

# A successive approximation method for quantum separability

Deren HAN<sup>1</sup>, Liqun QI<sup>2</sup>

<sup>1</sup> School of Mathematical Sciences, Jiangsu Key Laboratory for NSLSCS, Nanjing Normal University, Nanjing 210023, China

<sup>2</sup> Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China

© Higher Education Press and Springer-Verlag Berlin Heidelberg 2013

**Abstract** Determining whether a quantum state is separable or inseparable (entangled) is a problem of fundamental importance in quantum science and has attracted much attention since its first recognition by Einstein, Podolsky and Rosen [Phys. Rev., 1935, 47: 777] and Schrödinger [Naturwissenschaften, 1935, 23: 807–812, 823–828, 844–849]. In this paper, we propose a successive approximation method (SAM) for this problem, which approximates a given quantum state by a so-called separable state: if the given state is separable, this method finds its rank-one components and the associated weights; otherwise, this method finds the distance between the given state to the set of separable states, which gives information about the degree of entanglement in the system. The key task per iteration is to find a feasible descent direction, which is equivalent to finding the largest M-eigenvalue of a fourth-order tensor. We give a direct method for this problem when the dimension of the tensor is 2 and a heuristic cross-hill method for cases of high dimension. Some numerical results and experiences are presented.

**Keywords** Quantum system, entanglement, tensor, successive approximation, M-eigenvalue, cross-hill

**MSC** 15A18, 15A69

## 1 Introduction

Quantum entanglement, first recognized by Einstein et al. [5] and Schrödinger [22], plays a key role in many of the most interesting applications of quantum computation and quantum information. Entangled states are useful in

---

Received July 14, 2011; accepted December 12, 2012

Corresponding author: Deren HAN, E-mail: handeren@njnu.edu.cn

quantum cryptography, quantum teleportation, and quantum key distribution [13]. These applications rely on the fact that the quantum entanglement implies the existence of pure entangled states which produce nonclassical phenomena. Thus, it is of fundamental importance to know which states, usually mixed rather than pure ones in laboratory due to the uncontrolled interaction with the environment, are separable or entangled. Here, we say that a given state  $\varrho$  acting on a finite-dimensional product Hilbert space  $\mathcal{H}_1 \otimes \mathcal{H}_2$  is separable if it can be written as a convex linear combination of product density operators, i.e.,

$$\varrho = \sum_n p_n \varrho_n^{(1)} \otimes \varrho_n^{(2)}, \quad (1)$$

where, for all  $n$ ,  $p_n$  are nonnegative coefficients satisfying the condition  $\sum_n p_n = 1$  and  $\varrho_n^{(i)}$  are pure states on  $\mathcal{H}_i$ ,  $i = 1, 2$ .

Despite its wide-spread importance, it is a very complicated problem to determine whether a quantum state is separable or entangled. Recently, Gurvits proved that this problem is NP-hard [6], and there does not exist an efficient, general solution method. The efficient tests for separability, which based on the partial transposition and was first introduced by Peres [14], are ‘one-sided’ in the sense that they are either necessary conditions, or sufficient conditions for separability, but not both. They are sufficient and necessary only for some special cases [9]. There have also been some attempts to more general approaches to the identification of separability by use of the natural geometrical structure of the problem, where the Euclidean metric (Hilbert-Schmidt or Frobenius norm) of the Hermitian matrices can be used to give a geometrical characterization of the set of separable matrices [15,23].

Some schemes to test separability numerically have also been suggested, using modern technique from convex optimization. Specially, using the relation between positive partial transpose of the state and the semidefinite programs, Doherty et al. [4] proposed to distinguish separable and entangled states via solving a series of semidefinite programs, which can be solved efficiently via interior point methods. If a state  $\varrho$  on  $\mathcal{H}_1 \otimes \mathcal{H}_2$  is separable, then there is an extension  $\tilde{\varrho}$  on  $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_1$ , such that  $\pi \tilde{\varrho} \pi = \varrho$  ( $\pi$  is a projector onto the symmetric subspace), and the partial transpose of  $\tilde{\varrho}$  is positive. Thus, it generates a hierarchy of necessary conditions for separability. First, it tests the positivity of the partial transpose for a bipartite density matrix  $\varrho$ . If the test fails, the state is entangled; otherwise, the state could be separable or entangled and the test goes on to the next step, testing the positivity of an extension of the current state. Since each test in the sequence is at least as powerful as the previous one, the results become more and more accurate.

Recently, Ioannou et al. [10] proposed another numerical method, i.e., analytical cutting-plane method, to detect quantum separability and entanglement. Their method is based on the fact that a state  $\varrho$  is entangled if and only if there exists an entanglement witness detecting it. An entanglement witness is a traceless operator  $A \in \mathcal{H}_{M,N}$ , for which there

exists a state  $\varrho \in \mathcal{D}_{M,N}$ , such that

$$\mathrm{tr}(A\sigma) < \mathrm{tr}(A\varrho), \quad \forall \sigma \in \mathcal{S}_{M,N},$$

where  $\mathcal{H}_{M,N}$  denotes the vector space of all Hermitian operators acting on  $\mathcal{H}_1 \otimes \mathcal{H}_2$ ,  $M$  and  $N$  are dimensions of  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively,  $\mathcal{S}_{M,N}$  is the set of all separable states in  $\mathcal{H}_{M,N}$ , and  $\mathcal{D}_{M,N}$  denotes the set of all states in  $\mathcal{H}_{M,N}$ . Initially, the potential witnesses are contained in the ball

$$\mathcal{W} := \{A \in \mathcal{H}_{M,N} : \mathrm{tr}(A) = 0, \mathrm{tr}(A^2) \leq 1\}.$$

Then, they find the entanglement witness by reducing  $\mathcal{W}$  step by step: if  $\varrho$  is entangled, then any element in the set of the last step is an entanglement witness; otherwise, the set will be too small to contain an entanglement witness and the state is separable.

In some applications, it is also important to find the closest separable state. In fact, if a given state is separable, then the closest separable state is itself; otherwise, if the given state is entangled, then the distance between it and the closest separable state can be used as a measure of the degree of entanglement. In [3], Dahl et al. firstly considered this problem. After giving some geometrical description of separability, they proposed a Frank-Wolfe method for solving such a problem.

In this paper, in the same theme as [3], we propose a feasible descent method for finding the closest separable state of the given state. At each iteration, we need to find the largest M-eigenvalues of a fourth-order tensor. If the largest M-eigenvalue is nonpositive, then the given state is separable; otherwise, the M-eigenvector corresponding to the largest M-eigenvalue can form a feasible descent direction. Along such a direction, we can find a closer separable state than the current one and the method goes on. Eventually, we either find that the given state is separable or we find the closest separable state.

This paper is organized as follows. In the next section, we give the mathematical description of the problem formally and summarize some necessary preliminaries. In Section 3, we propose our feasible descent method and prove its convergence. The main task per iteration is to find the largest M-eigenvalue and the associated M-eigenvector of a fourth-order tensor. We thus pay our attention to solving such problems in Section 4: in Section 4.1, we describe a direct method for the case of dimension 2; in Section 4.2, we propose a direct method for the special case with biquadrate coefficient; and in Section 4.3, a heuristic cross-hill method is described. Some numerical results are reported in Section 5, where comparison with the method in [3] is made. Finally, we conclude the paper in Section 6.

## 2 Problem description and preliminaries

Mathematically, the mixed quantum states are described in terms of density matrices, which are positive semidefinite with trace one. A density matrix that

is a projection on a single vector is referred to as representing a pure state, while other density matrices represent mixed states.

Hereafter, we discuss in detail only the case of real vectors and real matrices. This will simplify our discussion and will not make our results loss of any generality due to the close correspondence between transposition of real matrices and conjugation of complex matrices. In particular, the set of real symmetric matrices is expanded to the set of Hermitian matrices, which is still a real vector space with the same positive definite real inner product  $\text{tr}(AB)$ , where  $\text{tr}(C)$  denotes the trace of a matrix  $C$ , equaling to the sum of its diagonal elements.

Let  $\mathbb{R}^n$  denote the Euclidean space of real vectors of length  $n$ , which is equipped with the standard inner product  $\langle \cdot, \cdot \rangle$  and the Euclidean norm  $\| \cdot \|$ . Vectors are treated as column vectors, and the transpose of a vector  $x$  is denoted by  $x^\top$ . The convex hull of a set  $S$  is the intersection of all convex sets containing  $S$ , and is denoted by  $\text{conv}(S)$ . We use  $I$  to denote the identity matrix of order  $n$  and  $n$  should be clear from the context.

Let  $\varphi^n$  denote the space of all real symmetric  $n \times n$  matrices. For any matrix  $A \in \varphi^n$ , we denote its  $(i, j)$ -th element with  $a_{ij}$ ,  $i, j = 1, \dots, n$ .  $\varphi^n$  is equipped with the standard inner product

$$\langle A, B \rangle = \text{tr}(AB) = \sum_{i,j=1}^n a_{ij}b_{ij}, \quad A, B \in \varphi^n.$$

The associated matrix norm is the Frobenius norm

$$\|A\|_F = \langle A, A \rangle^{1/2} = (\text{tr}(A^2))^{1/2} = \left( \sum_{i,j=1}^n a_{ij}^2 \right)^{1/2}.$$

A matrix  $A \in \varphi^n$  is positive semidefinite if for any  $x \in \mathbb{R}^n$ , we have  $x^\top Ax \geq 0$ . We use  $A \succeq 0$  to denote that  $A$  is positive semidefinite. All symmetric positive semidefinite matrices of order  $n$  form the positive semidefinite cone

$$\varphi_+^n = \{A \in \varphi^n : A \succeq 0\}.$$

A density matrix is a matrix in  $\varphi_+^n$  with trace 1. Let  $\mathcal{J}_+^n$  denote the set of all density matrices of order  $n$ :

$$\mathcal{J}_+^n = \{A \in \varphi_+^n : \text{tr}(A) = 1\}.$$

Then the set  $\mathcal{J}_+^n$  is convex and has the following properties.

**Lemma 1** [3, Theorem 2.1] *The set  $\mathcal{J}_+^n$  of density matrices has the following properties:*

(i)  $\mathcal{J}_+^n$  is the intersection of the positive semidefinite cone  $\varphi_+^n$  and the hyperplane

$$H = \{A \in \varphi^n : \langle A, I \rangle = 1\}.$$

- (ii)  $\mathcal{J}_+^n$  is a compact convex set of dimension  $\frac{n(n+1)}{2} - 1$ .
- (iii) The extreme points of  $\mathcal{J}_+^n$  are the symmetric rank one matrices  $A = xx^\top$ , where  $x \in \mathbb{R}^n$  satisfies  $\|x\| = 1$ . Therefore,

$$\mathcal{J}_+^n = \text{conv}\{xx^\top : x \in \mathbb{R}^n, \|x\| = 1\}.$$

If  $A \in \mathbb{R}^{p \times p}$  and  $B \in \mathbb{R}^{q \times q}$ , then the tensor product  $A \otimes B$  is the square matrix of order  $pq$  given by its  $(i, j)$ -th block  $a_{ij}B$  ( $1 \leq i, j \leq p$ ). A matrix  $A \in \mathbb{R}^{n \times n}$  is separable if  $A$  can be written as a convex combination

$$A = \sum_{j=1}^N \lambda_j B_j \otimes C_j$$

for some positive integer  $N$ , matrices  $B_j \in \mathcal{J}_+^p, C_j \in \mathcal{J}_+^q, n = pq$ , and non-negative numbers  $\lambda_j$  satisfying  $\sum_{j=1}^N \lambda_j = 1$ . Let  $\mathcal{J}_+^{n, \otimes}$  denote the set of all separable matrices of order  $n = pq$ . For sets  $U$  and  $W$  of matrices, we use  $U \otimes W$  to denote the set of matrices that can be written as the tensor product of a matrix in  $U$  and a matrix in  $W$ . The following lemma summarizes some important properties of  $\mathcal{J}_+^{n, \otimes}$ .

**Lemma 2** [3, Theorem 2.2] *The set  $\mathcal{J}_+^{n, \otimes}$  of separable matrices have the following properties:*

- (i)  $\mathcal{J}_+^{n, \otimes} \subseteq \mathcal{J}_+^n$ .
- (ii)  $\mathcal{J}_+^{n, \otimes}$  is a compact convex set and

$$\mathcal{J}_+^{n, \otimes} = \text{conv}(\mathcal{J}_+^p \otimes \mathcal{J}_+^q).$$

- (iii) The extreme points of  $\mathcal{J}_+^{n, \otimes}$  are the symmetric rank one matrices

$$A = (x \otimes y)(x \otimes y)^\top,$$

where  $x \in \mathbb{R}^p$  and  $y \in \mathbb{R}^q$  both have Euclidean length one and

$$\mathcal{J}_+^{n, \otimes} = \text{conv}\{(x \otimes y)(x \otimes y)^\top : \|x\| = 1, \|y\| = 1\}.$$

The problem under consideration thus can be formulated as follows. For a given density matrix  $A \in \mathcal{J}_+^n$ , find a separable density matrix  $X \in \mathcal{J}_+^{n, \otimes}$  that minimizes the distance  $\|X - A\|_F$ . Thus, we need to solve the following optimization problem:

$$f(A) := \min \frac{1}{2} \|X - A\|_F^2, \tag{2}$$

$$\text{s.t. } X \in \mathcal{J}_+^{n, \otimes}. \tag{3}$$

Note that  $A$  is separable if and only if  $f(A) = 0$ . If  $A$  is separable, then by solving (2)-(3), we can find its separable components; otherwise, if  $A$  is entangled, we can find the degree on its entanglement by finding its distance to the separable set.

### 3 A feasible descent direction method

Note that the optimization problem (2)-(3) is a convex optimization problem, where the objective function (2) is a strictly convex quadratic function of  $X$  and the constraint set  $\mathcal{J}_+^{n,\otimes}$  is a compact and convex set. Thus, it has a unique solution [1], denoted as  $X^*$ , which is essentially the projection of  $A$  onto  $\mathcal{J}_+^{n,\otimes}$ . That is,

$$X^* = \text{Proj}_{\otimes}(A).$$

The projector  $\text{Proj}_{\otimes}(X)$  of a point  $X \in \mathcal{J}_+^n$  onto  $\mathcal{J}_+^{n,\otimes}$  has the following properties.

**Lemma 3** *Let  $A \in \mathcal{J}_+^n$  and  $X \in \mathcal{J}_+^{n,\otimes}$ . Then the following statements are equivalent:*

- (i)  $X = \text{Proj}_{\otimes}(A)$ ;
- (ii)  $\langle A - X, Y - X \rangle \leq 0, \forall Y \in \mathcal{J}_+^{n,\otimes}$ ;
- (iii) *the following inequality holds:*

$$\langle A - X, Y - X \rangle \leq 0, \quad \forall Y \in \text{Ext}(\mathcal{J}_+^{n,\otimes}), \quad (4)$$

where

$$\text{Ext}(\mathcal{J}_+^{n,\otimes}) := \{(x \otimes y)(x \otimes y)^\top : \|x\| = 1, \|y\| = 1\}$$

denotes the set of extreme points of  $\mathcal{J}_+^{n,\otimes}$ .

The optimization problem (2)-(3), although it is a convex optimization problem, is very difficult to solve, due to the difficulty in description the constraint set  $\mathcal{J}_+^{n,\otimes}$ . Lemma 3 (iii) states that to solve (2)-(3), we need only to consider the linear variational inequality problem of finding  $X \in \mathcal{J}_+^{n,\otimes}$ , such that (4) holds. Note that the first term in the inner product in (4),  $A - X$ , is nothing else but the minus gradient of the objective function of the optimization problem (2)-(3), and  $Y - X$  forms a feasible descent direction provided that it violates (4). Thus, we have the following algorithm.

#### Algorithm 3.1 A Feasible Descent Direction Method

**Step 1** Set an initial point  $X^0 \in \mathcal{J}_+^{n,\otimes}$  and set the number of iteration  $k = 0$ .

**Step 2** Check if  $X^k$  is the optimal solution, that is, check if (4) holds: if it holds, then the current iterative point  $X^k$  is the optimal solution of (2)-(3); otherwise, any  $Y^k \in \text{Ext}(\mathcal{J}_+^{n,\otimes})$  that violates it provides a feasible descent direction  $Y^k - X^k$  at  $X^k$ .

**Step 3** Let

$$g(\alpha) := \|A - X^k + \alpha(Y^k - X^k)\|_F^2,$$

and let  $\alpha_k$  be the optimal solution of the problem

$$\min g(\alpha), \quad \text{s.t.} \quad 0 \leq \alpha \leq 1. \quad (5)$$

Then, update the iterative point

$$X^{k+1} = X^k + \alpha_k(Y^k - X^k).$$

**Step 4** Set  $k := k + 1$  and go to Step 1.

Algorithm 3.1 is in fact a feasible descent direction method for the optimization problem (2)-(3), or its equivalent problem of linear variational inequality problem of finding  $X^* \in \mathcal{J}_+^{n,\otimes}$ , such that

$$\langle A - X^*, Y - X^* \rangle \leq 0, \quad \forall Y \in \mathcal{J}_+^{n,\otimes},$$

or (4). Since the function  $g$  is a quadratic function in variable  $\alpha$ , the solution of (5) can be found analytically. Thus, the method is in fact a feasible descent direction method with exact line search. As we have observed, the optimization problem (2)-(3) is a convex optimization problem, where the objective function (2) is strictly convex and the feasible set  $\mathcal{J}_+^{n,\otimes}$  is compact and convex, guaranteeing the existence and uniqueness of the optimal solution  $X^*$ . Thus, according to the convergence theorem in [1, Chapter 2], the sequence of iterative points generated by the feasible descent direction method converges to  $X^*$  globally, i.e., we have the following result.

**Theorem 1** *Algorithm 3.1 generates a sequence of matrices  $\{X^k\}$  that converges to the optimal solution of (2)-(3).*

#### 4 Finding a feasible descent direction

The most important task per iteration of the feasible descent direction method, Algorithm 3.1, is to find the feasible descent direction  $Y^k - X^k$ . Let

$$\gamma(X) := \max\{\langle A - X, Y - X \rangle : Y \in \text{Ext}(\mathcal{J}_+^{n,\otimes})\}. \quad (6)$$

At iteration  $k$ , we evaluate  $\gamma(X^k)$  to find if it is negative. If it is negative, then (4) holds and according to Lemma 3.1,  $X^k$  is already the optimal solution of (2)-(3) and the algorithm stops; otherwise, we find  $Y^k$  and  $\gamma(X^k) > 0$ , and  $Y^k - X^k$  is a descent direction of the objective function (2). The optimization problem in evaluating  $\gamma$  is a polynomial optimization problem, which is generally nonconvex and difficult to solve. In the rest of this section, we propose a direct method for solving it in Section 4.1, for the special case that  $p = q = 2$ . Then, in Section 4.2, we propose a direct method for the special case with biquadrate coefficient. Finally, in Section 4.3, we propose a heuristic cross-hill method, based on the algorithms in Section 4.1 and Section 4.2 for the special cases, to solve the general problem.

Let  $B := A - X^k$  and omit the constant term. Then evaluating  $\gamma$  is equivalent to solving the following optimization problem:

$$\max \langle B, Y \rangle, \quad (7)$$

$$\text{s.t. } Y \in \text{Ext}(\mathcal{J}_+^{n,\otimes}). \quad (8)$$

Partition  $Y$  to the following  $p \times p$  blocks:

$$Y = (x \otimes y)(x \otimes y)^\top = (x \otimes x^\top)(y \otimes y^\top) = \begin{bmatrix} Y_{11} & \cdots & Y_{1p} \\ \vdots & & \vdots \\ Y_{p1} & \cdots & Y_{pp} \end{bmatrix},$$

where

$$Y_{ij} = x_i x_j (y y^\top) \in \mathbb{R}^{q \times q} \quad (i, j \leq p).$$

Then, partition  $B$  conformably as

$$B = \begin{bmatrix} B_{11} & \cdots & B_{1p} \\ \vdots & & \vdots \\ B_{p1} & \cdots & B_{pp} \end{bmatrix},$$

and denote  $B_{ij}^{kl}$  ( $1 \leq i, j \leq p$ ,  $1 \leq k, l \leq q$ ) as the  $(k, l)$ -th element of block  $B_{ij}$ . Let  $\mathcal{C}$  be a fourth-order tensor, whose elements are

$$\mathcal{C}_{ijkl} = B_{ij}^{kl}.$$

From the symmetry of the matrix  $B$ ,  $\mathcal{C}$  has the following symmetric property:

$$\mathcal{C}_{ijkl} = \mathcal{C}_{jikl} = \mathcal{C}_{ijlk}, \quad 1 \leq i, j \leq p, 1 \leq k, l \leq q.$$

The optimization problem (7)-(8) can be rewritten as

$$\max g(x, y) := \sum_{i,j=1}^p \sum_{k,l=1}^q \mathcal{C}_{ijkl} x_i x_j y_k y_l, \quad (9)$$

$$\text{s.t. } \|x\|^2 = \|y\|^2 = 1. \quad (10)$$

Let  $\lambda$  and  $\mu$  be Lagrange multiplier to the constraint

$$\|x\|^2 = 1, \quad \|y\|^2 = 1,$$

respectively. The optimality condition of (9)-(10) is

$$\begin{cases} \sum_{j=1}^p \sum_{k,l=1}^q \mathcal{C}_{ijkl} x_j y_k y_l = \lambda x_i, & i = 1, \dots, p, \\ \sum_{i,j=1}^p \sum_{k=1}^q \mathcal{C}_{ijkl} x_i x_j y_k = \mu y_l, & l = 1, \dots, q, \\ x^\top x = 1, \\ y^\top y = 1. \end{cases} \quad (11)$$

Suppose that  $\lambda, \mu, x$ , and  $y$  satisfy (11). It is easy to see that

$$\lambda = \sum_{i,j=1}^p \sum_{k,l=1}^q B_{ij}^{kl} x_i x_j y_k y_l = \mu.$$

Thus, we may rewrite (11) as

$$\begin{cases} \sum_{j=1}^p \sum_{k,l=1}^q \mathcal{C}_{ijk} x_j y_k y_l = \lambda x_i, & i = 1, \dots, p, \\ \sum_{i,j=1}^p \sum_{k=1}^q \mathcal{C}_{ijk} x_i x_j y_k = \lambda y_l, & l = 1, \dots, q, \\ x^\top x = 1, \\ y^\top y = 1. \end{cases} \quad (12)$$

If  $\lambda \in \mathbb{R}$ ,  $x \in \mathbb{R}^p$ , and  $y \in \mathbb{R}^q$  satisfy (12), we call  $\lambda$  an M-eigenvalue of  $\mathcal{C}$ , and call  $x$  and  $y$  left and right M-eigenvectors of  $\mathcal{C}$ , associated with the M-eigenvalue  $\lambda$ , respectively. It is easy to see that M-eigenvalues always exist and by comparing all the M-eigenvalues, we can check the negativity of  $\gamma$ ; the M-eigenvectors corresponding to the largest M-eigenvalue can form a  $Y^k$ , and consequently, a feasible descent direction can be constructed.

The concept of M-eigenvalues and M-eigenvectors was introduced in [18] to study the strong ellipticity of a given tensor, see [7,11] for its recent developments. It has close relations to the concepts of Z-eigenvalues [16,17] and D-eigenvalues [21]. The Z-eigenvalues have found applications in global optimization problems of polynomials [12,19,20] and the D-eigenvalues have applications in diffusion kurtosis imaging (DKI) in biomedical [8].

#### 4.1 A direct method for dimension 2

In this subsection, we present a direct method to find all the M-eigenvalues and the associated M-eigenvectors. The global optimal objective value of (9) can be obtained by comparing these M-eigenvalues. The key idea here is to reduce the five variables system (12) to a system involving only two variables. Then, for this system of two variables, we may use the Sylvester formula of the resultant to find the solutions.

We have the following result.

**Theorem 2** *We have the following results on the M-eigenvalues and their corresponding M-eigenvector pairs.*

- (a) *If  $\mathcal{C}_{1112} = \mathcal{C}_{1121} = 0$ , then  $\lambda = \mathcal{C}_{1111}$  is an M-eigenvalue of  $\mathcal{A}$  and the corresponding M-eigenvector pair is  $x = y = (1, 0)^\top$ .*
- (b) *For any real roots  $(u, v)^\top$  of equations*

$$\begin{cases} \mathcal{C}_{1121}u^2 + (\mathcal{C}_{2121} - \mathcal{C}_{1111})uv - \mathcal{C}_{1121}v^2 = 0, \\ \mathcal{C}_{1112}u^2 + 2\mathcal{C}_{1122}uv + \mathcal{C}_{2122}v^2 = 0, \end{cases} \quad (13)$$

$$\lambda = \mathcal{C}_{1111}u^2 + 2\mathcal{C}_{1121}uv + \mathcal{C}_{2121}v^2$$

is an  $M$ -eigenvalue with the corresponding eigenvector pair

$$x = \frac{(u, v)^\top}{\sqrt{u^2 + v^2}}, \quad y = (\pm 1, 0)^\top.$$

(c) For any real roots  $(u, v)^\top$  of equations

$$\begin{cases} \mathcal{C}_{1121}u^2 + 2\mathcal{C}_{1122}uv + \mathcal{C}_{1222}v^2 = 0, \\ \mathcal{C}_{1112}u^2 + (\mathcal{C}_{1212} - \mathcal{C}_{1111})uv - \mathcal{C}_{1112}v^2 = 0, \end{cases} \quad (14)$$

$$\lambda = \mathcal{C}_{1111}u^2 + 2\mathcal{C}_{1112}uv + \mathcal{C}_{1212}v^2$$

is an  $M$ -eigenvalue with the corresponding eigenvector pair

$$x = (\pm 1, 0)^\top, \quad y = \frac{(u, v)^\top}{\sqrt{u^2 + v^2}}.$$

(d)

$$\lambda = \sum_{i,j=1}^p \sum_{k,l=1}^q \mathcal{C}_{ijkl}x_i x_j y_k y_l$$

is an  $M$ -eigenvalue and

$$x = \pm \frac{(u, 1)^\top}{\sqrt{u^2 + 1}}, \quad y = \pm \frac{(v, 1)^\top}{\sqrt{v^2 + 1}} \quad (15)$$

constitute an  $M$ -eigenvector pair, where  $u$  and  $v$  are real solutions of the following system of polynomial equations:

$$\begin{cases} \mathcal{C}_{1121}u^2v^2 + 2\mathcal{C}_{1122}u^2v + \mathcal{C}_{1222}u^2 + (\mathcal{C}_{2121} - \mathcal{C}_{1111})uv^2 - \mathcal{C}_{1121}v^2 \\ \quad + 2(\mathcal{C}_{2122} - \mathcal{C}_{1112})uv + (\mathcal{C}_{2222} - \mathcal{C}_{1212})u - 2\mathcal{C}_{1122}v - \mathcal{C}_{1222} = 0, \\ \mathcal{C}_{1112}u^2v^2 + (\mathcal{C}_{1212} - \mathcal{C}_{1111})u^2v - \mathcal{C}_{1112}u^2 + 2(\mathcal{C}_{1222} - \mathcal{C}_{1121})uv \\ \quad + 2\mathcal{C}_{1122}uv^2 - 2\mathcal{C}_{1122}u + \mathcal{C}_{2122}v^2 + (\mathcal{C}_{2222} - \mathcal{C}_{2121})v - \mathcal{C}_{2122} = 0. \end{cases} \quad (16)$$

All the  $M$ -eigenvalues and the associated  $M$ -eigenvector pairs are given by (a), (b), (c), and (d) if  $\mathcal{C}_{1112} = \mathcal{C}_{1121} = 0$ , and by (b), (c), and (d) otherwise.

*Proof* (a) If  $\mathcal{C}_{1112} = \mathcal{C}_{1121} = 0$ , then it is direct to check that (a) holds.

(b) If  $y_2 = 0$ , then  $y_1 = \pm 1$  and (12) becomes

$$\begin{cases} \mathcal{C}_{1111}x_1 + \mathcal{C}_{1121}x_2 = \lambda x_1, \\ \mathcal{C}_{1121}x_1 + \mathcal{C}_{2121}x_2 = \lambda x_2, \\ \mathcal{C}_{1111}x_1^2 + 2\mathcal{C}_{1121}x_1x_2 + \mathcal{C}_{2121}x_2^2 = \lambda, \\ \mathcal{C}_{1112}x_1^2 + 2\mathcal{C}_{1122}x_1x_2 + \mathcal{C}_{2122}x_2^2 = 0, \\ x_1^2 + x_2^2 = 1. \end{cases}$$

Eliminating  $\lambda$  from the first two equations, we get

$$\begin{cases} \mathcal{C}_{1121}x_1^2 + (\mathcal{C}_{2121} - \mathcal{C}_{1111})x_1x_2 - \mathcal{C}_{1121}x_2^2 = 0, \\ \mathcal{C}_{1112}x_1^2 + 2\mathcal{C}_{1122}x_1x_2 + \mathcal{C}_{2122}x_2^2 = 0, \\ x_1^2 + x_2^2 = 1. \end{cases}$$

Let

$$u = \frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \quad v = \frac{x_2}{\sqrt{x_1^2 + x_2^2}}.$$

Then the results of (b) follow immediately.

(c) It can be proved in a similar way as (b).

(d) If  $x_2 \neq 0$  and  $y_2 \neq 0$ , (12) becomes

$$\begin{cases} \mathcal{C}_{1111}x_1y_1^2 + 2\mathcal{C}_{1112}x_1y_1y_2 + \mathcal{C}_{1212}x_1y_2^2 \\ + \mathcal{C}_{1121}x_2y_1^2 + 2\mathcal{C}_{1122}x_2y_1y_2 + \mathcal{C}_{1222}x_2y_2^2 = \lambda x_1, \\ \mathcal{C}_{1121}x_1y_1^2 + 2\mathcal{C}_{1122}x_1y_1y_2 + \mathcal{C}_{1222}x_1y_2^2 \\ + \mathcal{C}_{2121}x_2y_1^2 + 2\mathcal{C}_{2122}x_2y_1y_2 + \mathcal{C}_{2222}x_2y_2^2 = \lambda x_2, \\ \mathcal{C}_{1111}x_1^2y_1 + \mathcal{C}_{1112}x_1^2y_2 + 2\mathcal{C}_{1121}x_1x_2y_1 \\ + 2\mathcal{C}_{1122}x_1x_2y_2 + \mathcal{C}_{2121}x_2^2y_1 + \mathcal{C}_{2122}x_2^2y_2 = \lambda y_1, \\ \mathcal{C}_{1112}x_1^2y_1 + \mathcal{C}_{1212}x_1^2y_2 + 2\mathcal{C}_{1122}x_1x_2y_1 \\ + 2\mathcal{C}_{1222}x_1x_2y_2 + \mathcal{C}_{2122}x_2^2y_1 + \mathcal{C}_{2222}x_2^2y_2 = \lambda y_2, \\ x_1^2 + x_2^2 = 1, \\ y_1^2 + y_2^2 = 1. \end{cases}$$

Let

$$u = \frac{x_1}{x_2}, \quad v = \frac{y_1}{y_2}.$$

Then from the first two equalities of the above system, we have

$$\begin{aligned} &\mathcal{C}_{1121}u^2v^2 + 2\mathcal{C}_{1122}u^2v + \mathcal{C}_{1222}u^2 + (\mathcal{C}_{2121} - \mathcal{C}_{1111})uv^2 - \mathcal{C}_{1121}v^2 \\ &+ 2(\mathcal{C}_{2122} - \mathcal{C}_{1112})uv + (\mathcal{C}_{2222} - \mathcal{C}_{1212})u - 2\mathcal{C}_{1122}v - \mathcal{C}_{1222} = 0, \end{aligned}$$

and from the third and the fourth equalities, we have

$$\begin{aligned} &\mathcal{C}_{1112}u^2v^2 + (\mathcal{C}_{1212} - \mathcal{C}_{1111})u^2v - \mathcal{C}_{1112}u^2 + 2(\mathcal{C}_{1222} - \mathcal{C}_{1121})uv \\ &+ 2\mathcal{C}_{1122}uv^2 - 2\mathcal{C}_{1122}u + \mathcal{C}_{2122}v^2 + (\mathcal{C}_{2222} - \mathcal{C}_{2121})v - \mathcal{C}_{2122} = 0. \end{aligned}$$

Combining the above two equalities and the assumption that

$$x_1^2 + x_2^2 = 1, \quad y_1^2 + y_2^2 = 1,$$

we get the assertion immediately. □

From Theorem 2, we can see that to find all the M-eigenvalues and the associated M-eigenvector pairs, we need to solve some systems of polynomial equations with two variables. To solve such systems, we can use the resultant method from algebraic geometry. For example, to solve (16), we may regard it as equations of  $u$ :

$$\begin{cases} \alpha_0 u^2 + \alpha_1 u + \alpha_2 = 0, \\ \beta_0 u^2 + \beta_1 u + \beta_2 = 0, \end{cases} \quad (17)$$

where

$$\begin{aligned} \alpha_0 &= \mathcal{C}_{1121}v^2 + 2\mathcal{C}_{1122}v + \mathcal{C}_{1222}, \\ \alpha_1 &= (\mathcal{C}_{2121} - \mathcal{C}_{1111})v^2 + 2(\mathcal{C}_{2122} - \mathcal{C}_{1112})v + (\mathcal{C}_{2222} - \mathcal{C}_{1212}), \\ \alpha_2 &= -(\mathcal{C}_{1121}v^2 + 2\mathcal{C}_{1122}v + \mathcal{C}_{1222}), \end{aligned} \quad (18)$$

and

$$\begin{aligned} \beta_0 &= \mathcal{C}_{1112}v^2 + \mathcal{C}_{1212}v - \mathcal{C}_{1112}, \\ \beta_1 &= 2\mathcal{C}_{1122}v^2 + 2(\mathcal{C}_{1222} - \mathcal{C}_{1121})v - 2\mathcal{C}_{1122}, \\ \beta_2 &= \mathcal{C}_{2122}v^2 + (\mathcal{C}_{2222} - \mathcal{C}_{2121})v - \mathcal{C}_{2122}. \end{aligned} \quad (19)$$

System (17) has solutions if and only if its resultant vanishes [2]. By the Sylvester theorem [2], its resultant can be calculated as a determinant of the following  $4 \times 4$  matrix:

$$\begin{pmatrix} \alpha_0 & \alpha_1 & \alpha_2 & 0 \\ 0 & \alpha_0 & \alpha_1 & \alpha_2 \\ \beta_0 & \beta_1 & \beta_2 & 0 \\ 0 & \beta_0 & \beta_1 & \beta_2 \end{pmatrix},$$

which is a polynomial of  $v$ . We can use Matlab to find all its complex roots and select the real roots. After this, we substitute them into (16) to find all the real solutions of  $u$ . And correspondingly, all the M-eigenvalues and the associated M-eigenvector pairs can be found.

#### 4.2 A direct method for biquadrate case

In this subsection, we consider the case that  $\mathcal{C}$  is a biquadrate tensor. A fourth-order tensor  $\mathcal{A}$  is called a biquadrate tensor if

$$\mathcal{A}_{ijkl} = 0, \quad (i - j)^2 + (k - l)^2 > 0. \quad (20)$$

Note that for  $\mathcal{C}_{ijkl} = B_{ij}^{kl}$ ,  $\mathcal{C}$  is a biquadrate tensor if and only if  $B$  is a diagonal matrix.

The following theorem presents a direct method for finding all M-eigenvalues and the associated M-Eigenvectors of the biquadrate tensor  $\mathcal{C}$ .

**Theorem 3** *Suppose that  $\mathcal{C}$  is a biquadrate tensor. Then, all the M-eigenvalues and the M-eigenvectors of  $\mathcal{C}$  can be obtained by solving the*

following linear system of equalities and inequalities:

$$\left\{ \begin{array}{ll} \sum_{k \in S_q} \mathcal{C}_{ijkl} \bar{y}_k = \lambda, & i \in S_p, \\ \sum_{i \in S_p} \mathcal{C}_{ijkl} \bar{x}_i = \lambda, & k \in S_q, \\ \bar{x}_i = 0, & i \notin S_p, \\ \bar{x}_i \geq 0, & i \in S_p, \\ \bar{y}_k = 0, & k \notin S_q, \\ \bar{y}_k \geq 0, & k \in S_q, \end{array} \right. \quad (21)$$

where  $S_p \subseteq \{1, 2, \dots, p\}$  and  $S_q \subseteq \{1, 2, \dots, q\}$  are nonempty subsets of integers from 1 to  $p$  and 1 to  $q$ , respectively. Solving (21) for each subset  $S_p$  and  $S_q$  with  $|S_p| \geq 1$  and  $|S_q| \geq 1$ , we find all the M-eigenvalues and the associated M-eigenvectors.

*Proof* Since  $\mathcal{C}$  is a biquadrate tensor, it follows from (20) that (9)-(10) reduces to

$$\begin{aligned} & \max \sum_{i=1}^p \sum_{k=1}^q \mathcal{C}_{iikk} x_i^2 y_k^2, \\ & \text{s.t. } \|x\|^2 = \|y\|^2 = 1. \end{aligned}$$

Consequently, all M-eigenvectors and M-eigenvalues should satisfy

$$\left\{ \begin{array}{ll} \sum_{k=1}^q \mathcal{C}_{iikk} x_i y_k^2 = \lambda x_i, & i = 1, \dots, p, \\ \sum_{i=1}^p \mathcal{C}_{iikk} x_i^2 y_k = \lambda y_k, & k = 1, \dots, q, \\ x^\top x = 1, \\ y^\top y = 1. \end{array} \right. \quad (22)$$

Letting

$$\bar{x}_i = x_i^2 \ (i = 1, \dots, p), \quad \bar{y}_k = y_k^2 \ (k = 1, \dots, q),$$

it is easy to see that the system of equations (22) is equivalent to (21) and the conclusion follows.  $\square$

### 4.3 A heuristic cross-hill method

In this subsection, we give a heuristic cross-hill method for solving the optimization problem (9)-(10), when  $p \geq 3$  or  $q \geq 3$ .

Suppose that we have found a local maximizer of (9)-(10):

$$x = x^{(1)}, \quad y = y^{(1)}.$$

We may find  $p-1$  unit vectors  $x^{(i)}$  of  $\mathbb{R}^p$  and  $q-1$  unit vectors  $y^{(j)}$  of  $\mathbb{R}^q$  such that  $\{x^{(i)}: i = 1, \dots, p\}$  and  $\{y^{(j)}: j = 1, \dots, q\}$  constitute the orthogonal bases of  $\mathbb{R}^p$  and  $\mathbb{R}^q$ , respectively. Then, for  $i = 2, \dots, p$  and  $j = 2, \dots, q$ , we can restrict problem (9)-(10) on the plane spanned by  $x$  and  $x^{(i)}$  and  $y$  and  $y^{(j)}$ , which is now a problem with dimension two, and the current local maximizer  $x$  and  $y$  of (9)-(10) is also a local maximizer of the restricted problem. By Theorem 2, we can use the direct methods in Section 4.1 and Section 4.2 to find the other local maximizers of the restricted problem and select one, denoted as  $s^{(i)}$ ,  $t^{(j)}$ , with the largest objective function value, for  $i = 2, \dots, p$  and  $j = 2, \dots, q$ . Since  $s^{(i)}$ ,  $t^{(j)}$  may not be a local maximizer of the original problem (9)-(10), and  $x, y$  and  $s^{(i)}$ ,  $t^{(j)}$  are separated by a ‘hill’ of the objective function value in two-dimensional case,  $s^{(i)}$ ,  $t^{(j)}$  is a good starting point for a conventional ascent optimization method, in the sense that starting from  $s^{(i)}$ ,  $t^{(j)}$ , we will find a local maximizer  $w, u$  of (9)-(10), which is different from  $x, y$ . Continue this procedure until no new maximizers can be found. The details are summarized in the following algorithm.

**Algorithm 4.1 A Cross-Hill Algorithm.**

**Step 1** Set an initial point  $x^0 \in \mathbb{R}^p$ ,  $y^0 \in \mathbb{R}^q$  with  $\|x^0\| = \|y^0\| = 1$  and set the number of iteration  $k = 0$ .

**Step 2** Use an ascent optimization method M to find a local maximizer  $x = x^{(1)}$ ,  $y = y^{(1)}$  such that  $g(x, y) \geq g(x^0, y^0)$ . Let  $P = \{x^{(1)}\}$  and  $Q = \{y^{(1)}\}$ .

**Step 3** Find  $p-1$  unit vectors  $x^{(i)} \in \mathbb{R}^p$ ,  $i = 2, \dots, p$ , and  $q-1$  unit vectors  $y^{(j)} \in \mathbb{R}^q$ ,  $j = 2, \dots, q$ , such that  $\{x^{(i)}: i = 1, \dots, p\}$  and  $\{y^{(j)}: j = 1, \dots, q\}$  constitute the orthogonal bases of  $\mathbb{R}^p$  and  $\mathbb{R}^q$ , respectively.

**Step 4** For  $i = 2, \dots, p$ , and  $j = 2, \dots, q$ , consider the restricted problem of (9)-(10) on the plane spanned by  $x$  and  $x^{(i)}$  and  $y$  and  $y^{(j)}$ , which is a problem with dimension two. Used the direct method in Section 4.1 to find another local maximizer  $s^{(i)}$ ,  $t^{(j)}$ . Then, use Algorithm M to find a local maximizer  $w^{(i)}$ ,  $u^{(j)}$  of (9)-(10) such that  $g(w^{(i)}, u^{(j)}) \geq g(s^{(i)}, t^{(j)})$ .

**Step 5** For each local maximizer  $\{w^{(i)}, u^{(j)}\}$  found in Step 4, if it is not in  $P, Q$ , then add it to  $P, Q$ , and put it as  $x^{(1)}, y^{(1)}$ , and repeat Steps 3 and 4, until no new maximizers can be found.

**Step 6** Compare the function value of  $g$  for all local maximizers found in  $P, Q$ . The point with largest value of  $g$  is the solution found by this algorithm.

Clearly, Algorithm 4.1 terminates in a finite number of iterations provided the local maximizers of (9)-(10) is finite.

## 5 Numerical results

In this section, we present some numerical results for the feasible descent method. The purpose of the first two examples is to compare the effectiveness of our direct method in Section 4.1 and the heuristic method in Section 4.3 to

the alternating maximum eigenvalue method in [3]. Thus, the first example is with dimension 2 and the second example is with dimension 3. In Section 5.3, we test the whole feasible descent direction method, where we use the direct method described in Section 4.1 to solve subproblem (9)-(10) for the case of dimension 2 and use the heuristic cross-hill method for general case, to find a feasible descent direction, when the current iteration is not a solution of the problem.

### 5.1 Example for $p = q = 2$

In this section, we consider the case of dimension 2, i.e.,  $p = q = 2$ . For such a case, we can use the direct method described in Section 4.1 to solve subproblem (9)-(10). The purpose of this example is to show that, for such a case, we can always find the global solution and the optimum objective value of (9)-(10), while the alternating maximum eigenvalues method in [3] usually traps into local critical points.

Let

$$A = \begin{bmatrix} 0.4691 & 0.1203 & -0.1203 & 0.4691 \\ 0.1203 & 0.0309 & -0.0309 & 0.1203 \\ -0.1203 & -0.0309 & 0.0309 & -0.1203 \\ 0.4691 & 0.1203 & -0.1203 & 0.4691 \end{bmatrix},$$

and let

$$X_0 = \begin{bmatrix} 0.4633 & 0.3097 & 0.1532 & 0.1276 \\ 0.3097 & 0.2168 & 0.1276 & 0.0363 \\ 0.1532 & 0.1276 & 0.1002 & -0.0534 \\ 0.1276 & 0.0363 & -0.0534 & 0.2197 \end{bmatrix}$$

be an arbitrarily selected initial point. Then,

$$B = \begin{bmatrix} 0.0058 & -0.1894 & -0.2736 & 0.3415 \\ -0.1894 & -0.1859 & -0.1585 & 0.0841 \\ -0.2736 & -0.1585 & -0.0693 & -0.0669 \\ 0.3415 & 0.0841 & -0.0669 & 0.2494 \end{bmatrix}.$$

Using the alternating maximum eigenvalues method in [3], we find that

$$x^* = (0.1470, 0.9891)^\top, \quad y^* = (-0.1026, 0.9947)^\top$$

with the objective function value 0.2689.

Using our direct method described in Section 4.1, we find all the M-eigenvalues and the corresponding M-eigenvectors associated to the fourth-order 2-dimensional tensor, which are listed in Table 1.

From Table 1, we can see that the true global optimum should be

$$x^* = (-0.7494, 0.6621)^\top, \quad y^* = (-0.8889, 0.4581)^\top$$

with the objective function value 0.3610 and the critical point got by the alternating maximum eigenvalue method in [3] is not a global one, which is exactly the eighth M-eigenvector and the corresponding M-eigenvalues in Table 1.

Table 1 M-eigenvalues and eigenvectors of  $\mathcal{C}$ 

	$x_1$	$x_2$	$y_1$	$y_2$	$\lambda$
1	0.6680	0.7441	0.9977	0.0681	-0.3098
2	-0.7494	0.6621	-0.8889	0.4581	0.3610
3	0.9390	0.3438	0.8350	0.5503	-0.2560
4	-0.2789	0.9603	0.8152	0.5792	-0.0061
5	-0.9830	0.1834	0.4633	0.8862	-0.3157
6	0.9307	0.3657	-0.5118	0.8591	-0.0045
7	-0.1783	0.9840	-0.4170	0.9089	0.2533
8	0.1470	0.9891	-0.1026	0.9947	0.2689

### 5.2 Results for $p = q = 3$

In this section, we consider the case of dimension 3, i.e.,  $p = q = 3$ . For such a case, we can use the heuristic cross-hill method described in Section 4.3 to find a feasible descent direction per iteration. The purpose of this example is to show that, for such a case, we can often find the global solution and the optimum objective function value of (9)-(10), while the alternating maximum eigenvalues method in [3] usually traps into local critical points. In other words, the objective function value found by our heuristic method is usually larger than that by the method in [3]. We illustrate this by 20 randomly generated examples and plot the objective function values of both methods in Figure 1.

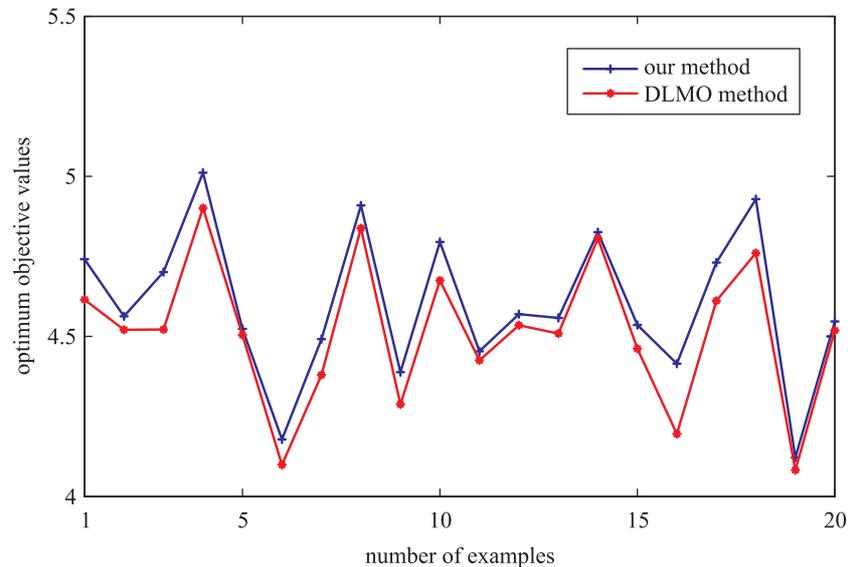


Fig. 1 Comparison of optimum objective function values found by our method and method in [3]

From Fig. 1, we can see that for almost all the randomly generated examples, the optimum objective function values found by our method are larger than those by the method in [3].

### 5.3 Results of feasible descent direction method

In this subsection, we test the performance of the the feasible descent direction method proposed in Section 3, where the subproblems are solved via the direct method for  $p = q = 2$  and the heuristic cross-hill method for general case. For comparison, we also use the Frank-Wolfe method in [3] to solve the same problem, where the subproblem of finding the descent direction was solved by the alternating maximum eigenvalues method in [3]. All codes were written in Matlab and run on a PC with 2.4 GHz CPU and 2.0 G of RAM. We terminate the iteration once the  $\gamma(X_k) < \varepsilon$ , where  $\varepsilon$  is set to be  $10^{-5}$ .

As in [3], we take  $p = q$  and use the maximally entangled matrices in the given dimensions, since the distance to the closest separable matrix in these special cases are known. Here, a maximally entangled matrix means a pure state  $A = uu^\top$ , where

$$u = \frac{1}{\sqrt{p}} \sum_{i=1}^p e_i \otimes f_i,$$

and  $\{e_i\}$  and  $\{f_i\}$  are two arbitrary sets of orthonormal basis vectors in  $\mathbb{R}^p$ . The closest separable state is

$$A' = \lambda A + \frac{1-\lambda}{p^2} I,$$

where  $\lambda = 1/(p+1)$  and the distance between  $A$  and  $A'$  is

$$d = \sqrt{\frac{p-1}{p+1}}.$$

For the method in [3], the maximum number of iteration of solving the subproblem is set to be 30. The stepsize in updating the iteration is calculated via solving an equation in one variable; that is, we use optimal stepsize. In fact, we also use some factors such as 1.2, 1.5, etc. to enlarge the stepsize. However, this can make the algorithm unstable.

The initial point  $X_0$  is generated via randomly generated  $x_0$  and  $y_0$ , where

$$\|x_0\| = \|y_0\| = 1.$$

Table 2 listed the computational results for  $p = q = 2$  to  $p = q = 10$ . In Table

Table 2 Comparison of proposed method and DLMO's method

$p, q$	$n = pq$	Algorithm 3.1		DLMO's method	
		No. of It.	Cpu	No. of It.	Cpu
2	4	2411	182	3209	301
3	9	2083	206	2984	286
4	16	2124	288	3053	353
5	25	2353	612	3255	823
6	36	2427	1923	3535	3233
7	49	2105	3258	2857	4534
8	64	1888	4785	2491	4636
9	81	1980	5529	2698	5207
10	100	2186	7831	2899	6335

2, ‘Cpu’ means the cpu time used in seconds, ‘No. of It.’ means number of iterations.

From Table 2, we can see that the number of iterations of our method is less than that of the method in [3]. The reason is that, as illustrated in the above two subsections, our direct method for the case  $p = q = 2$  and the heuristic cross-hill method for general case, always find ‘better’ solution to subproblem (7)-(8). Thus, at each iteration, the direction used by our method is ‘better’ than that adopted by DLMO’s method. We can also observe that the cpu time used by our method is also less than that by DLMO’s method for the case of small dimensions (from  $p = q = 2$  to  $p = q = 7$ ), and is more than DLMO’s method for the case with relatively high dimensions (from  $p = q = 7$  to  $p = q = 10$ ). The reason is that, on the one hand, to solve the system of equations (17)–(19), we need symbolic computations to find the determinant of the Sylvester matrix, which is time consuming; on the other hand, in our heuristic cross-hill method, we need to solve  $(p - 1)^{p-1}$  subproblems, which increases rapidly as  $p$  increases.

## 6 Final remarks

We proposed a feasible descent direction method for finding the distance between a given state to separable state, as well as the closest separable state to it. The method is globally convergent, provided the subproblem of finding the feasible descent direction is solved, which is, however, a hard problem due to its nonconvexity. We presented a direct method for the case of  $p = q = 2$ , which can solve the subproblem exactly, while for the general case, we proposed a heuristic cross-hill method. The numerical results show the proposed method is efficient.

**Acknowledgements** The first author’s work was supported by the National Natural Science Foundation of China (Grants Nos. 11071122, 11171159) and by the MOE of China (Grant No. 20103207110002). The second author’s work was supported by the Hong Kong Research Grant Council (Grant Nos. PolyU 501909, 502510, 502111, 501212).

## References

1. Bazarra M S, Sherali H D, Shetty C M. *Nonlinear Programming: Theory and Algorithms*. New York: John Wiley and Sons, Inc, 1993
2. Cox D, Little J, O’Shea D. *Using Algebraic Geometry*. New York: Springer-Verlag, 1998
3. Dahl D, Leinass J M, Myrheim J, Ovrum E. A tensor product matrix approximation problem in quantum physics. *Linear Algebra Appl*, 2007, 420: 711–725
4. Doherty A C, Parrilo P A, Spedalieri F M. Distinguishing separable and entangled states. *Phys. Rev. Lett.*, 2002, 88: 187904
5. Einstein A, Podolsky B, Rosen N. Can quantum-mechanical description of physical reality be considered complete? *Phys Rev*, 1935, 47: 777

6. Gurvits L. Classical deterministic complexity of Edmonds' problem and quantum entanglement. IN: Proceedings of the Thirty-Fifth ACM Symposium on Theory of Computing. New York: ACM, 2003, 10–19
7. Han D, Dai H, Qi L. Conditions for strong ellipticity of anisotropic elastic materials. *J Elasticity*, 2009, 97: 1–13
8. Han D, Qi L, Wu Ed. Extreme diffusion values for non-Gaussian diffusions. *Optim Methods Softw*, 2008, 23: 703–716
9. Horodecki M, Horodecki P, Horodecki R. Separability of mixed states: Necessary and sufficient conditions. *Phys Lett A*, 1996, 223: 1–8
10. Ioannou L M, Travaglione B C, Cheung D, Ekert K. Improved algorithm for quantum separability and entanglement detection. *Phys Rev A*, 2004, 70: 060303
11. Ling C, Nie J, Qi L, Ye Y. Bi-quadratic optimization over unit spheres and semidefinite programming relaxations. *SIAM J Optim*, 2009, 20: 1286–1310
12. Ng M, Qi L, Zhou G. Finding the largest eigenvalue of a non-negative tensor. *SIAM J Matrix Anal Appl*, 2009, 31: 1090–1099
13. Nielsen M N, Chuang I L. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press, 2000
14. Peres A. Separability criterion for density matrices. *Phys Rev Lett*, 1996, 77: 1413–1415
15. Pittenger A O, Rubin M H. Geometry of entanglement witness and local detection of entanglement. *Phys Rev A*, 2003, 67: 012327
16. Qi L. Rank and eigenvalues of a supersymmetric tensor, a multivariate homogeneous polynomial and an algebraic surface defined by them. *J Symbolic Comput*, 2006, 41: 1309–1327
17. Qi L. Eigenvalues and invariants of tensors. *J Math Anal Appl*, 2007, 325: 1363–1377
18. Qi L, Dai H, Han D. Conditions for strong ellipticity. *Front Math China*, 2009, 4: 349–364
19. Qi L, Wang F, Wang Y. A global homogenous polynomial optimization problem over the unit sphere. Working Paper, Department of Applied Mathematics, The Hong Kong Polytechnic University, 2007
20. Qi L, Wang F, Wang Y. Z-eigenvalue methods for a global polynomial optimization problem. *Math Program*, 2009, 118: 301–316
21. Qi L, Wang Y, Wu Ed. D-eigenvalues of diffusion kurtosis tensors. *J Comp Appl Math*, 2008, 221: 150–157
22. Schrödinger E. Die gegenwärtige situation in der quantenmechanik. *Naturwissenschaften*, 1935, 23: 807–812, 823–828, 844–849
23. Verstraete F, Dehaene J, De Moor B. On the geometry of entangled states. *J Mod Opt*, 2002, 49: 1277–1287