  $\mathcal{B}$ ,  $\lambda_0$ , and  $x_0$  be as in Theorem 3.1, and let  $x^* = x_0 / \|x_0\|_d$ . Then,  $x^*$  is a global solution of  $(\bar{P}_+^1)$ .*

*Proof.* From the conditions of this theorem, we have

$$(3.7) \quad \mathcal{B}(x^*)^{d-1} = \lambda_0 (x^*)^{[d-1]}, \quad x^* \in \mathfrak{R}_{++}^n, \quad \text{and } \|x^*\|_d = 1.$$

We consider the following optimization problem:

$$(\bar{P}_+^1) \quad \begin{aligned} & \max f_{\mathcal{B}}(x) \\ & \text{s.t. } \|x\|_d = 1, x \in \mathfrak{R}_+^n. \end{aligned}$$

The optimality conditions of  $(\bar{P}_+^1)$  are as follows: There exists a  $\lambda \in \mathfrak{R}$  such that

$$\mathcal{B}x^{d-1} = \lambda x^{[d-1]}, \quad x \in \mathfrak{R}^n, \quad \text{and } \|x\|_d = 1.$$

So, for any local maximizer  $\bar{x}$  of  $(\bar{P}_+^1)$ , there exists a  $\bar{\lambda} \in \mathfrak{R}$  such that

$$(3.8) \quad \mathcal{B}\bar{x}^{d-1} = \bar{\lambda} \bar{x}^{[d-1]}, \quad \bar{x} \in \mathfrak{R}^n, \quad \text{and } \|\bar{x}\|_d = 1.$$

Hence, we obtain

$$(3.9) \quad f_{\mathcal{B}}(\bar{x}) = \bar{x}^T (\mathcal{B}\bar{x}^{d-1}) = \bar{x}^T (\bar{\lambda} \bar{x}^{[d-1]}) = \bar{\lambda} \sum_{i=1}^n \bar{x}_i^d = \bar{\lambda}.$$

As for (3.9), by (3.7), we have  $f_{\mathcal{B}}(x^*) = \lambda_0$ . It follows from (3.9) that  $(\bar{\lambda}, \bar{x})$  is an eigenvalue-eigenvector of  $\mathcal{B}$ . So, by Theorem 3.1,  $|\bar{\lambda}| \leq \lambda_0$ . Hence,  $f_{\mathcal{B}}(x^*) \geq |f_{\mathcal{B}}(\bar{x})|$ , which means that  $x^*$  is a global maximizer of  $(\bar{P}_+^1)$ .  $\square$

We now look at how to find  $\lambda_0$  and  $x_0$  in Theorem 3.1. By Theorem 3.2, it is clear that  $\lambda_0$  and  $x_0$  can be obtained by solving the following optimization problem:

$$(3.10) \quad \min_{x \in \mathfrak{R}_{++}^n} \max_{1 \leq i \leq n} \frac{(\mathcal{B}x^{d-1})_i}{x_i^{d-1}}.$$

Let

$$\lambda = \max_{1 \leq i \leq n} \frac{(\mathcal{B}x^{d-1})_i}{x_i^{d-1}}.$$

Then, we have

$$(\mathcal{B}x^{d-1})_i \leq \lambda x_i^{d-1}, \quad i = 1, 2, \dots, n,$$

which can also be written as

$$\sum_{i_2, \dots, i_d=1}^n b_{i \ i_2 \dots i_d} x_{i_2} \cdots x_{i_d} \lambda^{-1} x_i^{1-d} \leq 1, \quad i = 1, 2, \dots, n.$$

Hence, (3.10) can be reformulated into the following problem:

$$(3.11) \quad \begin{aligned} (GP1) \quad & \min \lambda \\ & \text{s.t.} \quad \sum_{i_2, \dots, i_d=1}^n b_{i \ i_2 \dots i_d} x_{i_2} \cdots x_{i_d} \lambda^{-1} x_i^{1-d} \leq 1, \quad i = 1, 2, \dots, n, \\ & x \in \mathfrak{R}_{++}^n. \end{aligned}$$

Problem (GP1) is a geometric program which is extensively studied in [3, 4]. There are at least two major approaches to solving a geometric program using modern convex optimization techniques. One is the interior-point method as in [28], and the other is an infeasible algorithm as in [19]. The standard barrier-based interior-point method for convex optimization can be applied to GP in a straightforward way, with a worst-case polynomial-time complexity; see [4]. User-friendly software for GP is available on the Internet, such as the MOSEK package [26] and the GGPLAB package [13].

In the following, as for  $(\bar{P}_+^1)$ , we will reformulate  $(\bar{P}_+^2)$  and  $(\bar{P}_+^d)$  into GP problems, so  $(\bar{P}_+^2)$  and  $(\bar{P}_+^d)$  can also be solved by polynomial-time algorithms.

For a rectangular tensor  $\mathcal{C}$  defined as in (1.6), if  $\mathcal{C}$  is not partially symmetric, we can find a partially symmetric tensor  $\bar{\mathcal{C}}$  such that  $G_{\mathcal{C}}(x, y) = G_{\bar{\mathcal{C}}}(x, y)$ . Hence, we may always assume that  $\mathcal{C}$  is a partially symmetric tensor. For a partially symmetric rectangular tensor  $\mathcal{C}$ , from [8], we have the following definitions and results.

Let  $\mathcal{C}x^{p-1}y^q$  be a vector in  $\mathfrak{R}^n$  such that

$$(\mathcal{C}x^{p-1}y^q)_i = \sum_{i_2, \dots, i_p=1}^n \sum_{j_1, \dots, j_q=1}^m c_{i \ i_2 \dots i_p \ j_1 \dots j_q} x_{i_2} \cdots x_{i_p} y_{j_1} \cdots y_{j_q}, \quad i = 1, 2, \dots, n.$$

Similarly, let  $\mathcal{C}x^p y^{q-1}$  be a vector in  $\mathfrak{R}^m$  such that

$$(\mathcal{C}x^p y^{q-1})_j = \sum_{i_1, \dots, i_p=1}^n \sum_{j_2, \dots, j_q=1}^m c_{i_1 \dots i_p \ j \ j_2 \dots j_q} x_{i_1} \cdots x_{i_p} y_{j_2} \cdots y_{j_q}, \quad j = 1, 2, \dots, m.$$

By simple computation, we have

$$(3.12) \quad \nabla G_C(x, y) = \begin{pmatrix} p\mathcal{C}x^{p-1}y^q \\ q\mathcal{C}x^py^{q-1} \end{pmatrix}$$

and

$$(3.13) \quad G_C(x, y) = x^T(\mathcal{C}x^{p-1}y^q) = y^T(\mathcal{C}x^py^{q-1}).$$

For any  $j = 1, 2, \dots, m$ , let  $\mathcal{C}_{\bullet j} = (c_{i_1 \dots i_p j \dots j})$  be a  $p$ th order  $n$ -dimensional square tensor. For any  $i = 1, 2, \dots, n$ , let  $\mathcal{C}_{i \bullet} = (c_{i \dots i j_1 \dots j_q})$  be a  $q$ th order  $m$ -dimensional square tensor.

DEFINITION 3.3 (see [8]). *A nonnegative rectangular tensor  $\mathcal{C}$  is called irreducible if all the square tensors  $\mathcal{C}_{\bullet j}, j = 1, \dots, m$ , and  $\mathcal{C}_{i \bullet}, i = 1, \dots, n$ , are irreducible.*

DEFINITION 3.4 (see [8]). *If there exist a complex number  $\lambda$  and nonzero complex vectors  $x$  and  $y$  such that*

$$(3.14) \quad \begin{cases} \mathcal{C}x^{p-1}y^q = \lambda x^{[d-1]}, \\ \mathcal{C}x^py^{q-1} = \lambda y^{[d-1]}, \quad d = p + q, \end{cases}$$

then we say that  $\lambda$  is a singular value of  $\mathcal{C}$ , and  $x$  and  $y$  are, respectively, left and right eigenvectors of  $\mathcal{C}$ , associated with the singular value  $\lambda$ .

THEOREM 3.4 (see [8]). *If  $\mathcal{C}$  is an irreducible nonnegative rectangular tensor, then there exist  $\lambda_0 > 0, x_0 \in \mathfrak{R}_{++}^n$ , and  $y_0 \in \mathfrak{R}_{++}^m$  such that*

$$(3.15) \quad \begin{cases} \mathcal{C}x_0^{p-1}y_0^q = \lambda_0 x_0^{[d-1]}, \\ \mathcal{C}x_0^py_0^{q-1} = \lambda_0 y_0^{[d-1]}. \end{cases}$$

Moreover, for all singular values  $\lambda$  of  $\mathcal{C}$ ,  $|\lambda| \leq \lambda_0$ .

Clearly, it follows from this result that  $\lambda_0$  is the largest singular value of  $\mathcal{C}$ .

THEOREM 3.5 (see [8]). *Assume that  $\mathcal{C}$  is an irreducible nonnegative rectangular tensor. Then,*

$$\lambda_0 = \min_{(x,y) \in \mathfrak{R}_{++}^n \times \mathfrak{R}_{++}^m} \max_{i,j} \left( \frac{(\mathcal{C}x^{p-1}y^q)_i}{x_i^{d-1}}, \frac{(\mathcal{C}x^py^{q-1})_j}{y_j^{d-1}} \right),$$

where  $\lambda_0$  is the unique positive singular value corresponding to strongly positive left and right eigenvectors.

THEOREM 3.6. *Let  $\mathcal{C}, \lambda_0, x_0$ , and  $y_0$  be as in Theorem 3.4, and let  $x^* = x_0/\|x_0\|_d$  and  $y^* = y_0/\|y_0\|_d$ . Then,  $(x^*, y^*)$  is a global solution of  $(\bar{P}_+^2)$ .*

*Proof.* From the conditions of this theorem, we have

$$(3.16) \quad \begin{cases} \mathcal{C}(x^*)^{p-1}(y^*)^q = \lambda_0(\|x_0\|_d/\|y_0\|_d)^q(x^*)^{[d-1]}, \\ \mathcal{C}(x^*)^p(y^*)^{q-1} = \lambda_0(\|y_0\|_d/\|x_0\|_d)^p(y^*)^{[d-1]}, \\ \|x^*\|_d = 1, x^* \in \mathfrak{R}_{++}^n, \|y^*\|_d = 1, y^* \in \mathfrak{R}_{++}^m. \end{cases}$$

So we have

$$G_C(x^*, y^*) = (x^*)^T(\mathcal{C}(x^*)^{p-1}(y^*)^q) = \lambda_0(\|x_0\|_d/\|y_0\|_d)^q \sum_{i=1}^n (x^*)_i^d = \lambda_0(\|x_0\|_d/\|y_0\|_d)^q$$

and

$$G_C(x^*, y^*) = (y^*)^T (\mathcal{C}(x^*)^p (y^*)^{q-1}) = \lambda_0 (\|y_0\|_d / \|x_0\|_d)^p \sum_{j=1}^m (y^*)_j^d = \lambda_0 (\|y_0\|_d / \|x_0\|_d)^p.$$

From the above two equalities, we have  $\|x_0\|_d = \|y_0\|_d$ . Hence, we obtain

$$G_C(x^*, y^*) = \lambda_0.$$

We consider the following optimization problem:

$$\begin{aligned} (\bar{P}^2) \quad & \max G_C(x, y) \\ & \text{s.t. } \|x\|_d = 1, x \in \mathfrak{R}^n, \\ & \|y\|_d = 1, y \in \mathfrak{R}^m. \end{aligned}$$

The optimality conditions of  $(\bar{P}^2)$  are as follows: There exist  $\lambda, \mu \in \mathfrak{R}$  such that

$$\begin{cases} \mathcal{C}x^{p-1}y^q = \lambda x^{[d-1]}, \\ \mathcal{C}x^p y^{q-1} = \mu y^{[d-1]}, \\ \|x\|_d = 1, x \in \mathfrak{R}^n, \quad \|y\|_d = 1, y \in \mathfrak{R}^m. \end{cases}$$

So, for any local maximizer  $(\bar{x}, \bar{y})$  of  $(\bar{P}^2)$ , there exist  $\bar{\lambda}, \bar{\mu} \in \mathfrak{R}$  such that

$$(3.17) \quad \begin{cases} \mathcal{C}\bar{x}^{p-1}\bar{y}^q = \bar{\lambda}\bar{x}^{[d-1]}, \\ \mathcal{C}\bar{x}^p\bar{y}^{q-1} = \bar{\mu}\bar{y}^{[d-1]}, \\ \|\bar{x}\|_d = 1, \bar{x} \in \mathfrak{R}^n, \quad \|\bar{y}\|_d = 1, \bar{y} \in \mathfrak{R}^m. \end{cases}$$

Hence, we obtain

$$(3.18) \quad G_C(\bar{x}, \bar{y}) = \bar{x}^T (\mathcal{C}\bar{x}^{p-1}\bar{y}^q) = \bar{x}^T (\bar{\lambda}\bar{x}^{[d-1]}) = \bar{\lambda} \sum_{i=1}^n \bar{x}_i^d = \bar{\lambda}$$

and

$$(3.19) \quad G_C(\bar{x}, \bar{y}) = \bar{y}^T (\mathcal{C}\bar{x}^p\bar{y}^{q-1}) = \bar{y}^T (\bar{\mu}\bar{y}^{[d-1]}) = \bar{\mu} \sum_{i=1}^m \bar{y}_i^d = \bar{\mu}.$$

From the above two equalities, we have  $\bar{\lambda} = \bar{\mu}$ . This means that  $\bar{\lambda}$  is a singular value of  $\mathcal{C}$  corresponding to the left and right eigenvectors  $\bar{x}$  and  $\bar{y}$ , respectively. By Theorem 3.4,  $|\bar{\lambda}| \leq \lambda_0$ . Hence,  $G_C(x^*, y^*) \geq |G_C(\bar{x}, \bar{y})|$ , which means that  $(x^*, y^*)$  is a global maximizer of  $(\bar{P}^2)$ . Since  $x^* \in \mathfrak{R}_{++}^n$  and  $y^* \in \mathfrak{R}_{++}^m$ ,  $(x^*, y^*)$  is a global maximizer of  $(\bar{P}_+^2)$ .  $\square$

By Theorem 3.5,  $\lambda_0$ ,  $x_0$ , and  $y_0$  can be obtained by solving the following optimization problem:

$$(3.20) \quad \min_{(x,y) \in \mathfrak{R}_{++}^n \times \mathfrak{R}_{++}^m} \max_{i,j} \left( \frac{(\mathcal{C}x^{p-1}y^q)_i}{x_i^{d-1}}, \frac{(\mathcal{C}x^p y^{q-1})_j}{y_j^{d-1}} \right).$$

Let

$$\lambda = \max_{i,j} \left( \frac{(\mathcal{C}x^{p-1}y^q)_i}{x_i^{d-1}}, \frac{(\mathcal{C}x^p y^{q-1})_j}{y_j^{d-1}} \right).$$

Then, we have

$$\begin{aligned} (\mathcal{C}x^{p-1}y^q)_i &\leq \lambda x_i^{d-1}, \quad i = 1, 2, \dots, n, \\ (\mathcal{C}x^p y^{q-1})_j &\leq \lambda y_j^{d-1}, \quad j = 1, 2, \dots, m, \end{aligned}$$

which can also be written as

$$\begin{aligned} \sum_{i_2, \dots, i_p=1}^n \sum_{j_1, \dots, j_q=1}^m c_{i_2 \dots i_p j_1 \dots j_q} x_{i_2} \cdots x_{i_p} y_{j_1} \cdots y_{j_q} \lambda^{-1} x_i^{1-d} &\leq 1, \quad i = 1, 2, \dots, n, \\ \sum_{i_1, \dots, i_p=1}^n \sum_{j_2, \dots, j_q=1}^m c_{i_1 \dots i_p j_2 \dots j_q} x_{i_1} \cdots x_{i_p} y_{j_2} \cdots y_{j_q} \lambda^{-1} y_j^{1-d} &\leq 1, \quad j = 1, 2, \dots, m. \end{aligned}$$

Hence, problem (3.20) can be reformulated into the following GP problem:

$$\begin{aligned} (GP2) \quad &\min \lambda \\ &\text{s.t.} \quad \sum_{i_2, \dots, i_p=1}^n \sum_{j_1, \dots, j_q=1}^m c_{i_2 \dots i_p j_1 \dots j_q} x_{i_2} \cdots x_{i_p} y_{j_1} \cdots y_{j_q} \lambda^{-1} x_i^{1-d} \\ &\leq 1, \quad i = 1, 2, \dots, n, \\ &\sum_{i_1, \dots, i_p=1}^n \sum_{j_2, \dots, j_q=1}^m c_{i_1 \dots i_p j_2 \dots j_q} x_{i_1} \cdots x_{i_p} y_{j_2} \cdots y_{j_q} \lambda^{-1} y_j^{1-d} \\ &\leq 1, \quad j = 1, 2, \dots, m, \\ &x \in \mathfrak{R}_{++}^n, y \in \mathfrak{R}_{++}^m. \end{aligned}$$

We now move on to consider problem  $(\bar{P}_+^d)$ . For the nonnegative tensor  $\mathcal{A}$  defined in (1.2), we have, from [12], the following definitions and results. Let  $\nabla_i F_{\mathcal{A}}$  be a vector in  $\mathfrak{R}^{n_i}$ ,  $i = 1, 2, \dots, d$ , such that

$$\begin{aligned} &\nabla_i F_{\mathcal{A}}(x^1, x^2, \dots, x^d) \\ &= \left( \sum_{\substack{1 \leq j_k \leq n_1, \\ k \in \{1, 2, \dots, d\} \setminus \{i\}}} a_{j_1 \dots j_{i-1} j_{i+1} \dots j_d} x_{j_1}^1 \cdots x_{j_{i-1}}^{i-1} x_{j_{i+1}}^{i+1} \cdots x_{j_d}^d \right)_{1 \leq j_i \leq n_i}. \end{aligned}$$

By simple computation, we have

$$(3.21) \quad \nabla F_{\mathcal{A}}(x^1, x^2, \dots, x^d) = \begin{pmatrix} \nabla_1 F_{\mathcal{A}}(x^1, x^2, \dots, x^d) \\ \nabla_2 F_{\mathcal{A}}(x^1, x^2, \dots, x^d) \\ \vdots \\ \nabla_d F_{\mathcal{A}}(x^1, x^2, \dots, x^d) \end{pmatrix}$$

and

$$(3.22) \quad \begin{aligned} F_{\mathcal{A}}(x^1, x^2, \dots, x^d) &= (x^1)^T [\nabla_1 F_{\mathcal{A}}(x^1, x^2, \dots, x^d)] \\ &= (x^2)^T [\nabla_2 F_{\mathcal{A}}(x^1, x^2, \dots, x^d)] \\ &\quad \vdots \\ &= (x^d)^T [\nabla_d F_{\mathcal{A}}(x^1, x^2, \dots, x^d)]. \end{aligned}$$

The *weak irreducibility* for tensor  $\mathcal{A}$  can be defined by requiring a graph associated with the tensor  $\mathcal{A}$  to be connected; see [12]. The tensor  $\mathcal{A}$  is associated with an undirected  $d$ -partite graph  $\mathcal{G}(\mathcal{A}) = (V, E(\mathcal{A}))$ , the vertex set of which is the disjoint union  $V = \cup_{j=1}^d V_j$ , with  $V_j = \{v_1^j, v_2^j, \dots, v_{n_j}^j\}, j = 1, 2, \dots, d$ . The edge  $(v_{i_k}^k, v_{i_l}^l) \in V_k \times V_l, k \neq l$  belongs to  $E(\mathcal{A})$  if and only if  $a_{i_1 i_2 \dots i_d} > 0$  for some  $d - 2$  indices  $\{i_1, \dots, i_d\} \setminus \{i_k, i_l\}$ . The tensor  $\mathcal{A}$  is *weakly irreducible* if the graph  $\mathcal{G}(\mathcal{A})$  is connected. We have the following theorems.

**THEOREM 3.7.** *If  $\mathcal{A}$  is a weakly irreducible nonnegative tensor, then there exist  $\lambda_0 > 0, x^1 \in \mathbb{R}_{++}^{n_1}, \dots, x^d \in \mathbb{R}_{++}^{n_d}$ , such that*

$$(3.23) \quad \nabla_i F_{\mathcal{A}}(x^1, x^2, \dots, x^d) = \lambda_0 (x^i)^{[d-1]}, \quad i = 1, 2, \dots, d.$$

Moreover, if there are a complex number  $\sigma$  and nonzero complex vectors  $v^1, v^2, \dots, v^d$  such that  $\nabla_i F_{\mathcal{A}}(v^1, v^2, \dots, v^d) = \sigma (v^i)^{[d-1]}, i = 1, 2, \dots, d$ , then  $|\sigma| \leq \lambda_0$ .

*Proof.* By Theorem 4.1 and Corollary 4.3 in [12], this theorem holds. □

**THEOREM 3.8.** *Let  $\mathcal{A}$  and  $\lambda_0$  be as in Theorem 3.7. Then,*

$$\lambda_0 = \min_{(x^1, \dots, x^d) \in \mathbb{R}_{++}^{n_1} \times \dots \times \mathbb{R}_{++}^{n_d}} \max_{1 \leq i \leq d} \left( \frac{(\nabla_i F_{\mathcal{A}}(x^1, x^2, \dots, x^d))_j}{(x_j^i)^{d-1}}, j = 1, 2, \dots, n_i \right).$$

*Proof.* By Corollary 4.2 of [12], this theorem holds. □

**THEOREM 3.9.** *Let  $\mathcal{A}, \lambda_0$ , and  $x^1, \dots, x^d$  be as in Theorem 3.7, and let  $x^{i*} = x^i / \|x^i\|_d, i = 1, 2, \dots, d$ . Then,  $(x^{1*}, \dots, x^{d*})$  is a global solution of  $(\bar{P}_+^d)$ .*

*Proof.* By a similar argument as in Theorem 3.6, this theorem holds. □

By Theorem 3.8,  $\lambda_0$  and  $(x^1, \dots, x^d)$  can be obtained by solving the following optimization problem:

$$(3.24) \quad \min_{(x^1, \dots, x^d) \in \mathbb{R}_{++}^{n_1} \times \dots \times \mathbb{R}_{++}^{n_d}} \max_{1 \leq i \leq d} \left( \frac{(\nabla_i F_{\mathcal{A}}(x^1, x^2, \dots, x^d))_j}{(x_j^i)^{d-1}}, j = 1, 2, \dots, n_i \right).$$

Let

$$\lambda = \max_{1 \leq i \leq d} \left( \frac{(\nabla_i F_{\mathcal{A}}(x^1, x^2, \dots, x^d))_j}{(x_j^i)^{d-1}}, j = 1, 2, \dots, n_i \right).$$

Then, we have

$$(\nabla_i F_{\mathcal{A}}(x^1, x^2, \dots, x^d))_j \leq \lambda (x_j^i)^{d-1}, \quad i = 1, 2, \dots, d, j = 1, \dots, n_i,$$

which can also be written as

$$\sum_{\substack{1 \leq j_k \leq n_k, \\ k \in \{1, \dots, d\} \setminus \{i\}}} a_{j_1 \dots j_{i-1} j j_{i+1} \dots j_d} x_{j_1}^1 \dots x_{j_{i-1}}^{i-1} x_{j_{i+1}}^{i+1} \dots x_{j_d}^d \lambda^{-1} (x_j^i)^{1-d} \leq 1, \\ i = 1, 2, \dots, n, j = 1, \dots, n_i.$$

Hence, problem (3.24) can be reformulated into the following GP problem:

$$(GPd) \quad \min \lambda \\ \text{s.t.} \quad \sum_{\substack{1 \leq j_k \leq n_k, \\ k \in \{1, \dots, d\} \setminus \{i\}}} a_{j_1 \dots j_{i-1} j j_{i+1} \dots j_d} x_{j_1}^1 \dots x_{j_{i-1}}^{i-1} x_{j_{i+1}}^{i+1} \dots x_{j_d}^d \lambda^{-1} (x_j^i)^{1-d} \leq 1, \\ i = 1, 2, \dots, n, j = 1, \dots, n_i.$$

**3.2. Power methods.** In this subsection, we present PMs for solving the relaxations  $(\bar{P}_+^1)$ ,  $(\bar{P}_+^2)$ , and  $(\bar{P}_+^d)$ . We will first give a PM for  $(\bar{P}_+^1)$ . Then, we will propose PMs for  $(\bar{P}_+^2)$  and  $(\bar{P}_+^d)$ .

Let  $\mathcal{B}$ ,  $\lambda_0$ , and  $x_0$  be as in Theorem 3.1, and let  $x^* = x_0/\|x_0\|_d$ . By Theorem 3.3,  $x^*$  is a global solution of  $(\bar{P}_+^1)$ .  $\lambda_0$  and  $x_0$  can be computed by the following power algorithm.

ALGORITHM 3.1.

**Step 0.** Choose  $x^{(1)} \in \mathfrak{R}_{++}^n$ , and set  $k := 1$ .

**Step 1.** Compute

$$\begin{aligned} g^{(k)} &= \nabla f_{\mathcal{B}}(x^{(k)}), \\ \underline{\lambda}_k &= \min_{x_i^{(k)} > 0} \frac{g_i^{(k)}}{(x_i^{(k)})^{d-1}}, \\ \bar{\lambda}_k &= \max_{x_i^{(k)} > 0} \frac{g_i^{(k)}}{(x_i^{(k)})^{d-1}}, \\ x^{(k+1)} &= \frac{(g^{(k)})^{[\frac{1}{d-1}]}}{\left\| (g^{(k)})^{[\frac{1}{d-1}]}\right\|_d}. \end{aligned}$$

**Step 2.** If  $\bar{\lambda}_k = \underline{\lambda}_k$ , then stop. Otherwise, replace  $k$  by  $k + 1$  and go to Step 1.

Algorithm 3.1 has been studied recently in [7, 23, 29, 38, 39]. In particular, the convergence of this power algorithm for *primitive* nonnegative tensors has been established in [7]. In [23, 39], an updated version of this algorithm is proposed, and it has been proved that the updated algorithm is always convergent for any irreducible nonnegative tensors. The linear convergence results have been given in [38, 39].

Let  $\mathcal{C}$ ,  $\lambda_0$ ,  $x_0$ , and  $y_0$  be as in Theorem 3.4, and let  $x^* = x_0/\|x_0\|_d$ , where  $d = p+q$  and  $y^* = y_0/\|y_0\|_d$ . Then, by Theorem 3.6,  $(x^*, y^*)$  is a global solution of  $(\bar{P}_+^2)$ . As for Algorithm 3.1,  $\lambda_0$ ,  $x_0$ , and  $y_0$  can be obtained by the following power algorithm.

ALGORITHM 3.2.

**Step 0.** Choose  $x^{(1)} \in \mathfrak{R}_{++}^n$ ,  $y^{(1)} \in \mathfrak{R}_{++}^m$ , and set  $k := 1$ .

**Step 1.** Compute

$$\begin{aligned} \xi^{(k)} &= \nabla_x G_{\mathcal{C}}(x^{(k)}, y^{(k)}), \\ \eta^{(k)} &= \nabla_y G_{\mathcal{C}}(x^{(k)}, y^{(k)}), \\ \underline{\mu}_k &= \min_{x_i^{(k)} > 0, y_j^{(k)} > 0} \left\{ \frac{\xi_i^{(k)}}{(x_i^{(k)})^{d-1}}, \frac{\eta_j^{(k)}}{(y_j^{(k)})^{d-1}} \right\}, \\ \bar{\mu}_k &= \max_{x_i^{(k)} > 0, y_j^{(k)} > 0} \left\{ \frac{\xi_i^{(k)}}{(x_i^{(k)})^{d-1}}, \frac{\eta_j^{(k)}}{(y_j^{(k)})^{d-1}} \right\}. \end{aligned}$$

**Step 2.** If  $\bar{\mu}_k = \underline{\mu}_k$ , then stop. Otherwise, compute

$$x^{(k+1)} = \frac{(\xi^{(k)})^{[\frac{1}{d-1}]}}{\left\| (\xi^{(k)}, \eta^{(k)})^{[\frac{1}{d-1}]} \right\|},$$

$$y^{(k+1)} = \frac{(\eta^{(k)})^{[\frac{1}{d-1}]}}{\left\| (\xi^{(k)}, \eta^{(k)})^{[\frac{1}{d-1}]} \right\|},$$

and replace  $k$  by  $k + 1$  and go to Step 1.

Algorithm 3.2 has been proposed recently in [8]. An updated version of this algorithm is presented in [41], and it is proved that the updated algorithm is always convergent for any irreducible nonnegative tensor  $\mathcal{B}$ .

We now move on to present a PM for  $(\bar{P}_+^d)$ . Let  $\mathcal{A}$ ,  $\lambda_0$ , and  $(x^1, \dots, x^d)$  be as in Theorem 3.7, and let  $x^{i*} = x^i / \|x^i\|_d, i = 1, 2, \dots, d$ . Then, by Theorem 3.9,  $(x^{1*}, \dots, x^{d*})$  is a global solution of  $(\bar{P}_+^d)$ . Recently, in [12], a power algorithm is proposed for finding  $\lambda_0$ , and  $(x^1, \dots, x^d)$  in Theorem 3.7, and the linear convergence of this algorithm has also been established. We state this algorithm as follows.

ALGORITHM 3.3.

**Step 0.** Choose  $x^{(i,1)} \in \mathfrak{R}_{++}^{n_i}, i = 1, 2, \dots, d$ , and set  $k := 1$ .

**Step 1.** Compute

$$g_i^{(k)} = \nabla_i F_{\mathcal{A}}(x^{(1,k)}, x^{(2,k)}, \dots, x^{(d,k)}), \quad i = 1, \dots, d,$$

$$\underline{\sigma}_k = \min \left\{ \frac{(g_i^{(k)})_j}{(x_j^{(i,k)})^{d-1}} : x_j^{(i,k)} > 0, i = 1, 2, \dots, d, j = 1, \dots, n_i \right\},$$

$$\bar{\sigma}_k = \max \left\{ \frac{(g_i^{(k)})_j}{(x_j^{(i,k)})^{d-1}} : x_j^{(i,k)} > 0, i = 1, 2, \dots, d, j = 1, \dots, n_i \right\}.$$

**Step 2.** If  $\bar{\sigma}_k = \underline{\sigma}_k$ , then stop. Otherwise, compute

$$x^{(i,k+1)} = \frac{(g_i^{(k)})^{[\frac{1}{d-1}]}}{\left\| (g_1^{(k)}, \dots, g_d^{(k)})^{[\frac{1}{d-1}]} \right\|}, \quad i = 1, 2, \dots, d,$$

replace  $k$  by  $k + 1$  and go to Step 1.

**3.3. Smoothing Newton methods.** In this subsection, we will present an SNM for solving the relaxations  $(\bar{P}_+^1), (\bar{P}_+^2)$ , and  $(\bar{P}_+^d)$ . It has been shown in section 2 that solving  $(\bar{P}_+^1), (\bar{P}_+^2)$ , and  $(\bar{P}_+^d)$  is equivalent to solving  $(\hat{P}_+^1), (\hat{P}_+^2)$ , and  $(\hat{P}_+^d)$ , respectively. Since  $(\hat{P}_+^1), (\hat{P}_+^2)$ , and  $(\hat{P}_+^d)$ , defined as in (2.11), (2.13), and (2.8), respectively, are convex optimization problems, they can be solved by many state-of-the-art algorithms, such as interior-point methods [28, 37] and Newton-type methods [11, 31].

In the following, we will present an SNM for solving  $(\hat{P}_+^1)$ . Clearly, solving  $(\hat{P}_+^1)$



is equivalent to solving the following problem:

$$(3.25) \quad \begin{aligned} \min \quad & -f_{\mathcal{B}}(y^{[1/d]}) \\ \text{s.t.} \quad & \sum_{i=1}^n y_i = 1, y \geq 0, y \in \mathfrak{R}^n. \end{aligned}$$

Since (3.25) is a convex optimization problem, solving (3.25) is equivalent to solving the following KKT system of (3.25):

$$(3.26) \quad \begin{aligned} \sum_{i=1}^n y_i - 1 &= 0, \\ -\nabla f_{\mathcal{B}}(y^{[1/d]}) + \beta \mathbf{1} - z &= 0, \quad \mathbf{1} = [1, 1, \dots, 1]^T, z \in \mathfrak{R}^n, \\ y_i z_i &= 0, \quad y_i, z_i \geq 0, \quad i = 1, 2, \dots, n. \end{aligned}$$

Suppose that  $(y^*, \beta^*, z^*)$  is a solution of (3.26). Then,  $y^*$  is a solution of  $(\hat{P}_+^1)$ . The system (3.26) can be solved by many efficient algorithms; see [11]. In this paper, we will apply the SNM proposed in [31] to solve the system (3.26). Under some conditions, the SNM [31] is superlinearly convergent. See [31] for details about the SNM for solving (3.26).  $(\hat{P}_+^2)$  and  $(\hat{P}_+^d)$  can also be solved in a similar way by an SNM as for  $(\hat{P}_+^1)$ . So we omit the details about the SNMs for  $(\hat{P}_+^2)$  and  $(\hat{P}_+^d)$ .

In this section, we have proposed three methods for solving the relaxations  $(\bar{P}_+^1)$ ,  $(\bar{P}_+^2)$ , and  $(\bar{P}_+^d)$ , including the GP method, the PM, and the SNM. To conclude this section, we remark that the irreducibility of the tensors  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{C}$  is assumed to ensure that these methods are convergent. By Theorems 3.1, 3.4, and 3.7, the irreducibility condition can also ensure the positivity of the solutions of  $(\bar{P}_+^1)$ ,  $(\bar{P}_+^2)$ , and  $(\bar{P}_+^d)$ . Recently, a PM has been proposed in [16] for computing the largest eigenvalue for reducible nonnegative tensors. This method may be used to solve  $(\bar{P}_+^1)$  when  $\mathcal{B}$  is a reducible nonnegative tensor. When  $\mathcal{A}$  and  $\mathcal{C}$  are reducible, how do we solve  $(\bar{P}_+^2)$  and  $(\bar{P}_+^d)$ ? We leave it as one of our future research topics.

**4. Numerical experiments.** In this section, we are going to test the performance of the approximation algorithms proposed. We will focus on the cases  $d = 3$  and 4. All algorithms are implemented in MATLAB (R2008b) and all the numerical computations are conducted using an Intel 3.20 GHz computer with 1.93 GB of RAM. All test problems are randomly generated.

**4.1. First experiment.** In our first experiment, we compare the efficiency of the algorithms proposed in section 3 for solving the relaxations  $(\bar{P}_+^1)$ ,  $(\bar{P}_+^2)$ , and  $(\bar{P}_+^d)$ . We only tested algorithms for solving  $(\bar{P}_+^1)$  for some test problems with  $d = 3$ . We implemented three algorithms for solving  $(\bar{P}_+^1)$  proposed in section 3. We use the **ggplab** [13] to solve the geometric programming problem (GP1) defined in (3.11). For convenience of comparison, we refer to the algorithm used in **ggplab** [13] as *GP*. We let the power method proposed in section 3 be denoted by *PM* and the smoothing Newton method by *SNM*. Our numerical results are reported in Tables 2 and 3 and Figure 1. In Tables 2 and 3, *Ite* denotes the number of iterations of PM and SNM,  $f_{\mathcal{B}}(x^*)$  denotes the value of  $f_{\mathcal{B}}(x)$  at the final iteration, and *CPU(s)* denotes the total computer time in seconds used to solve the problem. From Tables 2 and 3, we can see that these three algorithms can solve all the test problems with similar optimal values. The results in Figure 1 (average for 10 data sets of each size) show that PM has better performance than SNM and GP. In the second experiment, we will use PMs to solve the relaxations  $(\bar{P}_+^1)$ ,  $(\bar{P}_+^2)$ , and  $(\bar{P}_+^d)$ .

TABLE 2  
Numerical results of GP, PM, and SNM for small size problems.

Problem		PM		SNM		GP	
n	d	$f_{\mathcal{B}}(x^*)$	CPU(s)	$f_{\mathcal{B}}(x^*)$	CPU(s)	$f_{\mathcal{B}}(x^*)$	CPU(s)
10	3	9.59e+3	0.00	9.59e+3	0.11	9.59e+3	22.33
15	3	1.77e+4	0.00	1.77e+4	0.11	1.77e+4	158.25
20	3	7.82e+4	0.00	7.82e+4	0.16	7.82e+4	663.66
25	3	1.08e+5	0.00	1.08e+5	0.41	1.08e+5	2034.66
30	3	1.03e+5	0.03	1.03e+5	0.73	1.03e+5	5117.14
35	3	1.82e+5	0.03	1.82e+5	1.14	1.82e+5	11164.05
40	3	2.28e+5	0.05	2.28e+5	1.77	2.28e+5	22026.02

TABLE 3  
Numerical results of PM and SNM for large size problems.

Problem		PM			SNM		
n	d	Ite	$f_{\mathcal{B}}(x^*)$	CPU(s)	Ite	$f_{\mathcal{B}}(x^*)$	CPU(s)
60	3	7	4.66e+5	0.19	12	4.66e+5	3.03
80	3	6	9.07e+5	0.22	12	9.07e+5	5.25
100	3	7	1.60e+6	0.42	13	1.60e+6	8.64
120	3	7	2.29e+6	0.64	15	2.29e+6	14.63
140	3	7	2.49e+6	1.00	16	2.49e+6	22.39
160	3	7	3.51e+6	3.61	14	3.51e+6	57.25
180	3	7	4.27e+6	4.09	15	4.27e+6	69.13
200	3	7	7.24e+6	6.77	15	7.24e+6	115.48

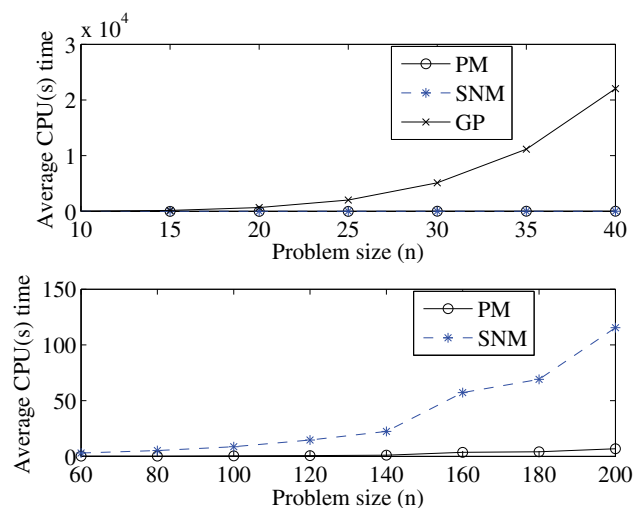


FIG. 1. Average CPU time in seconds obtained by GP, PM, and SNM.

**4.2. Upper bounds.** In our second experiment, we will test the quality of the approximation solutions obtained by the relaxations  $(\bar{P}_+^1)$ ,  $(\bar{P}_+^2)$ , and  $(\bar{P}_+^d)$ . To this end, we first consider the upper bounds of the optimal values of  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$ .

By Theorem 2.2, if  $(\hat{x}^1, \hat{x}^2, \dots, \hat{x}^d)$  is a global solution of  $(\bar{P}_+^d)$ , then  $F_A(\hat{x}^1, \hat{x}^2, \dots, \hat{x}^d)$  is an upper bound of the optimal value of  $(P^d)$ . Similarly, if  $\hat{x}$  is a global solution of  $(\bar{P}_+^1)$ , then  $f_{\mathcal{B}}(\hat{x})$  is an upper bound of the optimal value of  $(P^1)$ . If  $(\hat{x}, \hat{y})$  is a global solution of  $(\bar{P}_+^2)$ , then  $G_{\mathcal{C}}(\hat{x}, \hat{y})$  is an upper bound of the optimal value of  $(P^2)$ .

Upper bounds of the optimal values of  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$  can also be computed as follows; see [14]. Problems  $(P^1)$ ,  $(P^2)$ , and  $(P^d)$  can be relaxed to

$$\begin{aligned}
 (UP^1) \quad & \max \sum_{i_1, i_2, \dots, i_d=1}^n b_{i_1 i_2 \dots i_d} Z_{i_1 i_2 \dots i_{d-1}} x_{i_d} \\
 & \text{s.t. } \|x\| = 1, x \in \mathfrak{R}^n, \\
 & \|Z\| = 1, Z \in \mathfrak{R}^{\overbrace{n \times \dots \times n}^{d-1}}, \\
 (UP^2) \quad & \max \sum_{i_1, \dots, i_p=1}^n \sum_{j_1, \dots, j_q=1}^m c_{i_1 \dots i_p j_1 \dots j_q} Z_{i_1 \dots i_p j_1 \dots j_{q-1}} y_{j_q} \\
 & \text{s.t. } \|Z\| = 1, Z \in \mathfrak{R}^{\overbrace{n \times \dots \times n}^p \times \overbrace{m \times \dots \times m}^{q-1}}, \\
 & \|y\| = 1, y \in \mathfrak{R}^m,
 \end{aligned}$$

and

$$\begin{aligned}
 (UP^d) \quad & \max \sum_{1 \leq i_1 \leq n_1, 1 \leq i_2 \leq n_2, \dots, 1 \leq i_d \leq n_d} a_{i_1 i_2 \dots i_d} Z_{i_1 \dots i_{d-1}} x_{i_d}^d \\
 & \text{s.t. } \|x^d\| = 1, x^d \in \mathfrak{R}^{n_d}, \\
 & \|Z\| = 1, Z \in \mathfrak{R}^{n_1 \times n_2 \times \dots \times n_{d-1}},
 \end{aligned}$$

respectively. Some algorithms for  $(UP^1)$ ,  $(UP^2)$ , and  $(UP^d)$  have been proposed in [14]. The optimal values of  $(UP^1)$ ,  $(UP^2)$ , and  $(UP^d)$  are upper bounds of the optimal values of  $(P^1)$ ,  $(P^2)$  and  $(P^d)$ , respectively.

**4.3. Test procedures.** Based on the discussion in section 4.2, we describe, in the following, our test procedures for our second experiment.

**Test Procedure 1**

1. Solve the relaxation  $(\bar{P}_+^1)$ , and let  $\hat{x}$  be a global solution of  $(\bar{P}_+^1)$ . Solve the relaxation  $(UP^1)$ , and denote its optimal value as  $\hat{v}_1$ . Let  $\bar{v}_1 = \min\{\hat{v}_1, f_B(\hat{x})\}$ . Then,  $\bar{v}_1$  is an upper bound of the optimal value of  $(P^1)$ .
2. Let  $x^* = \frac{\hat{x}}{\|\hat{x}\|}$ . Then,  $x^*$  is an approximation solution of  $(P^1)$ . Report  $\tau_1 = \frac{f_B(x^*)}{\bar{v}_1}$ , the approximation ratio of the solution  $x^*$ .

**Test Procedure 2**

1. Solve the relaxation  $(\bar{P}_+^2)$ , and let  $(\hat{x}, \hat{y})$  be a global solution of  $(\bar{P}_+^2)$ . Solve the relaxation  $(UP^2)$ , and denote its optimal value as  $\hat{v}_2$ . Let  $\bar{v}_2 = \min\{\hat{v}_2, G_C(\hat{x}, \hat{y})\}$ . Then,  $\bar{v}_2$  is an upper bound of the optimal value of  $(P^2)$ .
2. Let  $(x^*, y^*) = (\frac{\hat{x}}{\|\hat{x}\|}, \frac{\hat{y}}{\|\hat{y}\|})$ . Then,  $(x^*, y^*)$  is an approximation solution of  $(P^2)$ . Report  $\tau_2 = \frac{G_C(x^*, y^*)}{\bar{v}_2}$ , the approximation ratio of the solution  $(x^*, y^*)$ .

**Test Procedure 3**

1. Solve the relaxation  $(\bar{P}_+^d)$ , and let  $(\hat{x}^1, \hat{x}^2, \dots, \hat{x}^d)$  be a global solution of  $(\bar{P}_+^d)$ . Solve the relaxation  $(UP^d)$ , and denote its optimal value as  $\hat{v}_3$ . Let  $\bar{v}_3 = \min\{\hat{v}_3, F_A(\hat{x}^1, \hat{x}^2, \dots, \hat{x}^d)\}$ . Then,  $\bar{v}_3$  is an upper bound of the optimal value of  $(P^d)$ .
2. Let  $((x^1)^*, \dots, (x^d)^*) = (\frac{\hat{x}^1}{\|\hat{x}^1\|}, \dots, \frac{\hat{x}^d}{\|\hat{x}^d\|})$ . Then,  $((x^1)^*, \dots, (x^d)^*)$  is an approximation solution of  $(P^d)$ . Report  $\tau_3 = \frac{F_A((x^1)^*, \dots, (x^d)^*)}{\bar{v}_3}$ , the approximation ratio of the solution  $((x^1)^*, \dots, (x^d)^*)$ .

TABLE 4  
*Numerical results of Test Procedures 1–3 for test problems with  $d = 3$ .*

$n$	20	30	40	50	60	70	80	90	100
$\tau_1$	0.9985	0.9991	0.9993	0.9994	0.9996	0.9996	0.9996	0.9997	0.9997
$\tau_2$	0.9955	0.9972	0.9978	0.9985	0.9986	0.9988	0.9990	0.9991	0.9992
$\tau_3$	0.9924	0.9944	0.9961	0.9968	0.9974	0.9976	0.9979	0.9982	0.9983

TABLE 5  
*Numerical results of Test Procedures 1–3 for test problems with  $d = 4$ .*

$n$	10	20	30	40	50
$\tau_1$	0.9988	0.9995	0.9997	0.9998	0.9998
$\tau_2$	0.9949	0.9977	0.9985	0.9989	0.9991
$\tau_3$	0.9836	0.9917	0.9944	0.9958	0.9967

**4.4. Second experiment.** In our second experiment, we implemented Test Procedures 1–3 and tested them for the test problems with  $d = 3$  and 4. The relaxations  $(\bar{P}_+^1)$ ,  $(\bar{P}_+^2)$ , and  $(\bar{P}_+^d)$  are solved by using the PMs proposed in section 3. We use the algorithm DR2 of [14] to solve  $(UP^1)$ ,  $(UP^2)$ , and  $(UP^d)$ . Our numerical results are reported in Tables 4 and 5. The results in these two tables (which are average for 10 data sets of each size) show that our numerical approximation ratios are close to 1, which means our proposed approximation methods can produce very high quality approximation solutions.

**5. Conclusion.** Nonnegative polynomial optimization over unit spheres is a challenging problem because it is NP-hard. In this paper, we have proposed polynomial-time approximation algorithms with new approximation bounds for this optimization problem; see Table 1. In addition, unlike the SDP relaxations in [14, 22], the convex optimization relaxations used in this paper have the same size as the original problems. This means that our proposed approximation algorithms can be used to solve large size problems. Numerical results reported in section 4 showed that the proposed approximation algorithms are practical and they produce very high quality solutions.

**Acknowledgments.** The authors are grateful to the associate editor and the anonymous referees for their valuable comments which led to several improvements of this paper.

#### REFERENCES

- [1] L. BARATCHART, M. BERTHOD, AND L. POTTIER, *Optimization of positive generalized polynomials under  $l_p$  constraints*, J. Convex Anal., 5 (1998), pp. 353–379.
- [2] L. BLOY AND R. VERMA, *On computing the underlying fiber directions from the diffusion orientation distribution function*, in Medical Image Computing and Computer-Assisted Intervention MICCAI 2008, Springer, Berlin, Heidelberg, 2008, pp. 1–8.
- [3] S. BOYD, S.-J. KIM, L. VANDENBERGHE, AND A. HASSIBI, *A tutorial on geometric programming*, Optim. Eng., 8 (2007), pp. 67–127.
- [4] S. BOYD AND L. VANDENBERGHE, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.
- [5] S. R. BULÒ AND M. PELILLO, *New bounds on the clique number of graphs based on spectral hypergraph theory*, in Learning and Intelligent Optimization, T. Stützel, ed., Springer-Verlag, Berlin, 2009, pp. 45–48.
- [6] K. C. CHANG, K. PEARSON, AND T. ZHANG, *Perron-Frobenius theorem for nonnegative tensors*, Commun. Math. Sci., 6 (2008), pp. 507–520.

- [7] K. C. CHANG, K. PEARSON, AND T. ZHANG, *Primitivity, the convergence of the NQZ method, and the largest eigenvalue for nonnegative tensors*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 806–819.
- [8] K. C. CHANG, L. QI, AND G. ZHOU, *Singular values of a real rectangular tensor*, J. Math. Anal. Appl., 370 (2010), pp. 284–294.
- [9] A. CICHOCKI, R. ZDUNEK, A.-H. PHAN, AND S. AMARI, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, John Wiley, New York, 2009.
- [10] L. DE LATHAUWER, B. DE MOOR, AND J. VANDEWALLE, *On the best rank-1 and rank- $(R_1, \dots, R_N)$  approximation of higher-order tensors*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1324–1342.
- [11] F. FACCHINEI AND J. S. PANG, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Vols. I and II, Springer-Verlag, New York, 2003.
- [12] S. FRIEDLAND, S. GAUBERT, AND L. HAN, *Perron-Frobenius theorem for nonnegative multilinear forms and extensions*, Linear Algebra Appl., to appear.
- [13] *GGPLAB: A Simple MATLAB Toolbox for Geometric Programming*, [www.stanford.edu/~boyd/ggplab/](http://www.stanford.edu/~boyd/ggplab/) (2006).
- [14] S. HE, Z. LI, AND S. ZHANG, *Approximation algorithms for homogeneous polynomial optimization with quadratic constraints*, Math. Program., 125 (2010), pp. 353–383.
- [15] S. HU AND L. QI, *Algebraic connectivity of an even uniform hypergraph*, J. Combinatorial Optimization, to appear.
- [16] S. HU, Z. HUANG, AND L. QI, *Finding the spectral radius of a nonnegative tensor*, arxiv: 1111.2138v1 [math.NA], 2011.
- [17] 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.
- [18] T. G. KOLDA AND B. W. BADER, *Tensor decompositions and applications*, SIAM Rev., 51 (2009), pp. 455–500.
- [19] K. O. KORTANEK, X. XU, AND Y. YE, *An infeasible interior-point algorithm for solving primal and dual geometric programs*, Math. Programming, 76 (1996), pp. 155–181.
- [20] J. B. LASSERRE, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., 11 (2001), pp. 796–817.
- [21] L.-H. LIM, *Singular values and eigenvalues of tensors: A variational approach*, in Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP '05), Vol. 1, IEEE, Piscataway, NJ, 2005, pp. 129–132.
- [22] C. LING, J. NIE, L. QI, AND Y. YE, *Biquadratic optimization over unit spheres and semidefinite programming relaxations*, SIAM J. Optim., 20 (2009), pp. 1286–1310.
- [23] Y. LIU, G. ZHOU, AND N. F. IBRAHIM, *An always convergent algorithm for the largest eigenvalue of an irreducible nonnegative tensor*, J. Comput. Appl. Math., 235 (2010), pp. 286–292.
- [24] 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.
- [25] T. S. MOTZKIN AND E. G. STRAUS, *Maxima for graphs and a new proof of a theorem of Turan*, Canad. J. Math., 17 (1965), pp. 533–540.
- [26] *MOSEK Optimization Toolbox*, [www.mosek.com](http://www.mosek.com) (2012).
- [27] Y. NESTEROV, *Random Walk in a Simplex and Quadratic Optimization over Convex Polytopes*, CORE Discussion Paper 2003/71, CORE, Catholic University of Louvain, Louvain-la-Neuve, Belgium, 2003.
- [28] Y. NESTEROV AND A. NEMIROVSKY, *Interior Point Polynomial Algorithms in Convex Programming*, SIAM, Philadelphia, 1994.
- [29] M. NG, L. QI, AND G. ZHOU, *Finding the largest eigenvalue of a nonnegative tensor*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 1090–1099.
- [30] L. QI, *Eigenvalues of a real supersymmetric tensor*, J. Symbolic Comput., 40 (2005), pp. 1302–1324.
- [31] L. QI, D. SUN, AND G. ZHOU, *A new look at smoothing Newton methods for nonlinear complementarity problems and box constrained variational inequalities*, Math. Program., 87 (2000), pp. 1–35.
- [32] L. QI AND K. L. TEO, *Multivariate polynomial minimization and its applications in signal processing*, J. Global Optim., 26 (2003), pp. 419–433.
- [33] L. QI, F. WANG, AND Y. WANG, *Z-eigenvalue methods for a global polynomial optimization problem*, Math. Program., 118 (2009), pp. 301–316.
- [34] L. QI, Y. WANG, AND E. X. WU, *D-eigenvalues of diffusion kurtosis tensor*, J. Comput. Appl. Math., 221 (2008), pp. 150–157.
- [35] A. SHASHUA AND T. HAZAN, *Non-negative tensor factorization with applications to statistics*

- and computer vision*, in Proceedings of the 22nd International Conference on Machine Learning (ICML), ACM, New York, 2005, pp. 792–799.
- [36] Q. YANG AND Y. YANG, *Further results for Perron–Frobenius theorem for nonnegative tensors II*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 1236–1250.
  - [37] Y. YE, *Interior Point Methods: Theory and Analysis*, John Wiley, New York, 1997.
  - [38] L. ZHANG AND L. QI, *Linear convergence of an algorithm for computing the largest eigenvalue of a nonnegative tensor*, Numer. Linear Algebra Appl., to appear.
  - [39] L. ZHANG, L. QI, AND Y. XU, *Finding the largest eigenvalue of an irreducible nonnegative tensor and linear convergence for weakly positive tensors*, J. Comput. Math., 30 (2012), pp. 24–33.
  - [40] X. ZHANG, L. QI, AND Y. YE, *The cubic spherical optimization problems*, Math. Comp., 81 (2012), pp. 1513–1525.
  - [41] G. ZHOU, L. CACCETTA, AND L. QI, *Convergence of an algorithm for the largest singular value of a nonnegative rectangular tensor*, Linear Algebra Appl., to appear.