A hierarchy of semidefinite relaxations for completely positive tensor optimization problems

Anwa Zhou[1] · Jinyan Fan[2]

Abstract

In this paper, we study the completely positive (CP) tensor program, which is a linear optimization problem with the cone of CP tensors and some linear constraints. We reformulate it as a linear program over the cone of moments, then construct a hierarchy of semidefinite relaxations for solving it. We also discuss how to find a best CP approximation of a given tensor. Numerical experiments are presented to show the efficiency of the proposed methods.

Keywords CP tensor program · Best CP tensor approximation · Linear optimization with moments · Nonnegative decomposition · Semidefinite program

Mathematics Subject Classification Primary: 15A69 · 44A60 · 90C22 · 90C26

1 Introduction

For given positive integers m and n, an m-way, n-dimension tensor $\mathcal{A}$ is a multidimensional array indexed by integer tuples $(i_1, \ldots, i_m)$ with $1 \leq i_1, \ldots, i_m \leq n$, i.e.,

$$\mathcal{A} = (\mathcal{A}_{i_1 i_2 \ldots i_m})_{1 \leq i_1, \ldots, i_m \leq n}.$$

Let $\mathbb{R}$ be the set of real numbers and $\mathrm{T}^m(\mathbb{R}^n)$ be the set of all such real-valued tensors. A tensor $\mathcal{A} \in \mathrm{T}^m(\mathbb{R}^n)$ is called symmetric if each entry $\mathcal{A}_{i_1 i_2 \ldots i_m}$ is invariant with respect to all permutations of $(i_1, \ldots, i_m)$. Let $\mathrm{S}^m(\mathbb{R}^n)$ be the subspace of all symmetric tensors in $\mathrm{T}^m(\mathbb{R}^n)$.

✉ Jinyan Fan
  jyfan@sjtu.edu.cn

  Anwa Zhou
  zhouanwa@shu.edu.cn

[1] Department of Mathematics, Shanghai University, Shanghai 200444, People's Republic of China

[2] School of Mathematical Sciences, and MOE-LSC, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China

For a vector $u \in \mathbb{R}^n$, denote $u^{\otimes m}$ the m-way n-dimensional symmetric outer product tensor such that

$$(u^{\otimes m})_{i_1,\dots,i_m} = u_{i_1} \cdots u_{i_m}.$$

A tensor in $S^m(\mathbb{R}^n)$ is called rank-1 if it has the form $\lambda u^{\otimes m}$ where $\lambda \in \mathbb{R}$ and $u \in \mathbb{R}^n$. Clearly, if m is odd or $\lambda > 0$, then the mth real root of λ always exists, so we can rewrite the tensor as

$$\lambda u^{\otimes m} = v^{\otimes m} \text{ with } v = \sqrt[m]{\lambda} u.$$

For the other case, the scalar λ cannot be absorbed into the vector v. Comon et al. [5] showed that every symmetric tensor $\mathcal{A} \in S^m(\mathbb{R}^n)$ can be decomposed as

$$\mathcal{A} = \sum_{k=1}^{r} \lambda_k (u^k)^{\otimes m}, \tag{1.1}$$

where $\lambda_k \in \mathbb{R}$ and $u^k \in \mathbb{R}^n$ for $k = 1, \dots, r$. Interested readers are referred to [4,6,14,16–18,25,26] for numerical methods for computing real eigenvalues of symmetric tensors, tensor decompositions, as well as applications.

Let $\mathbb{R}_+^n$ be the nonnegative orthant of $\mathbb{R}^n$. A symmetric tensor $\mathcal{A} \in S^m(\mathbb{R}^n)$ is completely positive (CP), if there exist vectors $v^1, \dots, v^r \in \mathbb{R}_+^n$ such that

$$\mathcal{A} = \sum_{k=1}^{r} (v^k)^{\otimes m}, \tag{1.2}$$

where r is called the length of the decomposition (1.2) (cf. [30,31]). The smallest r in the above is called the cp-rank of $\mathcal{A}$. If $\mathcal{A}$ is CP, (1.2) is called a CP decomposition (or nonnegative decomposition) of $\mathcal{A}$. The CP tensor is a natural extension of the CP matrix. We refer to [1,2,21,35,36] for work on CP matrices and tensors.

For $\mathcal{B} \in S^m(\mathbb{R}^n)$, we define

$$\mathcal{B}x^m := \sum_{1 \leq i_1,\dots,i_m \leq n} \mathcal{B}_{i_1 i_2 \dots i_m} x_{i_1} \cdots x_{i_m}.$$

If $\mathcal{B}x^m \geq 0$ for all $x \in \mathbb{R}_+^n$, we call $\mathcal{B}$ a copositive tensor (cf. [31]). The copositive tensor is an extension of the copositive matrix. Obviously, both symmetric nonnegative tensors and positive semidefinite tensors are copositive tensors.

For $\mathcal{A}, \mathcal{B} \in S^m(\mathbb{R}^n)$, the inner product of $\mathcal{A}$ and $\mathcal{B}$ is defined as

$$\mathcal{A} \bullet \mathcal{B} = \sum_{1 \leq i_1,\dots,i_m \leq n} \mathcal{A}_{i_1 i_2 \dots i_m} \mathcal{B}_{i_1 i_2 \dots i_m},$$

and the norm of a tensor $\mathcal{A}$ is the square root of the sum of squares of its entries, i.e.,

$$\|\mathcal{A}\| := \sqrt{\sum_{1 \leq i_1,\dots,i_m \leq n} \mathcal{A}_{i_1 i_2 \dots i_m}^2}.$$

For a cone $C \subseteq S^m(\mathbb{R}^n)$, the dual cone of C is defined as

$$C^* := \{\mathcal{B} \in S^m(\mathbb{R}^n) : \mathcal{A} \bullet \mathcal{B} \geq 0 \text{ for all } \mathcal{A} \in C\}.$$

Denote by $CP_{m,n}$ and $COP_{m,n}$ the sets of m-way, n-dimension CP tensors and copositive tensors, respectively. Both $CP_{m,n}$ and $COP_{m,n}$ are proper cones, and they are dual to each other [30].

CP tensor and nonnegative decomposition have wide applications in statistics, computer vision, exploratory multiway data analysis and blind source separation [3,32,33]. Generally, it is hard to check whether a tensor is CP or not. The problem is NP-hard even for the CP matrix case [8]. Kolda [18] assumed the length of the CP decomposition is known and formulated the CP tensor decomposition problem as a nonnegative constrained least squares problem, then solved it by a nonlinear optimization software. Qi et al. [30] showed that a strongly symmetric hierarchically dominated nonnegative tensor is CP and presented a hierarchical elimination algorithm for checking this. For a general symmetric tensor, we transformed the problem of checking whether it is CP to a truncated moment problem, then proposed a semidefinte algorithm for solving it [9].

Recently, Penã et. al [28] showed that under certain conditions a general polynomial optimization problem (POP), which are not necessarily quadratic, can be formulated as a linear optimization problem with the cone of CP tensors and some linear equality constraints. In [15], Kuang and Zuluaga gave the CP tensor relaxation for the POP and showed that it provides a better bound than the Lagrangian relaxation. They also proved that the CP tensor relaxation yields a better bound than the CP matrix relaxation for quadratic reformulations of some class of polynomial optimization problems. Since the checking of CP tensors is NP-hard, it is difficult to deal with the cone of CP tensors directly. A general way is to approximate it by some tractable cones, such as the cones of nonnegative tensors, the positive semidefinite tenors and the doubly nonnegative tensors [15]. However, this approach generally only provides a bound for the original problem and gives an approximate solution of the problem.

The aim of this paper is to develop a new method for the CP tensor program, which is stated as:

$$
\begin{aligned}
\min_{\mathcal{X}} \;& \mathcal{A} \bullet \mathcal{X} \\
\text{s.t. } & \mathcal{A}_i \bullet \mathcal{X} = b_i, \, i = 1, \ldots, l_1, \\
& \mathcal{A}_i \bullet \mathcal{X} \geq b_i, \, i = l_1 + 1, \ldots, l, \\
& \mathcal{X} \in CP_{m,n},
\end{aligned}
\tag{1.3}
$$

where $\mathcal{A}, \mathcal{A}_i \in S^m(\mathbb{R}^n), b_i \in \mathbb{R}, i = 1, \ldots, l$. We formulate it as a linear optimization problem with the cone of moments, then construct a hierarchy of semidefinite relaxations for solving it. If (1.3) is infeasible, we can get a certificate for that. If it is feasible, a CP decomposition of the optimal solution can also be obtained. We also discuss the best CP tensor approximation problem, which is to find a tensor in the intersection of a set of linear constraints and the cone of CP tensors such that it is close to a given tensor as much as possible. It is an extension of the CP-matrix approximation problem [10]. We transform the problem to a conic linear program over the cone of moments and the second-order cone. A hierarchical semidefinite relaxation algorithm is also presented for it.

The paper is organized as follows. In Sect. 2, we characterize the CP tensor as a moment sequence. In Sect. 3, we formulate the CP tensor program as a linear optimization problem with the cone of moments, then propose a smidefinite algorithm for solving it. The convergence properties of the algorithm are also studied. In Sect. 4, we show how to solve the best CP tensor approximation problem. Some computational experiments are presented in Sect. 5.

2 Moments and outer approximations

In this section, we characterize a CP tensor as a truncated moment sequence, then give some necessary and sufficient conditions for a symmetric tensor to be CP. A hierarchy of outer approximations of the cone of CP tensors is also introduced.

2.1 Moments and flat extension

As we know, a symmetric matrix can be identified by a vector that consists of its upper triangular entries. Similarly, a symmetric tensor $\mathcal{A} \in S^m(\mathbb{R}^n)$ can also be identified by a vector $\text{vech}(\mathcal{A})$ that consists of its upper triangular entries, i.e. the entries

$$\mathcal{A}_{i_1 i_2 \ldots i_m} \text{ with } 1 \le i_1 \le \cdots \le i_m \le n.$$

Let $\mathbb{N}$ be the set of nonnegative integers. For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, denote $|\alpha| := \alpha_1 + \cdots + \alpha_n$. Let

$$\mathbb{T} := \{\alpha \in \mathbb{N}^n : |\alpha| = m\}. \tag{2.1}$$

Then, there is a one-to-one correspondence between the index pair $(i_1, \ldots i_m)$ and the vector

$$\alpha = e_{i_1} + e_{i_2} + \cdots + e_{i_m} \in \mathbb{T},$$

where e_i is the i-th unit vector in $\mathbb{R}^n$. Hence, the identifying vector of $\mathcal{A} \in S^m(\mathbb{R}^n)$ can also be written as

$$\mathbf{a} = (\mathbf{a}_\alpha)_{\alpha \in \mathbb{T}} \in \mathbb{R}^{\mathbb{T}} \text{ with } \mathbf{a}_\alpha = \mathcal{A}_{i_1 i_2 \ldots i_m}, \quad \text{if } \alpha = e_{i_1} + e_{i_2} + \cdots + e_{i_m},$$

where $\mathbb{R}^{\mathbb{T}}$ denotes the space of real vectors indexed by $\alpha \in \mathbb{T}$. We call $\mathbf{a}$ a $\mathbb{T}$-truncated moment sequence ($\mathbb{T}$-tms).

Let

$$K = \{x \in \mathbb{R}^n : x^T x - 1 = 0, x \ge 0\}. \tag{2.2}$$

Note that every nonnegative vector is a multiple of a vector in K. By (1.2), $\mathcal{A} \in CP_{m,n}$ if and only if there exist $\lambda_1, \ldots, \lambda_r > 0$ and $u^1, \ldots, u^r \in K$ such that

$$\mathcal{A} = \lambda_1 (u^1)^{\otimes m} + \cdots + \lambda_r (u^r)^{\otimes m}. \tag{2.3}$$

We say a $\mathbb{T}$-tms $\mathbf{a}$ admits a K-measure μ, i.e., a nonnegative Borel measure μ supported in K such that

$$\mathbf{a}_\alpha = \int_K x^\alpha d\mu, \quad \forall \alpha \in \mathbb{T},$$

where $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. A measure μ satisfying the above is called a K-representing measure for $\mathbf{a}$. A measure is called finitely atomic if its support is a finite set, and is called r-atomic if its support consists of at most r distinct points. We refer to [22] for representing measures of truncated moment sequences.

Hence, by (2.3), a tensor $\mathcal{A} \in S^m(\mathbb{R}^n)$, with the identifying vector $\mathbf{a} \in \mathbb{R}^{\mathbb{T}}$, is CP if and only if $\mathbf{a}$ admits an r-atomic K-measure, i.e.,

$$\mathbf{a} = \lambda_1 [u^1]_{\mathbb{T}} + \cdots + \lambda_r [u^r]_{\mathbb{T}}, \tag{2.4}$$

where each $\lambda_i > 0$, $u^i \in K$ and

$$[u]_{\mathbb{T}} := (u^\alpha)_{\alpha \in \mathbb{T}}.$$

Denote

$$\mathcal{M}_{\mathbb{T},K} = \{\mathbf{a} \in \mathbb{R}^{\mathbb{T}} : \mathbf{a} \text{ admits a } K\text{-measure}\}. \tag{2.5}$$

Then, $\mathcal{M}_{\mathbb{T},K}$ is the CP tensor cone. So,

$$\mathcal{A} \in CP_{m,n} \quad \text{if and only if} \quad \mathbf{a} \in \mathcal{M}_{\mathbb{T},K}. \tag{2.6}$$

Denote

$$\mathbb{R}[x]_{\mathbb{T}} := \text{span}\{x^\alpha : \alpha \in \mathbb{T}\}.$$

For a polynomial $p \in \mathbb{R}[x]_{\mathbb{T}}$, $p|_K \geq 0$ (resp., $p|_K > 0$) denotes $p \geq 0$ (resp., $p > 0$) on K. Then the dual cone of $\mathcal{M}_{\mathbb{T},K}$ is

$$\mathcal{P}_{\mathbb{T},K} = \{p \in \mathbb{R}[x]_{\mathbb{T}} : p|_K \geq 0\}, \tag{2.7}$$

which is the copositive tensor cone $COP_{m,n}$ (cf. [20,23]). We say $\mathbb{R}[x]_{\mathbb{T}}$ is K-full if there exists a polynomial $p \in \mathbb{R}[x]_{\mathbb{T}}$ such that $p|_K > 0$ (cf. [11]). It is easy to check that $\mathbb{R}[x]_{\mathbb{T}}$ is K-full.

For a $\mathbb{T}$-tms $\mathbf{a} \in \mathbb{R}^{\mathbb{T}}$, define a linear Riesz functional $\mathcal{L}_{\mathbf{a}}$ acting on $\mathbb{R}[x]_{\mathbb{T}}$ as

$$\mathcal{L}_{\mathbf{a}}\left(\sum_{\alpha \in \mathbb{T}} p_\alpha x^\alpha\right) := \sum_{\alpha \in \mathbb{T}} p_\alpha \mathbf{a}_\alpha. \tag{2.8}$$

We also denote $\langle p, \mathbf{a} \rangle := \mathcal{L}_{\mathbf{a}}(p)$ for convenience. Let

$$\mathbb{N}_d^n := \{\alpha \in \mathbb{N}^n : |\alpha| \leq d\},$$

and

$$\mathbb{R}[x]_d := \text{span}\{x^\alpha : \alpha \in \mathbb{N}_d^n\}.$$

For $s \in \mathbb{R}^{\mathbb{N}_{2k}^n}$ and $q \in \mathbb{R}[x]_{2k}$, the k-th localizing matrix of q generated by s is the symmetric matrix $L_q^{(k)}(s)$ satisfying

$$\mathcal{L}_s(qp^2) = \text{vec}(p)^T (L_q^{(k)}(s)) \text{vec}(p), \quad \forall p \in \mathbb{R}[x]_{k-\lceil deg(q)/2 \rceil}, \tag{2.9}$$

where $\text{vec}(p)$ denotes the coefficient vector of p in the graded lexicographical ordering, and $\lceil t \rceil$ denotes the smallest integer that is not smaller than t. In particular, when $q = 1$, $L_1^{(k)}(s)$ is called a k-th order moment matrix and denoted as $M_k(s)$.
In fact, we have

$$L_q^{(k)}(s) = \left(\sum_\alpha q_\alpha s_{\alpha+\beta+\gamma}\right)_{\beta,\gamma \in \mathbb{N}_{k-\lceil deg(q)/2 \rceil}^n},$$

and

$$M_k(s) = L_1^{(k)}(s) = (s_{\beta+\gamma})_{\beta,\gamma \in \mathbb{N}_k^n}.$$

We refer to [19,20,22] for more details about localizing matrices and moment matrices.
Denote the polynomials:

$$h(x) := x^T x - 1, \ g_0(x) := 1, \ g_1(x) := x_1, \ \ldots, \ g_n(x) := x_n.$$

Note that K given in (2.2) is nonempty compact. It can also be described equivalently as

$$K = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \geq 0\}, \tag{2.10}$$

where $g(x) = (g_0(x), g_1(x), \ldots, g_n(x))$. As shown in Nie [22], for $s \in \mathbb{R}^{\mathbb{N}_{2k}^n}$, if

$$L_h^{(k)}(s) = 0 \quad \text{and} \quad L_{g_j}^{(k)}(s) \succeq 0, \quad j = 0, 1, \ldots, n, \tag{2.11}$$

$$\operatorname{rank}(M_{k-1}(s)) = \operatorname{rank}(M_k(s)), \tag{2.12}$$

then s admits a unique K-measure, which is $\operatorname{rank} M_k(s)$-atomic (cf. Curto and Fialkow [7]). We say s is flat if both (2.11) and (2.12) are satisfied.

Given two tms' $y \in \mathbb{R}^{\mathbb{N}_d^n}$ and $z \in \mathbb{R}^{\mathbb{N}_e^n}$, we say z is an extension of y, if $d \leq e$ and $y_\alpha = z_\alpha$ for all $\alpha \in \mathbb{N}_d^n$. We denote by $z|_{\mathbb{T}}$ the subvector of z, whose entries are indexed by $\alpha \in \mathbb{T}$. For convenience, we denote by $z|_d$ the subvector $z|_{\mathbb{N}_d^n}$. If z is flat and extends y, we say z is a flat extension of y. It is shown in [22] that a $\mathbb{T}$-tms $\mathbf{a} \in \mathbb{R}^{\mathbb{T}}$ admits a K-measure if and only if it is extendable to a flat tms $z \in \mathbb{R}^{\mathbb{N}_{2k}^n}$ for some k. So, by (2.6),

$$\mathcal{A} \in CP_{m,n} \quad \text{if and only if} \quad \mathbf{a} \text{ has a flat extension.} \tag{2.13}$$

2.2 Outer approximations

We call a subset $I \subseteq \mathbb{R}[x]$ an ideal if $I + I \subseteq I$ and $I \cdot \mathbb{R}[x] \subseteq I$. For a tuple $\chi = (\chi_1, \ldots, \chi_l)$ of polynomials in $\mathbb{R}[x]$, denoted by $I(\chi)$ the ideal generated by $\chi_1, \ldots, \chi_l$. The smallest ideal containing all χ_i is the set $\chi_1 \mathbb{R}[x] + \cdots + \chi_l \mathbb{R}[x]$. A polynomial $f \in \mathbb{R}[x]$ is called a sum of squares (SOS) if there exist $f_1, \ldots, f_t \in \mathbb{R}[x]$ such that $f = f_1^2 + \cdots + f_t^2$.

Let h and g be as in (2.10). Denote

$$I_{2k}(h) = \{h(x)\phi(x) : \deg(h\phi) \leq 2k\}, \tag{2.14}$$

and

$$Q_k(g) = \left\{ \sum_{j=0}^n g_j \varphi_j : \text{each } \deg(g_j \varphi_j) \leq 2k \text{ and } \varphi_j \text{ is SOS} \right\}. \tag{2.15}$$

Then, $I(h) = \bigcup_{k \in \mathbb{N}} I_{2k}(h)$ is the ideal generated by h, and $Q(g) = \bigcup_{k \in \mathbb{N}} Q_k(g)$ is the quadratic module generated by g (cf. [23]). We say $I(h) + Q(g)$ is archimedean if there exists a number $N > 0$ such that $N - \|x\|^2 \in I(h) + Q(g)$. Clearly, if $f \in I(h) + Q(g)$, then $f|_K \geq 0$. Conversely, if $f|_K > 0$ and $I(h) + Q(g)$ is archimedean, then $f \in I(h) + Q(g)$. This is due to Putinar's Positivstellensatz (cf. [29]).

For each $k \in \mathbb{N}$ and $k \geq \lceil m/2 \rceil$, denote

$$\Phi_k = \{p \in \mathbb{R}[x]_{\mathbb{T}} : p \in I_{2k}(h) + Q_k(g)\}. \tag{2.16}$$

Recall that $\mathbb{R}[x]_{\mathbb{T}}$ is K-full, $\mathbb{T}$ is finite and $I(h) + Q(g)$ is archimedean because $1 - \|x\|^2 = -h(x) \in I(h) + Q(g)$. By Nie [23, Propositions 3.5], we have

$$\Phi_{\lceil m/2 \rceil} \subseteq \cdots \subseteq \Phi_k \subseteq \Phi_{k+1} \subseteq \cdots \subseteq \mathcal{P}_{\mathbb{T},K}. \tag{2.17}$$

Moreover,

$$\operatorname{cl}\left(\bigcup_{k=\lceil m/2 \rceil}^\infty \Phi_k \right) = \mathcal{P}_{\mathbb{T},K}, \tag{2.18}$$

where $\operatorname{cl}(\cdot)$ denotes the closure of a set.

Correspondingly, for each $k \in \mathbb{N}$ and $k \geq \lceil m/2 \rceil$, denote

$$\Gamma_k = \left\{ s \in \mathbb{R}^{\mathbb{N}^n_{2k}} : L_h^{(k)}(s) = 0, L_{g_j}^{(k)}(s) \succeq 0, j = 0, 1, \ldots, n \right\}, \tag{2.19}$$

and

$$\Omega_k = \{ s|_{\mathbb{T}} : s \in \Gamma_k \}. \tag{2.20}$$

Since $\mathbb{T}$ is finite, $\mathbb{R}[x]_{\mathbb{T}}$ is K-full and $I(h) + Q(g)$ is archimedean, by Nie [23, Proposition 3.3], we have

$$\Omega_{\lceil m/2 \rceil} \supseteq \cdots \supseteq \Omega_k \supseteq \Omega_{k+1} \supseteq \cdots \supseteq \mathcal{M}_{\mathbb{T}, K}, \tag{2.21}$$

and

$$\bigcap_{k=\lceil m/2 \rceil}^{\infty} \Omega_k = \mathcal{M}_{\mathbb{T}, K}. \tag{2.22}$$

Moreover, Φ_k and Ω_k are dual to each other.

As shown above, the hierarchy of Φ_k provides inner approximations of $\mathcal{P}_{\mathbb{T}, K}$ and converges monotonically and asymptotically to $\mathcal{P}_{\mathbb{T}, K}$, while the hierarchy of Ω_k provides outer approximations of $\mathcal{M}_{\mathbb{T}, K}$ and converges monotonically and asymptotically to $\mathcal{M}_{\mathbb{T}, K}$. So, Φ_k and Ω_k can approximate $\mathcal{P}_{\mathbb{T}, K}$ and $\mathcal{M}_{\mathbb{T}, K}$, arbitrarily well, respectively.

3 The CP tensor program

In this section, we reformulate the CP tensor program as a linear optimization problem with the cone of moments, then construct a hierarchy of semidefinite relaxations for solving it. The convergence properties of the semidefinite algorithm are also discussed.

3.1 Reformulation and semidefinite relaxations

Denote by

$$\mathbf{x} = \operatorname{vech}(\mathcal{X})$$

the identifying vector of $\mathcal{X}$. Denote by $\mathbf{c} \in \mathbb{R}^{\mathbb{T}}$ a constant vector with the element $\mathbf{c}_\alpha = \sqrt{\frac{m!}{\alpha_1! \cdots \alpha_n!}}$ for each $\alpha \in \mathbb{T}$. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{\mathbb{T}}$, $\mathbf{x} \circ \mathbf{y}$ denotes their Hadamard product. For convenience, we denote $\mathbf{c} \circ \mathbf{c}$ by $\mathbf{c}^2$. Then, (1.3) can be formulated as the linear optimization problem:

$$
(P_1): \quad
\begin{aligned}
\vartheta_{P_1} = \min_{\mathbf{x}} \quad & (\mathbf{c}^2 \circ \operatorname{vech}(\mathcal{A}))^T \mathbf{x} \\
\text{s.t.} \quad & (\mathbf{c}^2 \circ \operatorname{vech}(\mathcal{A}_i))^T \mathbf{x} = b_i, i = 1, \ldots, l_1, \\
& (\mathbf{c}^2 \circ \operatorname{vech}(\mathcal{A}_i))^T \mathbf{x} \geq b_i, i = l_1 + 1, \ldots, l, \\
& \mathbf{x} \in \mathcal{M}_{\mathbb{T}, K},
\end{aligned}
$$

where $\mathcal{M}_{\mathbb{T}, K}$ is given by (2.5). The dual problem of (P_1) is

$$
(D_1): \quad
\begin{aligned}
\vartheta_{D_1} = \max_{\lambda, \mathbf{s}} \quad & b^T \lambda \\
\text{s.t.} \quad & \lambda_i \geq 0, \ i = l_1 + 1, \ldots, l, \\
& \mathbf{c}^2 \circ \operatorname{vech}(\mathcal{A}) - \sum_{i=1}^{l} \lambda_i \mathbf{c}^2 \circ \operatorname{vech}(\mathcal{A}_i) - \mathbf{s} = 0, \\
& \mathbf{s} \in \mathcal{P}_{\mathbb{T}, K},
\end{aligned}
$$

where $\mathcal{P}_{\mathbb{T},K}$ is given by (2.7). Denote by $\mathcal{F}(P)$ the feasible set of the problem (P). By weak duality, for all feasible points $\mathbf{x} \in \mathcal{F}(P_1)$ and $(\lambda, \mathbf{s}) \in \mathcal{F}(D_1)$, we have $\vartheta_{P_1} \geq \vartheta_{D_1}$.

By (2.19) and (2.20), for each $k \in \mathbb{N}$ and $k \geq \lceil m/2 \rceil$, the k-th order relaxation of (P_1) can be defined as

$$(P_1^k):\quad \begin{aligned} \vartheta_{P_1}^k = \min_{\mathbf{x},\tilde{\mathbf{x}}} \quad & (\mathbf{c}^2 \circ \operatorname{vech}(\mathcal{A}))^T \mathbf{x} \\ \text{s.t.} \quad & (\mathbf{c}^2 \circ \operatorname{vech}(\mathcal{A}_i))^T \mathbf{x} = b_i, \, i = 1, \ldots, l_1, \\ & (\mathbf{c}^2 \circ \operatorname{vech}(\mathcal{A}_i))^T \mathbf{x} \geq b_i, \, i = l_1 + 1, \ldots, l, \\ & \mathbf{x} = \tilde{\mathbf{x}}|_{\mathbb{T}}, \, \tilde{\mathbf{x}} \in \Gamma_k. \end{aligned}$$

The dual problem of (P_1^k) is

$$(D_1^k):\quad \begin{aligned} \vartheta_{D_1}^k = \max_{\lambda,\mathbf{s}} \quad & b^T \lambda \\ \text{s.t.} \quad & \lambda_i \geq 0, \, i = l_1 + 1, \ldots, l, \\ & \mathbf{c}^2 \circ \operatorname{vech}(\mathcal{A}) - \sum_{i=1}^{l} \lambda_i \mathbf{c}^2 \circ \operatorname{vech}(\mathcal{A}_i) - \mathbf{s} = 0, \\ & \mathbf{s} \in \Phi_k. \end{aligned}$$

Clearly, $\vartheta_{P_1}^k \leq \vartheta_{P_1}$ and $\vartheta_{D_1}^k \leq \vartheta_{D_1}$ for all k. Suppose that $(\mathbf{x}^{*,k}, \tilde{\mathbf{x}}^{*,k})$ is a minimizer of (P_1^k) and $(\lambda^{*,k}, \mathbf{s}^{*,k})$ is a maximizer of (D_1^k). If $\mathbf{x}^{*,k} = \tilde{\mathbf{x}}^{*,k}|_{\mathbb{T}} \in \mathcal{M}_{\mathbb{T},K}$, then $\vartheta_{P_1}^k = \vartheta_{P_1}$ and $\mathbf{x}^{*,k}$ is a minimizer of (P_1), i.e., the relaxation (P_1^k) is exact for solving (P_1). In this case, if $\vartheta_{P_1}^k = \vartheta_{D_1}^k$, then $\vartheta_{D_1}^k = \vartheta_{D_1}$ and $(\lambda^{*,k}, \mathbf{s}^{*,k})$ is a maximizer of (D_1). If the relaxation (P_1^k) is infeasible, then (P_1) is also infeasible.

3.2 A semidefinite algorithm

Based on the above, we propose a semidefinite algorithm for solving the CP tensor program (1.3).

Algorithm 3.1

Step 0. Input $\mathcal{A}, \mathcal{A}_i \in \mathrm{S}^m(\mathbb{R}^n)$, $b_i \in \mathbb{R}$, $i = 1, \ldots, l$ and K as (2.2). Let $k := \lceil m/2 \rceil$.
Step 1. Solve the relaxation problem (P_1^k). If (P_1^k) is infeasible, stop and output that (P_1) is infeasible; otherwise, compute an optimal solution $(\mathbf{x}^{*,k}, \tilde{\mathbf{x}}^{*,k})$ of (P_1^k). Let $t := 1$.
Step 2. Let $\hat{\mathbf{x}} := \tilde{\mathbf{x}}^{*,k}|_{2t}$. If the rank condition (2.12) is not satisfied, go to Step 4.
Step 3. Compute $\rho_i > 0$, $v^i \in K$ and $r = \operatorname{rank}(M_t(\hat{\mathbf{x}}))$. Output the CP decomposition of $\mathcal{X}^*$ as

$$\mathcal{X}^* = \sum_{i=1}^{r} \rho_i (v^i)^{\otimes m}.$$

Step 4. If $t < k$, set $t := t + 1$ and go to Step 2; otherwise, set $k := k + 1$ and go to Step 1.

We use Step 2 to check whether $\tilde{\mathbf{x}}^{*,k}|_{2t}$ is flat or not. We use Henrion and Lasserre's method in [12] to get an r-atomic K-measure for $\hat{\mathbf{x}}$, which can further produce the CP decomposition of the optimal solution.

We say (D_1) has a relative interior, if there exists a feasible point pair $(\lambda, \mathbf{s}) \in \mathcal{F}(D_1)$ such that $\mathbf{s} \in \operatorname{int}(\mathcal{P}_{\mathbb{T},K})$ and $\lambda_i > 0$, $i = l_1 + 1, \ldots, l$. The asymptotic convergence of Algorithm 3.1 is given as below.

Theorem 3.2 *Let $\mathbb{T}$ and K be as in (2.1) and (2.10), respectively. Suppose (P_1) is feasible and (D_1) has a relative interior point. Algorithm 3.1 has the following properties:*

(i) *For all k sufficiently large, (D_1^k) has a relative interior point and (P_1^k) has a minimizer $(\mathbf{x}^{*,k}, \tilde{\mathbf{x}}^{*,k})$.*

(ii) *The sequence $\{\mathbf{x}^{*,k}\}$ is bounded, and each of its accumulation points is a minimizer of (P_1). The sequence $\{\vartheta_{P_1}^k\}$ converges to the minimum of (P_1).*

Proof (i) Let $(\lambda^0, \mathbf{s}^0) \in \mathcal{F}(D_1)$ with $\mathbf{s}^0 \in \text{int}(\mathcal{P}_{\mathbb{T},K})$ and $\lambda_i^0 > 0 (i = l_1 + 1, \ldots, l)$. Then, $\mathbf{s}^0|_K > 0$ (cf. [23, Lemma 3.1]). Note that since K is compact, there exist $\epsilon > 0$ and $\delta > 0$ such that

$$\mathbf{s}|_K - \epsilon > \epsilon, \quad \forall \mathbf{s} \in B(\mathbf{s}^0, \delta).$$

By Nie [27, Theorem 6], there exists $N_0 > 0$ such that

$$\mathbf{s} - \epsilon \in I_{2N_0}(h) + Q_{N_0}(g), \quad \forall \mathbf{s} \in B(\mathbf{s}^0, \delta).$$

So, (D_1^k) has a relative interior point for all $k \geq N_0$. Thus, the strong duality holds for (P_1^k) and (D_1^k). As (P_1) is feasible, the relaxation problem (P_1^k) is also feasible. So, (P_1^k) has a minimizer $(\mathbf{x}^{*,k}, \tilde{\mathbf{x}}^{*,k})$.

(ii) We first show that $\{\mathbf{x}^{*,k}\}$ is bounded. Let $(\lambda^0, \mathbf{s}^0)$ and ϵ be as in the proof of (i). The set $I_{2N_0}(h) + Q_{N_0}(g)$ is dual to Γ_{N_0}. For all $k \geq N_0$, we have $\tilde{\mathbf{x}}^{*,k} \in \Gamma_{N_0}$ and

$$0 \leq \langle \mathbf{s}^0 - \epsilon, \tilde{\mathbf{x}}^{*,k} \rangle = \langle \mathbf{s}^0, \tilde{\mathbf{x}}^{*,k} \rangle - \epsilon \langle 1, \tilde{\mathbf{x}}^{*,k} \rangle,$$
$$\langle \mathbf{s}^0, \mathbf{x}^{*,k} \rangle \leq \vartheta_{P_1}^k - b^T \lambda^0.$$

Since $\vartheta_{P_1}^k \leq \vartheta_{P_1}$ and $\langle \mathbf{s}^0, \tilde{\mathbf{x}}^{*,k} \rangle = \langle \mathbf{s}^0, \mathbf{x}^{*,k} \rangle$, it holds that

$$\langle \mathbf{s}^0, \tilde{\mathbf{x}}^{*,k} \rangle \leq V_0 := \vartheta_{P_1} - b^T \lambda^0.$$

So,

$$0 \leq \langle \mathbf{s}^0 - \epsilon, \tilde{\mathbf{x}}^{*,k} \rangle \leq V_0 - \epsilon (\tilde{\mathbf{x}}^{*,k})_\mathbf{0},$$
$$(\tilde{\mathbf{x}}^{*,k})_\mathbf{0} \leq V_1 := V_0/\epsilon.$$

Note that $I(h) + Q(g)$ is archimedean, following the line of the proof of [23, Theorem 4.3 (ii)], we can obtain that the sequence $\{\mathbf{x}^{*,k}\}$ is bounded.

Suppose $\mathbf{x}^*$ is an accumulation point of $\{\mathbf{x}^{*,k}\}$. Without loss of generality, we assume

$$\mathbf{x}^{*,k} \to \mathbf{x}^*, \quad k \to \infty.$$

Since $\mathbf{x}^{*,k} \in \Omega_k$, by (2.21) and (2.22), we have $\mathbf{x}^* \in \bigcap_{k=\lceil m/2 \rceil}^\infty \Omega_k = \mathcal{M}_{\mathbb{T},K}$. Note that $\mathbf{x}^{*,k} \in \mathcal{F}(P_1^k)$, we further have $\mathbf{x}^* \in \mathcal{F}(P_1)$. Denote by $\vartheta_{P_1}^*$ the objective function value of (P_1) at $\mathbf{x}^*$. Then,

$$\vartheta_{P_1} \leq \vartheta_{P_1}^*. \tag{3.1}$$

Since (P_1^k) is a relaxation problem of (P_1) and $\mathbf{x}^{*,k}$ is a minimizer of (P_1^k), we have

$$\vartheta_{P_1} \geq \vartheta_{P_1}^k, \ k = \lceil m/2 \rceil, \lceil m/2 \rceil + 1, \ldots.$$

Taking $k \to \infty$, we get

$$\vartheta_{P_1} \geq \lim_{k \to \infty} \vartheta_{P_1}^k = \vartheta_{P_1}^*. \tag{3.2}$$

This, together with (3.1), implies that

$$\vartheta_{P_1} = \vartheta_{P_1}^*.$$

So, $\mathbf{x}^*$ is a minimizer of (P_1), and the sequence $\{\vartheta_{P_1}^k\}$ converges to the minimum of (P_1). □

Remark 3.3 By Theorem 3.2, we know that Algorithm 3.1 can solve the CP tensor program (1.3). If it is feasible, a CP decomposition of the optimal solution can also be obtained. In particularly, under some general conditions, Algorithm 3.1 converges within finitely many steps (cf. [23,24]).

4 The best CP tensor approximation problem

In this section, we study the best CP tensor approximation problem, which approximates a given symmetric tensor by a CP tensor under linear constraints. We show that the problem can be transformed to a linear program over the cone of moments and the second order cone. A semidefinite algorithm is also presented for it.

Consider the best CP tensor approximation problem, stated as:

$$\begin{aligned}
\min_{\mathcal{X}} \quad & \|\mathcal{X} - \mathcal{A}\| \\
\text{s.t.} \quad & \mathcal{A}_i \bullet \mathcal{X} = b_i, i = 1, \ldots, l_1, \\
& \mathcal{A}_i \bullet \mathcal{X} \geq b_i, i = l_1 + 1, \ldots, l, \\
& \mathcal{X} \in CP_{m,n},
\end{aligned} \tag{4.1}$$

where $\mathcal{A}, \mathcal{A}_i \in S^m(\mathbb{R}^n)$, $b_i \in \mathbb{R}$, $i = 1, \ldots, l$. Compared to the CP tensor program (1.3), the objective function of (4.1) is nonlinear on the variable $\mathcal{X}$.

If there are no linear constraints, (4.1) is reduced to the CP tensor projection problem:

$$\begin{aligned}
\min_{\mathcal{X}} \quad & \|\mathcal{X} - \mathcal{A}\| \\
\text{s.t.} \quad & \mathcal{X} \in CP_{m,n}.
\end{aligned} \tag{4.2}$$

Clearly, (4.2) is always feasible and has a solution. If the minimum is zero, then the projection of $\mathcal{A}$ onto $CP_{m,n}$ is itself, which implies that $\mathcal{A}$ is CP. If the minimum is nonzero, then $\mathcal{A}$ is not CP. So, solving the CP tensor projection problem also provides a way to check whether a tensor is CP or not.

In the following, we show how to solve (4.1). Introducing a nonnegative variable $\gamma \in \mathbb{R}$, we transform (4.1) to the following problem:

$$\begin{aligned}
\min_{\mathcal{X}, \gamma} \quad & \gamma \\
\text{s.t.} \quad & \|\mathcal{X} - \mathcal{A}\| \leq \gamma, \\
& \mathcal{A}_i \bullet \mathcal{X} = b_i, i = 1, \ldots, l_1, \\
& \mathcal{A}_i \bullet \mathcal{X} \geq b_i, i = l_1 + 1, \ldots, l, \\
& \mathcal{X} \in CP_{m,n}.
\end{aligned} \tag{4.3}$$

Let $\mathcal{Y} = \mathcal{X} - \mathcal{A}$. Denote by

$$\mathbf{x} = \text{vech}(\mathcal{X}), \quad \mathbf{y} = \text{vech}(\mathcal{Y}),$$

the identifying vectors of $\mathcal{X}$ and $\mathcal{Y}$, respectively. It is easy to check that $\|\mathcal{X} - \mathcal{A}\| \leq \gamma$ if and only if $(\mathbf{c} \circ \mathbf{y}, \gamma) \in \mathcal{SOC}_{\mathbb{T}}$, where

$$\mathcal{SOC}_{\mathbb{T}} = \{(\mathbf{w}, \gamma) \in \mathbb{R}^{\mathbb{T}} \times \mathbb{R} : \|\mathbf{w}\|_2 \leq \gamma\}$$

is the second-order cone and is self-dual. Then, (4.3) can be formulated as the following conic linear program:

$$\vartheta_{P_2} = \min_{\mathbf{x},\mathbf{y},\gamma} \quad \gamma$$

$(P_2):$
$$\begin{aligned} \text{s.t.} \quad & \mathbf{x} - \mathbf{y} = \text{vech}(\mathcal{A}), \\ & (\mathbf{c}^2 \circ \text{vech}(\mathcal{A}_i))^T \mathbf{x} = b_i, \, i = 1, \ldots, l_1, \\ & (\mathbf{c}^2 \circ \text{vech}(\mathcal{A}_i))^T \mathbf{x} \geq b_i, \, i = l_1 + 1, \ldots, l, \\ & (\mathbf{x}, (\mathbf{c} \circ \mathbf{y}, \gamma)) \in \mathcal{M}_{\mathbb{T},K} \times \mathcal{SOC}_{\mathbb{T}}, \end{aligned}$$

where $\mathcal{M}_{\mathbb{T},K}$ is given by (2.5). The dual problem of (P_2) is

$$\vartheta_{D_2} = \max_{\lambda,\mathbf{s},\mathbf{z}} \quad b^T \lambda + (\mathbf{c}^2 \circ \mathbf{z})^T \text{vech}(\mathcal{A})$$

$(D_2):$
$$\begin{aligned} \text{s.t.} \quad & \sum_{i=1}^{l} \lambda_i \mathbf{c}^2 \circ \text{vech}(\mathcal{A}_i) + \mathbf{s} + \mathbf{c}^2 \circ \mathbf{z} = 0, \\ & \lambda_i \geq 0, \, i = l_1 + 1, \ldots, l, \\ & (\mathbf{s}, (\mathbf{c} \circ \mathbf{z}, 1)) \in \mathcal{P}_{\mathbb{T},K} \times \mathcal{SOC}_{\mathbb{T}}, \end{aligned}$$

where $\mathcal{P}_{\mathbb{T},K}$ is given by (2.7). By weak duality, for all $(\mathbf{x},\mathbf{y},\gamma) \in \mathcal{F}(P_2)$ and $(\lambda,\mathbf{s},\mathbf{z}) \in \mathcal{F}(D_2)$, we have $\vartheta_{P_2} \geq \vartheta_{D_2}$.

By (2.19) and (2.20), for each $k \geq \lceil m/2 \rceil$, the k-th order relaxation of (P_2) can be defined as

$$\vartheta_{P_2}^{k} = \min_{\mathbf{x},\mathbf{y},\gamma,\tilde{\mathbf{x}}} \quad \gamma$$

$(P_2^k):$
$$\begin{aligned} \text{s.t.} \quad & \mathbf{x} - \mathbf{y} = \text{vech}(\mathcal{A}), \\ & (\mathbf{c}^2 \circ \text{vech}(\mathcal{A}_i))^T \mathbf{x} = b_i, \, i = 1, \ldots, l_1, \\ & (\mathbf{c}^2 \circ \text{vech}(\mathcal{A}_i))^T \mathbf{x} \geq b_i, \, i = l_1 + 1, \ldots, l, \\ & \mathbf{x} = \tilde{\mathbf{x}}|_{\mathbb{T}}, \\ & (\tilde{\mathbf{x}}, (\mathbf{c} \circ \mathbf{y}, \gamma)) \in \Gamma_k \times \mathcal{SOC}_{\mathbb{T}}. \end{aligned}$$

The dual problem of (P_2^k) is

$$\vartheta_{D_2}^{k} = \max_{\lambda,\mathbf{s},\mathbf{z}} \quad b^T \lambda + (\mathbf{c}^2 \circ \mathbf{z})^T \text{vech}(\mathcal{A})$$

$(D_2^k):$
$$\begin{aligned} \text{s.t.} \quad & \sum_{i=1}^{l} \lambda_i \mathbf{c}^2 \circ \text{vech}(\mathcal{A}_i) + \mathbf{s} + \mathbf{c}^2 \circ \mathbf{z} = 0, \\ & \lambda_i \geq 0, \, i = l_1 + 1, \ldots, l, \\ & (\mathbf{s}, (\mathbf{c} \circ \mathbf{z}, 1)) \in \Phi_k \times \mathcal{SOC}_{\mathbb{T}}. \end{aligned}$$

Clearly, $\vartheta_{P_2}^{k} \leq \vartheta_{P_2}$ and $\vartheta_{D_2}^{k} \leq \vartheta_{D_2}$ for all k. Suppose that $(\mathbf{x}^{*,k}, \mathbf{y}^{*,k}, \gamma^{*,k}, \tilde{\mathbf{x}}^{*,k})$ is a minimizer of (P_2^k) and $(\lambda^{*,k}, \mathbf{s}^{*,k}, \mathbf{z}^{*,k})$ is a maximizer of (D_2^k). If $\mathbf{x}^{*,k} = \tilde{\mathbf{x}}^{*,k}|_{\mathbb{T}} \in \mathcal{M}_{\mathbb{T},K}$, then $\vartheta_{P_2}^{k} = \vartheta_{P_2}$ and $(\mathbf{x}^{*,k}, \mathbf{y}^{*,k}, \gamma^{*,k})$ is a minimizer of (P_2), i.e., the relaxation (P_2^k) is exact for solving (P_2). In this case, if $\vartheta_{P_2}^{k} = \vartheta_{D_2}^{k}$, then $\vartheta_{D_2}^{k} = \vartheta_{D_2}$ and $(\lambda^{*,k}, \mathbf{s}^{*,k}, \mathbf{z}^{*,k})$ is a maximizer of (D_2). If the relaxation (P_2^k) is infeasible, then (P_2) is also infeasible.

A semidefinite algorithm for solving the best CP tensor approximation problem (4.1) is presented as follows.

Algorithm 4.1

Step 0. Input $\mathcal{A}, \mathcal{A}_i \in \mathbb{S}^m(\mathbb{R}^n)$, $b_i \in \mathbb{R}$, $i = 1, \ldots, l$ and K as (2.2). Let $k := \lceil m/2 \rceil$.

Step 1. Solve the relaxation problem (P_2^k). If (P_2^k) is infeasible, stop and output that (P_2) is infeasible; otherwise, compute an optimal solution $(\mathbf{x}^{*,k}, \mathbf{y}^{*,k}, \gamma^{*,k}, \tilde{\mathbf{x}}^{*,k})$ of (P_2^k). Let $t := 1$.

Step 2. Let $\hat{\mathbf{x}} := \tilde{\mathbf{x}}^{*,k}|_{2t}$. If the rank condition (2.12) is not satisfied, go to Step 4.

Step 3. Compute $\rho_i > 0$, $v^i \in K$ and $r = \text{rank}(M_t(\hat{\mathbf{x}}))$. Output the CP decomposition of $\mathcal{X}^*$ as

$$\mathcal{X}^* = \sum_{i=1}^{r} \rho_i (v^i)^{\otimes m}.$$

Step 4. If $t < k$, set $t := t+1$ and go to Step 2; otherwise, set $k := k+1$ and go to Step 1.

We say (D_2) has a relative interior, if there exists a feasible point pair $(\lambda, \mathbf{s}, \mathbf{z}) \in \mathcal{F}(D_2)$ such that $(\mathbf{s}, (\mathbf{c} \circ \mathbf{z}, 1)) \in \text{int}(\mathcal{P}_{\mathbb{T},K} \times \mathcal{SOC})$ and $\lambda_i > 0$, $i = l_1 + 1, \ldots, l$. The convergence properties of Algorithm 4.1 is similar to those of Algorithm 3.1.

Theorem 4.2 *Let $\mathbb{T}$ and K be as in (2.1) and (2.10), respectively. Suppose (P_2) is feasible and (D_2) has a relative interior point. Algorithm 4.1 has the following properties:*

(i) *For all k sufficiently large, (D_2^k) has a relative interior point and (P_2^k) has a minimizer $(\mathbf{x}^{*,k}, \mathbf{y}^{*,k}, \gamma^{*,k}, \tilde{\mathbf{x}}^{*,k})$.*

(ii) *The sequence $\{(\mathbf{x}^{*,k}, \mathbf{y}^{*,k}, \gamma^{*,k})\}$ is bounded, and each of its accumulation points is a minimizer of (P_2). The sequence $\{\gamma^{*,k}\}$ converges to the minimum of (4.1).*

Proof (i) Let $(\lambda^0, \mathbf{s}^0, \mathbf{z}^0) \in \mathcal{F}(D_2)$ with $\lambda_i^0 > 0 (i = l_1 + 1, \ldots, l)$, and $(\mathbf{s}^0, (\mathbf{c} \circ \mathbf{z}^0, 1)) \in \text{int}(\mathcal{P}_{\mathbb{T},K} \times \mathcal{SOC})$. Then, $\mathbf{s}^0|_K > 0$. Since K is compact, there exist $\epsilon > 0$ and $\delta > 0$ such that

$$\mathbf{s}|_K - \epsilon > \epsilon, \quad \forall \mathbf{s} \in B(\mathbf{s}^0, \delta).$$

By [27, Theorem 6], there exists $N_0 > 0$ such that

$$\mathbf{s} - \epsilon \in I_{2N_0}(h) + Q_{N_0}(g), \quad \forall \mathbf{s} \in B(\mathbf{s}^0, \delta).$$

So, (D_2^k) has a relative interior point for all $k \geq N_0$. Thus, the strong duality holds for (P_2^k) and (D_2^k). As (P_2) is feasible, the relaxation problem (P_2^k) is also feasible. So, (P_2^k) has a minimizer $(\mathbf{x}^{*,k}, \mathbf{y}^{*,k}, \gamma^{*,k}, \tilde{\mathbf{x}}^{*,k})$.

(ii) We first show $\{(\mathbf{x}^{*,k}, \mathbf{y}^{*,k}, \gamma^{*,k})\}$ is bounded. Let $(\lambda^0, \mathbf{s}^0, \mathbf{z}^0)$ and ϵ be as in the proof of (i). The set $I_{2N_0}(h) + Q_{N_0}(g)$ is dual to Γ_{N_0}. For all $k \geq N_0$, we have $\tilde{\mathbf{x}}^{*,k} \in \Gamma_{N_0}$ and

$$0 \leq \langle \mathbf{s}^0 - \epsilon, \tilde{\mathbf{x}}^{*,k} \rangle = \langle \mathbf{s}^0, \tilde{\mathbf{x}}^{*,k} \rangle - \epsilon \langle 1, \tilde{\mathbf{x}}^{*,k} \rangle,$$

$$\langle (\mathbf{s}^0, (\mathbf{c} \circ \mathbf{z}^0, 1)), (\mathbf{x}^{*,k}, (\mathbf{c} \circ \mathbf{y}^{*,k}, \gamma^{*,k})) \rangle \leq \gamma^{*,k} - b^T \lambda^0 - (\mathbf{c}^2 \circ \mathbf{z}^0)^T \text{vech}(\mathcal{A}).$$

Because $\gamma^{*,k} \leq \vartheta_{P_2}$ and

$$\langle \mathbf{s}^0, \tilde{\mathbf{x}}^{*,k} \rangle = \langle \mathbf{s}^0, \mathbf{x}^{*,k} \rangle \leq \langle (\mathbf{s}^0, (\mathbf{c} \circ \mathbf{z}^0, 1)), (\mathbf{x}^{*,k}, (\mathbf{c} \circ \mathbf{y}^{*,k}, \gamma^{*,k})) \rangle,$$

we have

$$\langle \mathbf{s}^0, \tilde{\mathbf{x}}^{*,k} \rangle \leq U_0 := \vartheta_{P_2} - b^T \lambda^0 - (\mathbf{c}^2 \circ \mathbf{z}^0)^T \text{vech}(\mathcal{A}).$$

So,

$$0 \leq \langle \mathbf{s}^0 - \epsilon, \tilde{\mathbf{x}}^{*,k} \rangle \leq U_0 - \epsilon(\tilde{\mathbf{x}}^{*,k})_0,$$

$$(\tilde{\mathbf{x}}^{*,k})_0 \leq U_1 := U_0/\epsilon.$$

Note that $I(h) + Q(g)$ is archimedean, we can further prove that the sequence $\{\mathbf{x}^{*,k}\}$ is bounded. By the definitions of $\mathbf{x}$, $\mathbf{y}$ and γ, we know $\{(\mathbf{x}^{*,k}, \mathbf{y}^{*,k}, \gamma^{*,k})\}$ is bounded.

Suppose $(\mathbf{x}^*, \mathbf{y}^*, \gamma^*)$ is an accumulation point of $\{(\mathbf{x}^{*,k}, \mathbf{y}^{*,k}, \gamma^{*,k})\}$. Without loss of generality, we assume

$$(\mathbf{x}^{*,k}, \mathbf{y}^{*,k}, \gamma^{*,k}) \to (\mathbf{x}^*, \mathbf{y}^*, \gamma^*), \quad k \to \infty.$$

Since $\mathbf{x}^{*,k} \in \Omega_k$, by (2.21) and (2.22), we have $\mathbf{x}^* \in \bigcap_{k=\lceil m/2 \rceil}^{\infty} \Omega_k = \mathcal{M}_{\mathbb{T},K}$. Note that $(\mathbf{x}^{*,k}, \mathbf{y}^{*,k}, \gamma^{*,k}) \in \mathcal{F}(P_2^k)$, we further have $(\mathbf{x}^*, \mathbf{y}^*, \gamma^*) \in \mathcal{F}(P_2)$. So,

$$\vartheta_{P_2} \le \gamma^*. \tag{4.4}$$

Because (P_2^k) is a relaxation problem of (P_2) and $(\mathbf{x}^{*,k}, \mathbf{y}^{*,k}, \gamma^{*,k})$ is a minimizer of (P_2^k), we get

$$\vartheta_{P_2} \ge \gamma^{*,k}, \ k = \lceil m/2 \rceil, \lceil m/2 \rceil + 1, \dots$$

Taking $k \to \infty$, we obtain

$$\vartheta_{P_2} \ge \lim_{k \to \infty} \gamma^{*,k} = \gamma^*. \tag{4.5}$$

This, together with (4.4), implies that

$$\vartheta_{P_2} = \gamma^*.$$

So, $(\mathbf{x}^*, \mathbf{y}^*, \gamma^*)$ is a minimizer of (P_2), and the sequence $\{\gamma^{*,k}\}$ converges to the minimum of (P_2). $\qquad\square$

Remark 4.3 It is worth to point out that, though the objective function of (4.1) is nonlinear, we can transform it to a linear optimization problem with the cone of moments and the second order cone. So, it is essentially a conic linear program.

If (P_2) is feasible, Algorithm 4.1 generally converges within finitely many steps (cf. [23,24]).

5 Numerical experiments

In this section, we present some numerical experiments for solving the CP tensor program (1.3) and the best CP tensor approximation problem (4.1) by Algorithms 3.1 and 4.1, respectively. The CP decompositions of the optimal solutions are also given. We use the softwares GloptiPoly 3 [13] and SeDuMi [34] to solve the semidefinite relaxation problems (P_1^k) and (P_2^k). The experiments are implemented on a laptop with an Intel Core i5-2520M CPU (2.50 GHz) and 8 GB of RAM, using Matlab R2014b. We display 4 decimal digits for numerical numbers.

5.1 CP tensor programs

We use Algorithm 3.1 to solve the CP tensor programs (1.3).

Example 5.1 Consider the tensor $\mathcal{A} \in \mathrm{S}^3(\mathbb{R}^{10})$ given as:

$$\mathcal{A}_{i,j,k} = \sin(i + j + k).$$

Case 1. Consider (1.3) with the CP cone and the linear constraints $\mathcal{A}_i \bullet \mathcal{X} = b_i (i = 1, 2)$, where

$$(\mathcal{A}_1)_{i,j,k} = \frac{(-1)^i}{i} + \frac{(-1)^j}{j} + \frac{(-1)^k}{k}, \quad (\mathcal{A}_2)_{i,j,k} = \frac{i+j+k}{i \times j \times k},$$

$$b_1 = -1, \quad b_2 = 1.$$

We apply Algorithm 3.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 2$ with $\vartheta_{P_1}^k = -12.5084$. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{2} \rho_i (v^i)^{\otimes 3}$, where

$$\rho_1 = 12.2876, \ v^1 = (0.0000, 0.0000, 0.0000, 0.0000, 0.0000,$$
$$0.0000, 0.0000, 0.0000, 0.0314, 0.9995)^T,$$
$$\rho_2 = 6.3969, \ v^2 = (0.0000, 0.0000, 0.0000, 0.0000, 0.0000,$$
$$0.0000, 0.7535, 0.0000, 0.6574, 0.0000)^T.$$

The computation takes about 7 s.

Case 2. Consider (1.3) with the CP cone and the linear constraints $\mathcal{A}_i \bullet \mathcal{X} = b_i (i = 1, 2)$, where $\mathcal{A}_1, \mathcal{A}_2, b_2$ are the same as in Case 1 and

$$b_1 = 1.$$

We apply Algorithm 3.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 2$ with $\vartheta_{P_1}^k = -16.2281$. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{2} \rho_i (v^i)^{\otimes 3}$, where

$$\rho_1 = 5.1872, \ v^1 = (0.0000, 0.0000, 0.0000, 0.0000,$$
$$0.0000, 0.0000, 0.7535, 0.0000, 0.6574, 0.0000)^T,$$
$$\rho_2 = 15.9174, \ v^2 = (0.0000, 0.0000, 0.0000, 0.0000, 0.0000,$$
$$0.0000, 0.0000, 0.0000, 0.0314, 0.9995)^T.$$

It takes about 6 s.

Case 3. Consider (1.3) with the CP cone and the linear constraints $\mathcal{A}_1 \bullet \mathcal{X} \geq b_1, \mathcal{A}_2 \bullet \mathcal{X} = b_2$, where $\mathcal{A}_i, b_i (i = 1, 2)$ are the same as in Case 2. We apply Algorithm 3.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 2$ with $\vartheta_{P_1}^k = -32.9344$. The optimal solution is $\mathcal{X}^* = \rho(v)^{\otimes 3}$, where $\rho = 33.3333$ and

$$v = (0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 1.0000)^T.$$

It takes about 5 s.

Case 4. Consider (1.3) with the CP cone and the linear constraints $\mathcal{A}_i \bullet \mathcal{X} = b_i (i = 1, 2)$, where $\mathcal{A}_1, \mathcal{A}_2, b_1$ are the same as in Case 1 and

$$b_2 = \frac{1}{10}.$$

We apply Algorithm 3.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 2$ and the relaxation problem (P^k) is infeasible, so the problem (1.3) is infeasible. It takes about 5 s.

Example 5.2 Consider the tensor $\mathcal{A} \in S^5(\mathbb{R}^5)$ given as:

$$\mathcal{A}_{i_1, i_2, i_3, i_4, i_5} = \cos(i_1 + i_2 + i_3 + i_4 + i_5).$$

Case 1. Consider (1.3) with the CP cone and the linear constraints $\mathcal{A}_i \bullet \mathcal{X} = b_i (i = 1, 2)$, where

$$(\mathcal{A}_1)_{i_1,i_2,i_3,i_4,i_5} = \log\left(\frac{i_1}{5}\right) + \log\left(\frac{i_2}{5}\right) + \log\left(\frac{i_3}{5}\right) + \log\left(\frac{i_4}{5}\right) + \log\left(\frac{i_5}{5}\right),$$

$$(\mathcal{A}_2)_{i_1,i_2,i_3,i_4,i_5} = \sin(i_1 + i_2 + i_3 + i_4 + i_5)/(\cos(i_1 + i_2 + i_3 + i_4 + i_5) + 1),$$

$$b_1 = -1, \quad b_2 = 1.$$

We apply Algorithm 3.1 and choose $d = 6$ and $k = 3$ in Step 0. It terminates at $k = 3$ with $\vartheta_{P_1}^k = -0.2992$. The optimal solution is $\mathcal{X}^* = \rho(v)^{\otimes 5}$, where $\rho = 0.3858$ and

$$v = (0.0000, 0.0000, 1.0000, 0.0075, 0.0000)^T.$$

It takes about 2 s.

Case 2. Consider (1.3) with the CP cone and the linear constraints $\mathcal{A}_i \bullet \mathcal{X} = b_i (i = 1, 2)$, where $\mathcal{A}_1, \mathcal{A}_2, b_1$ are the same as in Case 1 and

$$b_2 = -1.$$

We apply Algorithm 3.1 and choose $d = 6$ and $k = 3$ in Step 0. It terminates at $k = 3$ with $\vartheta_{P_1}^k = -0.3175$. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{2} \rho_i(v^i)^{\otimes 5}$, where

$$\rho_1 = 0.0091, \quad v^1 = (0.0000, 0.0000, 0.0000, 0.7914, 0.6113)^T,$$

$$\rho_2 = 0.3091, \quad v^2 = (0.0000, 0.0000, 0.9917, 0.1283, 0.0000)^T.$$

It takes about 2 s.

Case 3. Consider (1.3) with the CP cone and the linear constraints $\mathcal{A}_1 \bullet \mathcal{X} \geq b_1, \mathcal{A}_2 \bullet \mathcal{X} = b_2$, where $\mathcal{A}_1, \mathcal{A}_2, b_2$ are the same as in Case 1 and

$$b_1 = -10.$$

We apply Algorithm 3.1 and choose $d = 6$ and $k = 3$ in Step 0. It terminates at $k = 3$ with $\vartheta_{P_1}^k = -3.1521$. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{2} \rho_i(v^i)^{\otimes 5}$, where

$$\rho_1 = 0.0227, \quad v^1 = (0.0000, 0.0000, 0.0000, 0.7912, 0.6116)^T,$$

$$\rho_2 = 3.1225, \quad v^2 = (0.0000, 0.0000, 0.9917, 0.1283, 0.0000)^T.$$

It takes about 2 s.

Case 4. Consider (1.3) with the CP cone and the linear constraints $\mathcal{A}_i \bullet \mathcal{X} = b_i (i = 1, 2)$, where $\mathcal{A}_1, \mathcal{A}_2, b_2$ are the same as in Case 1 and

$$b_1 = 1.$$

We apply Algorithm 3.1 and choose $d = 6$ and $k = 3$ in Step 0. It terminates at $k = 3$ and the relaxation problem (P^k) is infeasible, so the problem (1.3) is infeasible. It takes about 2 s.

5.2 CP projection problems

We compute the CP projection of a given tensor by applying Algorithm 4.1 to (4.2).

Example 5.3 Consider the tensor $\mathcal{A} \in S^4(\mathbb{R}^5)$ given as:

$$\mathcal{A}_{i_1, i_2, i_3, i_4} = \sin(i_1 + i_2 + i_3 + i_4).$$

We apply Algorithm 4.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 4$ with $\gamma^{*,k} = 17.3001$. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{5} \rho_i (v^i)^{\otimes 5}$, where

$$\rho_1 = 0.9230, \ v^1 = (0.0000, 0.0000, 0.0000, 0.0448, 0.9990)^T,$$

$$\rho_2 = 0.1135, \ v^2 = (0.9082, 0.0000, 0.0000, 0.0000, 0.4185)^T,$$

$$\rho_3 = 2.3795, \ v^3 = (0.0000, 0.0254, 0.6951, 0.7153, 0.0674)^T,$$

$$\rho_4 = 0.0067, \ v^4 = (0.0033, 0.0817, 0.5878, 0.7758, 0.0387)^T,$$

$$\rho_5 = 2.4899, \ v^5 = (0.4513, 0.8000, 0.3947, 0.0002, 0.0000)^T.$$

It takes about 45 s.

Example 5.4 Consider the tensor $\mathcal{A} \in S^5(\mathbb{R}^5)$ given as:

$$\mathcal{A} = \sum_{k=1}^{7} (u^k)^{\otimes 5}, \tag{5.1}$$

where u^k is generated randomly as follows:

$$(u^1, u^2, u^3, u^4, u^5, u^6, u^7)$$
$$= \begin{pmatrix} -0.3627 & 0.3481 & -0.2680 & 0 & 0 & 0.5319 & -0.3411 \\ 0.8941 & -0.0326 & 0 & 0 & -0.0996 & -0.0122 & -0.7011 \\ -0.4760 & 0 & -0.2260 & -0.3299 & 0.1715 & 0.2673 & -0.3013 \\ -0.1280 & 0.8290 & 0.2555 & 0 & 0 & 0.6222 & 0 \\ -0.2392 & 0.0860 & 0 & 0.0009 & -0.8341 & 0.4437 & -0.6253 \end{pmatrix}.$$

We apply Algorithm 4.1 and choose $d = 6$ and $k = 3$ in Step 0. It terminates at $k = 5$ with $\gamma^{*,k} = 2.1140$. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{3} \rho_i (v^i)^{\otimes 5}$, where

$$\rho_1 = 0.3601, \ v^1 = (0.3131, 0.0000, 0.0000, 0.9497, 0.0000)^T,$$

$$\rho_2 = 0.9694, \ v^2 = (0.5285, 0.0000, 0.1938, 0.7357, 0.3767)^T,$$

$$\rho_3 = 0.4020, \ v^3 = (0.0000, 1.0000, 0.0000, 0.0000, 0.0000)^T.$$

It takes about 30 s.

Example 5.5 Consider the tensor $\mathcal{A} \in S^3(\mathbb{R}^{10})$ given as (cf. [30]):

$\mathcal{A}_{2,2,2} = 4, \mathcal{A}_{2,2,3} = 1, \mathcal{A}_{2,2,4} = 1, \mathcal{A}_{2,2,5} = 1, \mathcal{A}_{2,2,8} = 1, \mathcal{A}_{2,3,3} = 1,$
$\mathcal{A}_{2,3,8} = 1, \mathcal{A}_{2,4,4} = 1, \mathcal{A}_{2,4,5} = 1, \mathcal{A}_{2,5,5} = 1, \mathcal{A}_{2,8,8} = 1, \mathcal{A}_{3,3,3} = 6,$
$\mathcal{A}_{3,3,4} = 1, \mathcal{A}_{3,3,5} = 1, \mathcal{A}_{3,3,7} = 1, \mathcal{A}_{3,3,8} = 2, \mathcal{A}_{3,4,4} = 1, \mathcal{A}_{3,4,5} = 1,$
$\mathcal{A}_{3,5,5} = 1, \mathcal{A}_{3,7,7} = 1, \mathcal{A}_{3,7,8} = 1, \mathcal{A}_{3,8,8} = 2, \mathcal{A}_{4,4,4} = 7, \mathcal{A}_{4,4,5} = 2,$
$\mathcal{A}_{4,4,7} = 1, \mathcal{A}_{4,4,9} = 1, \mathcal{A}_{4,4,10} = 1, \mathcal{A}_{4,5,5} = 2, \mathcal{A}_{4,7,7} = 1, \mathcal{A}_{4,7,9} = 1,$
$\mathcal{A}_{4,9,9} = 1, \mathcal{A}_{4,10,10} = 1, \mathcal{A}_{5,5,5} = 4, \mathcal{A}_{7,7,7} = 4, \mathcal{A}_{7,7,8} = 1, \mathcal{A}_{7,7,9} = 1,$
$\mathcal{A}_{7,8,8} = 1, \mathcal{A}_{7,9,9} = 1, \mathcal{A}_{8,8,8} = 6, \mathcal{A}_{8,8,9} = 1, \mathcal{A}_{8,8,10} = 1, \mathcal{A}_{8,9,9} = 1,$
$\mathcal{A}_{8,9,10} = 1, \mathcal{A}_{8,10,10} = 1, \mathcal{A}_{9,9,9} = 4, \mathcal{A}_{9,9,10} = 1, \mathcal{A}_{9,10,10} = 1, \mathcal{A}_{10,10,10} = 3,$

and the other entries are zero, except permutations of the above indices. $\mathcal{A}$ is a strongly symmetric hierarchically dominated nonnegative tensor, so it is completely positive. A CP decomposition of length 15 of $\mathcal{A}$ is also given in [30].

We apply Algorithm 4.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 3$ with $\gamma^{*,k} = 3.2576 \times 10^{-7}$. Thus $\mathcal{A}$ is CP. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{14} \rho_i (v^i)^{\otimes 3}$, where

$$\rho_1 = 1.9999, \ v^1 = (0.0000, 0.0000, 0.0000, 0.0000, 0.0000,$$
$$0.0000, 0.0000, 0.0000, 0.0000, 1.0000)^T,$$

$$\rho_2 = 1.0000, \ v^2 = (0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 1.0000,$$
$$0.0000, 0.0000, 0.0000, 0.0000)^T,$$

$$\rho_3 = 1.0000, \ v^3 = (0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000,$$
$$0.0000, 1.0000, 0.0000, 0.0000)^T,$$

$$\rho_4 = 2.5647, \ v^4 = (0.0000, 0.4015, 0.9158, 0.0000, 0.0000, 0.0000,$$
$$0.0000, 0.0000, 0.0000, 0.0000)^T,$$

$$\rho_5 = 5.1962, \ v^5 = (0.0000, 0.5774, 0.0000, 0.0000, 0.0000, 0.0000,$$
$$0.0000, 0.5774, 0.0000, 0.5774)^T,$$

$$\rho_6 = 3.0406, \ v^6 = (0.0000, 0.9768, 0.2141, 0.0000, 0.0000, 0.0000,$$
$$0.0000, 0.0000, 0.0000, 0.0000)^T,$$

$$\rho_7 = 1.0000, \ v^7 = (0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000,$$
$$1.0000, 0.0000, 0.0000, 0.0000)^T,$$

$$\rho_8 = 2.8284, \ v^8 = (0.7071, 0.0000, 0.0000, 0.0000, 0.7071, 0.0000,$$
$$0.0000, 0.0000, 0.0000, 0.0000)^T,$$

$$\rho_9 = 5.1961, \ v^9 = (0.0000, 0.0000, 0.5774, 0.5774, 0.5773, 0.0000,$$
$$0.0000, 0.0000, 0.0000, 0.0000)^T,$$

$$\rho_{10} = 2.5647, \ v^{10} = (0.0000, 0.0000, 0.0000, 0.0000, 0.9158,$$
$$0.0000, 0.0000, 0.0000, 0.4015, 0.0000)^T,$$

$$\rho_{11} = 1.0000, \ v^{11} = (0.0000, 0.0000, 0.0000, 1.0000, 0.0000,$$
$$0.0000, 0.0000, 0.0000, 0.0000, 0.0000)^T,$$

$$\rho_{12} = 5.1961, \ v^{12} = (0.0000, 0.5774, 0.0000, 0.0000, 0.0000,$$
$$0.5774, 0.0000, 0.0000, 0.5774, 0.0000)^T,$$

$$\rho_{13} = 5.1962, \ v^{13} = (0.0000, 0.0000, 0.0000, 0.0000, 0.0000,$$
$$0.0000, 0.5774, 0.0000, 0.5774, 0.5774)^T,$$

$$\rho_{14} = 3.0406, \ v^{14} = (0.0000, 0.0000, 0.0000, 0.0000, 0.2141,$$
$$0.0000, 0.0000, 0.0000, 0.9768, 0.0000)^T.$$

We get a different CP decomposition from the one given in [30]. It takes about 4369 s.

5.3 Best CP approximation problems

In the following, we use Algorithm 4.1 to solve the best CP approximation problems (4.1).

Example 5.6 Consider the tensor $\mathcal{A} \in \mathrm{S}^4(\mathbb{R}^5)$ given as:

$$\mathcal{A}_{i_1,i_2,i_3,i_4} = \tan(i_1) + \tan(i_2) + \tan(i_3) + \tan(i_4).$$

Case 1. Consider (4.1) with the CP cone and the linear constraints $\mathcal{A}_i \bullet \mathcal{X} = b_i (i = 1, 2)$, where

$$\mathcal{A}_1 = \mathcal{I}, \quad (\mathcal{A}_2)_{i_1,i_2,i_3,i_4} = \sin(i_1 + i_2 + i_3 + i_4),$$
$$b_1 = 3, \quad b_2 = 10.$$

We apply Algorithm 4.1 and choose $d = 4$ and $k = 2$ in Step 0. $k = 4$ with $\gamma^{*,k} = 108.2775$. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{3} \rho_i (v^i)^{\otimes 4}$, where

$$\rho_1 = 6.0525, \quad v^1 = (0.0000, 0.0000, 0.6694, 0.7429, 0.0000)^T,$$
$$\rho_2 = 28.0878, \quad v^2 = (0.5717, 0.0666, 0.4256, 0.6983, 0.0000)^T,$$
$$\rho_3 = 0.8052, \quad v^3 = (0.0000, 1.0000, 0.0000, 0.0000, 0.0000)^T.$$

It takes about 21 s.

Case 2. Consider (4.1) with the CP cone and the linear constraints $\mathcal{A}_1 \bullet \mathcal{X} \geq b_1$, $\mathcal{A}_2 \bullet \mathcal{X} = b_2$, where $\mathcal{A}_i, b_i$, $(i = 1, 2)$ are the same as in Case 1. We apply Algorithm 4.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 3$ with $\gamma^{*,k} = 107.7770$. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{3} \rho_i (v^i)^{\otimes 4}$, where

$$\rho_1 = 0.2832, \quad v^1 = (0.0000, 1.0000, 0.0000, 0.0000, 0.0000)^T,$$
$$\rho_2 = 9.6388, \quad v^2 = (0.0000, 0.0000, 0.6396, 0.7687, 0.0000)^T,$$
$$\rho_3 = 31.6701, \quad v^3 = (0.6973, 0.0778, 0.3804, 0.6025, 0.0000)^T.$$

It takes about 18 s.

Case 3. Consider (4.1) with the CP cone and the linear constraints $\mathcal{A}_i \bullet \mathcal{X} \geq b_i (i = 1, 2)$, where $\mathcal{A}_1, \mathcal{A}_2, b_2$ are the same as in Case 1 and

$$b_1 = 8.$$

We apply Algorithm 4.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 3$ with $\gamma^{*,k} = 107.7815$. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{3} \rho_i (v^i)^{\otimes 4}$, where

$$\rho_1 = 0.2632, \quad v^1 = (0.0000, 1.0000, 0.0000, 0.0000, 0.0000)^T,$$
$$\rho_2 = 32.0446, \quad v^2 = (0.7069, 0.0783, 0.3765, 0.5937, 0.0000)^T$$
$$\rho_3 = 9.9369, \quad v^3 = (0.0000, 0.0000, 0.6376, 0.7704, 0.0000)^T.$$

It takes about 3 s.

Case 4. Consider (4.1) with the CP cone and the linear constraints $\mathcal{A}_i \bullet \mathcal{X} = b_i (i = 1, 2)$, where $\mathcal{A}_1, \mathcal{A}_2, b_2$ are the same as in Case 1 and

$$b_1 = -3.$$

We apply Algorithm 4.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 2$ and the relaxation problem (P^k) is infeasible, so the problem is infeasible. It takes about 2 s.

Example 5.7 Consider the tensor $\mathcal{A} \in S^3(\mathbb{R}^6)$ given as:

$$\mathcal{A} = \sum_{k=1}^{6} (u^k)^{\otimes 3}, \qquad (5.2)$$

where u^k is generated as:

$$(u^1, u^2, u^3, u^4, u^5, u^6) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Case 1. Consider (4.1) with the CP cone and the linear constraints $\mathcal{A}_i \bullet \mathcal{X} = b_i (i = 1, 2)$, where

$$(\mathcal{A}_1)_{i,j,k} = \frac{(-1)^i}{i} + \frac{(-1)^j}{j} + \frac{(-1)^k}{k}, \quad (\mathcal{A}_2)_{i,j,k} = \frac{i + j + k}{i \times j \times k},$$
$$b_1 = 5, \quad b_2 = 10.$$

We apply Algorithm 4.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 3$ with $\gamma^{*,k} = 3.2829$. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{9} \rho_i (v^i)^{\otimes 3}$, where

$$\rho_1 = 1.5410, \ v^1 = (0.3416, 0.9398, 0.0000, 0.0000, 0.0000, 0.0000)^T,$$
$$\rho_2 = 1.6540, \ v^2 = (0.3656, 0.0000, 0.0000, 0.0000, 0.0000, 0.9308)^T,$$
$$\rho_3 = 2.4480, \ v^3 = (0.0000, 0.0000, 0.6431, 0.7658, 0.0000, 0.0000)^T,$$
$$\rho_4 = 2.6323, \ v^4 = (0.0000, 0.0000, 0.0000, 0.7372, 0.6757, 0.0000)^T,$$
$$\rho_5 = 2.4737, \ v^5 = (0.0000, 0.7412, 0.6713, 0.0000, 0.0000, 0.0000)^T,$$
$$\rho_6 = 0.1779, \ v^6 = (0.0000, 0.6077, 0.5324, 0.5892, 0.0000, 0.0000)^T,$$
$$\rho_7 = 2.4998, \ v^7 = (0.0000, 0.0000, 0.0000, 0.0000, 0.7107, 0.7035)^T,$$
$$\rho_8 = 0.7884, \ v^8 = (0.0000, 0.5992, 0.0000, 0.5593, 0.2090, 0.5336)^T,$$
$$\rho_9 = 0.0660, \ v^9 = (0.0000, 0.6189, 0.0000, 0.0000, 0.3647, 0.6956)^T.$$

It takes about 8 s.

Case 2. Consider (4.1) with the CP cone and the linear constraints $\mathcal{A}_i \bullet \mathcal{X} = b_i (i = 1, 2)$, where $\mathcal{A}_1, \mathcal{A}_2, b_2$ are the same as in Case 1 and

$$b_1 = -5.$$

We apply Algorithm 4.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 3$ with $\gamma^{*,k} = 2.8225$. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{7} \rho_i (v^i)^{\otimes 3}$, where

$$\rho_1 = 0.8638, \ v^1 = (0.5276, 0.8495, 0.0000, 0.0000, 0.0000, 0.0000)^T,$$
$$\rho_2 = 2.1761, \ v^2 = (0.5882, 0.0000, 0.0000, 0.0000, 0.0000, 0.8088)^T,$$
$$\rho_3 = 2.6494, \ v^3 = (0.0144, 0.0000, 0.0000, 0.0000, 0.7504, 0.6608)^T,$$
$$\rho_4 = 2.5492, \ v^4 = (0.0000, 0.0000, 0.0000, 0.6849, 0.7286, 0.0000)^T,$$
$$\rho_5 = 2.0366, \ v^5 = (0.0000, 0.5931, 0.8051, 0.0000, 0.0000, 0.0000)^T,$$

$$\rho_6 = 0.0612, \quad v^6 = (0.0000, 0.0000, 0.7018, 0.0001, 0.7123, 0.0000)^T,$$

$$\rho_7 = 2.5081, \quad v^7 = (0.0000, 0.0000, 0.7285, 0.6850, 0.0000, 0.0000)^T.$$

It takes about 8 s.

Case 3. Consider (4.1) with the CP cone and the linear constraints $\mathcal{A}_1 \bullet \mathcal{X} \geq b_1$, $\mathcal{A}_2 \bullet \mathcal{X} = b_2$, where $\mathcal{A}_i, b_i (i = 1, 2)$ are the same as in Case 2. We apply Algorithm 4.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 3$ with $\gamma^{*,k} = 2.6451$. The optimal solution is $\mathcal{X}^* = \sum_{i=1}^{6} \rho_i (v^i)^{\otimes 3}$, where

$$\rho_1 = 1.2712, \quad v^1 = (0.4401, 0.8979, 0.0000, 0.0000, 0.0000, 0.0000)^T,$$

$$\rho_2 = 2.2492, \quad v^2 = (0.0000, 0.6644, 0.7474, 0.0000, 0.0000, 0.0000)^T,$$

$$\rho_3 = 2.6408, \quad v^3 = (0.0000, 0.0000, 0.0000, 0.7128, 0.7014, 0.0000)^T,$$

$$\rho_4 = 2.5398, \quad v^4 = (0.0000, 0.0000, 0.6976, 0.7165, 0.0000, 0.0000)^T,$$

$$\rho_5 = 2.6490, \quad v^5 = (0.0000, 0.0000, 0.0000, 0.0000, 0.7342, 0.6789)^T,$$

$$\rho_6 = 2.0055, \quad v^6 = (0.5197, 0.0000, 0.0000, 0.0000, 0.0000, 0.8543)^T.$$

It takes about 8 s.

Case 4. Consider (4.1) with the CP cone and the linear constraints $\mathcal{A}_i \bullet \mathcal{X} = b_i (i = 1, 2)$, where $\mathcal{A}_1, \mathcal{A}_2, b_1$ are the same as in Case 2 and

$$b_2 = \frac{1}{10}.$$

We apply Algorithm 4.1 and choose $d = 4$ and $k = 2$ in Step 0. It terminates at $k = 2$ and the relaxation problem (P^k) is infeasible, so the problem is infeasible. It takes about 0.5 s.

References

1. Berman, A., Shaked-Monderer, N.: Completely Positive Matrices. World Scientific, Singapore (2003)
2. Burer, S.: On the copositive representation of binary and continuous nonconvex quadratic programs. Math. Program., Ser. A **120**, 479–495 (2009)
3. Cichocki, A., Zdunek, R., Phan, A.H., Amari, S.: Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multiway Data Analysis and Blind Source Separation. Wiley, New York (2009)
4. Comon, P.: Tensor Decompositions: State of the Art and Applications, Mathematics in Signal Processing, V (Coventry, 2000), Institute for Mathematics and its Apllications Conference Series, vol. 71, pp. 1–24. Oxford University Press, Oxford (2002)
5. Comon, P., Golub, G., Lim, L.H., Mourrain, B.: Symmetric tensors and symmetric tensor rank. SIAM J. Matrix Anal. Appl. **30**, 1254–1279 (2008)
6. Cui, C., Dai, Y., Nie, J.: All real eigenvalues of symmetric tensors. SIAM J. Matrix Anal. Appl. **35**, 1582–1601 (2014)
7. Curto, R., Fialkow, L.: Truncated K-moment problems in several variables. J. Oper. Theory **54**, 189–226 (2005)
8. Dickinson, P.J., Gijben, L.: On the computational complexity of membership problems for the completely positive cone and its dual. Comput. Optim. Appl. **57**, 403–415 (2014)
9. Fan, J., Zhou, A.: A semidefinite algorithm for completely positive tensor decomposition. Comput. Optim. Appl. (2016). https://doi.org/10.1007/s10589-016-9870-9
10. Fan, J., Zhou, A.: The CP-matrix approximation problem. SIAM J. Matrix Anal. Appl. **37**, 171–194 (2016)
11. Fialkow, L., Nie, J.: The truncated moment problem via homogenization and flat extensions. J. Funct. Anal. **263**, 1682–1700 (2012)
12. Henrion, D., Lasserre, J.: Detecting global optimality and extracting solutions in GloptiPoly, positive polynomials in control. In: Lecture Notes in Control and Information Sciences, vol. 312, pp. 293–310. Springer, Berlin (2005)

13. Henrion, D., Lasserre, J., Loefberg, J.: GloptiPoly 3: moments, optimization and semidefinite programming. Optim. Methods Softw. **24**, 761–779 (2009)
14. Hu, Shenglong, Huang, Zheng-Hai, Qi, Liqun: Strictly nonnegative tensors and nonnegative tensor partition. Sci. China Math. **57**, 181–195 (2014)
15. Kuang, X., Zuluaga, L.: Completely positive and completely positive semidefinite tensor relaxations for polynomial optimization. J. Glob. Optim. **70**, 551–577 (2018)
16. Kolda, Tamara G., Bader, Brett W.: Tensor decompositions and applications. SIAM Rev. **51**, 455–500 (2009)
17. Kolda, Tamara G., Mayo, J.R.: Shifted power method for computing tensor eigenpairs. SIAM J. Matrix Anal. Appl. **32**, 1095–1124 (2011)
18. Kolda, Tamara G.: Numerical optimization for symmetric tensor decomposition. Math. Program., Ser. B **151**, 225–248 (2015)
19. Lasserre, J.B.: Moments, Positive Polynomials and Their Applications. Imperial College Press, London (2009)
20. Laurent, M.: Sums of squares, moment matrices and optimization over polynomials. In: Putinar, M., Sullivant, S. (eds.) Emerging Applications of Algebraic Geometry, IMA Volumes in Mathematics and its Applications, vol. 149, pp. 157–270. Springer, Berlin (2009)
21. Luo, Z., Qi, L.: Completely positive tensors: properties, easily checkable subclasses, and tractable relaxations. SIAM J. Matrix Anal. Appl. **37**, 1675–1698 (2016)
22. Nie, J.: The $\mathcal{A}$-truncated K-moment problem. Found. Comput. Math. **14**, 1243–1276 (2014)
23. Nie, J.: Linear optimization with cones of moments and nonnegative polynomials. Math. Program., Ser. B **153**, 247–274 (2015)
24. Nie, J.: Optimality conditions and finite convergence of Lasserre's hierarchy. Math. Program., Ser. A **146**, 97–121 (2014)
25. Nie, J., Wang, L.: Semidefinite relaxations for best rank-1 tensor approximations. SIAM J. Matrix Anal. Appl. **35**, 1155–1179 (2014)
26. Nie, J.: Generating polynomials and symmetric tensor decompositions. Found. Comput. Math. **17**, 423–465 (2017)
27. Nie, J., Schweighofer, M.: On the complexity of Putinar's Positivstellensatz. J. Complex. **23**, 135–150 (2007)
28. Peña, J., Vera, J.C., Zuluaga, L.F.: Completely positive reformulations for polynomial optimization. Math. Program., Ser. B **151**, 405–431 (2015)
29. Putinar, M.: Positive polynomials on compact semi-algebraic sets. Indiana Univ. Math. J. **42**, 969–984 (1993)
30. Qi, L., Xu, C., Xu, Y.: Nonnegative tensor factorization, completely positive tensors and an hierarchical elimination algorithm. SIAM J. Matrix Anal. Appl. **35**, 1227–1241 (2014)
31. Qi, L.: Symmetric nonnegative tensors and copositive tensors. Linear Algebra Appl. **439**, 228–238 (2013)
32. Shashua, A., Hazan, T.: Non-negative tensor factorization with applications to statistics and computer vision. In: ACM International Conference Proceeding Series: Proceedings of the 22nd International Conference on Machine Learning, vol. 119, pp. 792–799 (2005)
33. Shashua, A., Zass, R., Hazan, T.: Multiway clustering using supersymmetric nonnegative tensor factorization. In: Proceedings of the European Conference on Computer Vision, pp. 595–608 (2006)
34. Sturm, J.F.: SeDuMi 1.02: A MATLAB toolbox for optimization over symmetric cones. Optim. Methods Softw. **11 & 12**, 625–653 (1999)
35. Xu, C., Luo, Z., Qi, L., Chen, Z.: {0, 1} completely positive tensors and multi-hypergraphs. Linear Algebra Appl. **510**, 110–123 (2016)
36. Zhou, A., Fan, J.: The CP-matrix completion problem. SIAM J. Matrix Anal. Appl. **35**, 127–142 (2014)

Springer