

RESEARCH

Open Access



A globally convergent QP-free algorithm for nonlinear semidefinite programming

Jian-Ling Li¹, Zhen-Ping Yang¹ and Jin-Bao Jian^{2*}

*Correspondence:
jianjb@gxu.edu.cn

²School of Mathematics and Statistics, Guangxi Colleges and Universities Key Laboratory of Complex System Optimization and Big Data Processing, Yulin Normal University, Jiaoyu Road, Yulin, Guangxi 537000, China
Full list of author information is available at the end of the article

Abstract

In this paper, we present a QP-free algorithm for nonlinear semidefinite programming. At each iteration, the search direction is yielded by solving two systems of linear equations with the same coefficient matrix; l_1 penalty function is used as merit function for line search, the step size is determined by Armijo type inexact line search. The global convergence of the proposed algorithm is shown under suitable conditions. Preliminary numerical results are reported.

MSC: Primary 90C30; secondary 90C22

Keywords: nonlinear semidefinite programming; KKT conditions; QP-free algorithm; global convergence

1 Introduction

Consider the following nonlinear semidefinite programming (NLSDP for short):

$$\begin{aligned} & \min f(x) \\ & \text{s.t. } \mathcal{A}(x) \preceq 0; \\ & \quad h_j(x) = 0, \quad j \in \mathcal{E} = \{1, 2, \dots, l\}, \end{aligned} \tag{1.1}$$

where $f : R^n \rightarrow R$, $h_j (j \in \mathcal{E}) : R^n \rightarrow R^l$ and $\mathcal{A} : R^n \rightarrow \mathcal{S}^m$ are continuously differentiable functions, not necessarily convex. \mathcal{S}^m is a space whose elements are real symmetric matrices of size $m \times m$. \preceq denotes the negative semidefinite order, that is, $A \preceq B$ if and only if $A - B$ is a negative semidefinite matrix.

NLSDP (1.1) has a broad range of applications such as eigenvalue problems, control problems, optimal structural design, truss design problems (see [1–3]). So it is desired to develop numerical methods for solving NLSDP (1.1).

In recent years, NLSDPs have been attracting a great deal of research attention [1, 3–25]. As is well known, NLSDP (1.1) is an extension of nonlinear programming, some efficient numerical methods for the latter are generalized to solve NLSDP. For example, Correa and Ramirez [26] proposed an algorithm which used the sequential linear SDP method. Fares *et al.* [27] applied the sequential linear SDP method to robust control problems. Freund *et al.* [4] also studied a sequential SDP method. Kanzow *et al.* [9] presented a successive linearization method with a trust region-type globalization strategy.

In addition, Kovara and Stingl [10] developed a computer code PENNON for solving NLSDP (1.1), where the augmented Lagrangian function method was used. Sun *et al.* [20] and Luo *et al.* [11, 22] proposed an augmented Lagrangian method for NLSDP (1.1), respectively. Sun *et al.* [19] analyzed the rate of local convergence of the augmented Lagrangian method for NLSDPs. Yamashita *et al.* recently proposed a primal-dual interior point method for NLSDP (1.1) (see [23]). The algorithm is globally convergent and locally superlinearly convergent under suitable conditions. Very recently Aroztegui [28] proposed a feasible direction interior point algorithm for NLSDP (1.1) with only semidefinite matrix constraint.

As we know, QP-free (also called SSLE) method is a kind of efficient methods for standard nonlinear programs (see [16]-[13]). In this paper, motivated from QP-free method for standard nonlinear programs, based on techniques of perturbation and penalty function, we propose a globally convergent QP-free algorithm for NLSDP (1.1). The construction of systems of linear equations (SLE for short) is a key point. Based on KKT conditions of NLSDP (1.1) and techniques of perturbation, we construct two SLEs skillfully. At each iteration, the search direction is yielded by solving two SLEs with the same coefficient matrix; An exact penalty function is used as the merit function for line search and the step size is determined by suitable inexact line search. The global convergence of the proposed algorithm is shown under some mild conditions.

The paper is organized as follows. In Section 2 we restate some definitions and results on NLSDP and matrix analysis. In Section 3 the algorithm is presented and its feasibility is discussed. The global convergence is analyzed in Section 4. Some preliminary numerical results are reported in Section 5 and some concluding remarks are given in the final section.

2 Preliminaries

For the sake of convenience, some results on matrix analysis and NLSDP are restated in this section, which will be employed in the following analysis of the proposed algorithm. More introduction for theory of matrices should be seen in [21] and [6]. Denote by $R^{m \times n}$ the space of $m \times n$ real matrices, denote by S_+^m and S_{++}^m the sets of m -order symmetric positive semidefinite and positive definite matrices, respectively. The sets S_-^m and S_{--}^m are defined similarly.

Definition 2.1 For any $A = (a_{ij}), B = (b_{ij}) \in R^{m \times n}$, the inner product of A and B is defined by

$$\langle A, B \rangle = \text{Tr}(B^T A) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_{ij}, \tag{2.1}$$

where $\text{Tr}(P)$ means the trace of the matrix P .

Definition 2.2 ([6]) For any $M \in R^{m \times m}$, let

$$\text{sym}(M) = \frac{1}{2}(M + M^T), \quad \text{skw}(M) = \frac{1}{2}(M - M^T), \tag{2.2}$$

$\text{sym}(M)$ and $\text{skw}(M)$ are called the symmetric part and the skew part of M , respectively.

Given a matrix $A \in \mathcal{S}^m$, let $\bar{m} = \frac{1}{2}m(m + 1)$, define a map $\text{svec}: \mathcal{S}^m \rightarrow \mathbb{R}^{\bar{m}}$:

$$\text{svec}(A) = (a_{11}, \sqrt{2}a_{21}, \dots, \sqrt{2}a_{m1}, a_{22}, \sqrt{2}a_{32}, \dots, \sqrt{2}a_{m2}, \dots, a_{mm})^T,$$

and the map $\text{smat}: \mathbb{R}^{\bar{m}} \rightarrow \mathcal{S}^m$ is defined to be the inverse of svec . Then the inner product of matrices is indicated by

$$\langle A, B \rangle = \text{svec}(A)^T \text{svec}(B), \quad \text{for } A, B \in \mathcal{S}^m. \tag{2.3}$$

Definition 2.3 ([21]) For any $A, B \in \mathbb{R}^{m \times m}$, the symmetric Kronecker product, denoted by $A \otimes_s B$, is a mapping on a vector $u = \text{svec}(U)$ where U is an $m \times m$ symmetric matrix and is defined as

$$(A \otimes_s B)u = \frac{1}{2} \text{svec}(BUA^T + AUB^T). \tag{2.4}$$

For any matrix $U \in \mathcal{S}^m$, it is verified that the following equality is true:

$$(A \otimes_s B) \text{svec}(U) = \text{svec}(\text{sym}(BUA)). \tag{2.5}$$

Note that the linear operator $A \otimes_s B$ is defined implicitly in (2.4). In Appendix of [21] a matrix representation of $A \otimes_s B$ is given as follows:

$$A \otimes_s B = \frac{1}{2}Q(A \otimes B + B \otimes A)Q^T, \tag{2.6}$$

where $A \otimes B = [a_{ij}B]$ ($i, j = 1, 2, \dots, m$) is the Kronecker product of A and B , Q is an orthogonal $\bar{m} \times m^2$ matrix (i.e. $QQ^T = I_{\bar{m}}$), with the following property:

$$Q \text{vec}(U) = \text{svec}(U), \quad Q^T \text{svec}(U) = \text{vec}(U), \quad \forall U \in \mathcal{S}^m, \tag{2.7}$$

where $\text{vec}(U) = (u_{11}, u_{21}, \dots, u_{m1}, u_{12}, u_{22}, \dots, u_{m2}, \dots, u_{mm})^T$.

Remark 2.1 One choice for the matrix Q is given in the appendix of [21].

Lemma 2.1 ([21]) For any $A, B \in \mathcal{S}^m$, the following results are true:

- (1) $A \otimes_s B = B \otimes_s A$;
- (2) $(A \otimes_s B)^T = A^T \otimes_s B^T$;
- (3) $(A \otimes_s B)(C \otimes_s D) = \frac{1}{2}(AC \otimes_s BD + AD \otimes_s BC)$;
- (4) If A and B are symmetric positive definite, then $A \otimes_s B$ is positive definite.

Lemma 2.2 ([28]) If $A, B \in \mathcal{S}^m$, $A \succ 0$ and $AB + BA \prec 0$, then $B \prec 0$.

Lemma 2.3 If $A \in \mathcal{S}_{++}^m$, $B \in \mathcal{S}_{--}^m$, then all eigenvalues of AB are less than zero.

The proof is elementary and omitted here.

Lemma 2.4 ([28]) If $A \in \mathcal{S}_{++}^m$, $B \in \mathcal{S}_{--}^m$, and they commute, then $(A \otimes_s I_m)^{-1}(B \otimes_s I_m) \in \mathcal{S}_{--}^{\bar{m}}$.

Lemma 2.5 *Suppose $A \in S_{++}^m, B \in S_{--}^m$, and they commute, then $(A \otimes_s I_m)^{-1}(B \otimes_s I_m) \in S_{--}^m$.*

Proof Since $A \in S_{++}^m, B \in S_{--}^m$, and they commute, there exists an orthogonal matrix $P \in R^{m \times m}$ such that

$$A = PD_A P^{-1}, \quad B = PD_B P^{-1},$$

where D_A is a diagonal and positive definite matrix, and D_B is a diagonal and negative definite matrix. It follows from Lemma 2.1(3) that

$$A \otimes_s I_m = T D_A T^{-1}, \quad B \otimes_s I_m = T D_B T^{-1},$$

where $T = P \otimes_s P, D_A = D_A \otimes_s I_m$ and $D_B = D_B \otimes_s I_m$. We know from Lemma 2.1(2), (3) that T is orthogonal, from Lemma 2.1(4) that D_A is a diagonal and positive definite matrix, and D_B is a diagonal and negative definite matrix. Hence,

$$(A \otimes_s I_m)^{-1}(B \otimes_s I_m) = T D_A D_B T^{-1} \in S_{--}^m. \quad \square$$

In the rest of this section we state the first order optimality conditions for NLSDP (1.1). For the sake of convenience, we first introduce some notations. Given a matrix valued function $\mathcal{A}(\cdot)$, we use the notation

$$D\mathcal{A}(x) = \left(\frac{\partial \mathcal{A}(x)}{\partial x_1}, \dots, \frac{\partial \mathcal{A}(x)}{\partial x_n} \right)^T$$

for its differential operator evaluated at x , where $\frac{\partial \mathcal{A}(x)}{\partial x_i}$ denotes the partial derivative of $\mathcal{A}(x)$ with respect to x_i with components $\frac{\partial a_{pq}(x)}{\partial x_i}$ ($p, q = 1, \dots, m$). Then the derivative of $\mathcal{A}(\cdot)$ in the direction $d = (d_1, \dots, d_n)^T \in R^n$ at x denoted by $D\mathcal{A}(x)d$ is defined by

$$D\mathcal{A}(x)d = \sum_{i=1}^n d_i \frac{\partial \mathcal{A}(x)}{\partial x_i}. \tag{2.8}$$

If we denote

$$\nabla \mathcal{A}(x) := \left(\text{svec} \left(\frac{\partial \mathcal{A}(x)}{\partial x_1} \right), \dots, \text{svec} \left(\frac{\partial \mathcal{A}(x)}{\partial x_n} \right) \right)_{\overline{m} \times n}, \tag{2.9}$$

then by (2.8), the following equality is true:

$$\text{svec}(D\mathcal{A}(x)d) = \nabla \mathcal{A}(x)d. \tag{2.10}$$

The Lagrangian function of NLSDP (1.1) $L : R^n \times S^m \times R^l \rightarrow R$ is defined by

$$L(x, \Lambda, \mu) = f(x) + \langle \mathcal{A}(x), \Lambda \rangle + h(x)^T \mu, \tag{2.11}$$

where $h(x) = (h_1(x), h_2(x), \dots, h_l(x))^T$. In view of (2.3), the above equality can be rewritten as follows:

$$L(x, \lambda, \mu) = f(x) + \text{svec}(\mathcal{A}(x))^T \lambda + h(x)^T \mu,$$

where $\lambda := \text{svec}(\Lambda)$. The gradient of $L(x, \lambda, \mu)$ with respect to x is given as follows:

$$\nabla_x L(x, \lambda, \mu) = \nabla f(x) + \nabla \mathcal{A}(x)^T \lambda + \nabla h(x) \mu, \tag{2.12}$$

where $\nabla h(x) = (\nabla h_1(x), \nabla h_2(x), \dots, \nabla h_l(x))$.

We are now in a position to restate the definition of the first order optimality conditions for NLSDP (1.1).

Definition 2.4 ([18]) For $x \in R^n$, if there exist a matrix $\Lambda \in S^m$ and a vector $\mu (\in R^l)$ such that

$$\nabla_x L(x, \Lambda, \mu) = 0, \tag{2.13a}$$

$$\Lambda \mathcal{A}(x) = 0, \quad \Lambda \succeq 0, \tag{2.13b}$$

$$h(x) = 0, \quad \mathcal{A}(x) \preceq 0, \tag{2.13c}$$

then x is called a KKT point of NLSDP (1.1).

Remark 2.2 According to the Von Neumann-Theobald inequality, the complementarity condition $\Lambda \mathcal{A}(x) = 0$ has the following two useful equivalent forms:

$$\text{Tr}(\Lambda \mathcal{A}(x)) = 0,$$

$$\lambda_j(\Lambda) \lambda_j(\mathcal{A}(x)) = 0, \quad \forall j \in \{1, 2, \dots, m\}.$$

3 The algorithm

In this section, we present our algorithm and show it is well defined. For the sake of simplicity, we introduce some notations:

$$\Omega = \{x \in R^n : \mathcal{A}(x) \preceq 0, h(x) = 0\},$$

$$\mathcal{F} = \{x \in R^n : \mathcal{A}(x) \preceq 0\}, \quad \mathcal{F}_0 = \{x \in R^n : \mathcal{A}(x) \prec 0\},$$

that is, Ω is the feasible set of NLSDP (1.1).

In general, $\Lambda \mathcal{A}(x)$ is not guaranteed to be symmetric, so we consider $\text{sym}(\Lambda \mathcal{A}(x)) = 0$ instead of $\Lambda \mathcal{A}(x) = 0$. Then the three equalities of KKT condition (2.13a)-(2.13c) can be rewritten in the following form:

$$\nabla f(x) + \nabla \mathcal{A}(x)^T \lambda + \nabla h(x) \mu = 0,$$

$$\text{svec}(\text{sym}(\Lambda \mathcal{A}(x))) = 0, \tag{3.1}$$

$$h(x) = 0.$$

In order to solve (3.1) at each Newton iteration, we define a vector-value function $\varphi : R^{n+m+l} \rightarrow R^{n+m+l}$ as follows:

$$\varphi(x, \lambda, \mu) = \begin{pmatrix} \varphi_{Lg}(x, \lambda, \mu) \\ \varphi_C(x, \lambda, \mu) \\ \varphi_h(x, \lambda, \mu) \end{pmatrix} = \begin{pmatrix} \nabla f(x) + \nabla \mathcal{A}(x)^T \lambda + \nabla h(x) \mu \\ \text{svec}(\text{sym}(\Lambda \mathcal{A}(x))) \\ h(x) \end{pmatrix}.$$

It follows from (2.5) and Lemma 2.1 that

$$\varphi_C(x, \lambda, \mu) = \text{svec}(\text{sym}(I\Lambda\mathcal{A}(x))) = (I \otimes_s \mathcal{A}(x)) \text{svec}(\Lambda) = (\Lambda \otimes_s I) \text{svec}(\mathcal{A}(x)),$$

thus, the Jacobian of φ is

$$\nabla\varphi(x, \lambda, \mu) = \begin{pmatrix} \nabla_{xx}^2 L(x, \lambda, \mu) & \nabla\mathcal{A}(x)^T & \nabla h(x) \\ (\Lambda \otimes_s I)\nabla\mathcal{A}(x) & I \otimes_s \mathcal{A}(x) & 0 \\ \nabla h(x)^T & 0 & 0 \end{pmatrix}.$$

Instead of the Hessian $\nabla_{xx}^2 L(x, \lambda, \mu)$, we employ a positive definite matrix denoted by H which can be a quasi-Newton approximation or the identity matrix. A Newton-like iteration to solve (3.1) is given by the linear systems as follows:

$$\begin{pmatrix} H & \nabla\mathcal{A}(x)^T & \nabla h(x) \\ (\bar{\Lambda} \otimes_s I)\nabla\mathcal{A}(x) & I \otimes_s \mathcal{A}(x) & 0 \\ \nabla h(x)^T & 0 & 0 \end{pmatrix} \begin{pmatrix} x^0 - x \\ \lambda^0 - \bar{\lambda} \\ \mu^0 - \mu \end{pmatrix} = - \begin{pmatrix} \nabla f(x) + \nabla\mathcal{A}(x)^T \bar{\lambda} + \nabla h(x)\mu \\ \text{svec}(\text{sym}(\bar{\Lambda}\mathcal{A}(x))) \\ h(x) \end{pmatrix}, \tag{3.2}$$

where $(x, \bar{\Lambda}, \mu) \in \mathcal{F}_0 \times \mathcal{S}_{++}^m \times R^l$ is the current point, $(x^0, \Lambda^0, \mu^0) \in \mathcal{F} \times \mathcal{S}_{++}^m \times R^l$ is the new estimates given by the Newton-like iteration, $\bar{\lambda} := \text{svec}(\bar{\Lambda})$ and $\lambda^0 := \text{svec}(\Lambda^0)$. Let $d^0 = x - x^0$, we obtain from (3.2)

$$Hd^0 + \nabla\mathcal{A}(x)^T \lambda^0 + \nabla h(x)\mu^0 = -\nabla f(x), \tag{3.3a}$$

$$(\bar{\Lambda} \otimes_s I)\nabla\mathcal{A}(x)d^0 + (I \otimes_s \mathcal{A}(x))\lambda^0 = 0, \tag{3.3b}$$

$$\nabla h(x)^T d^0 = -h(x). \tag{3.3c}$$

If $d^0 = 0$, then we have

$$\nabla f(x) + \nabla\mathcal{A}(x)^T \lambda^0 + \nabla h(x)\mu^0 = 0,$$

$$(I \otimes_s \mathcal{A}(x))\lambda^0 = 0, \quad h(x) = 0.$$

Since $\mathcal{A}(x) \prec 0$, $I \otimes_s \mathcal{A}(x)$ is nonsingular and we have $\Lambda^0 := \text{smat}(\lambda^0) = 0$, which implies that $\Lambda^0 \mathcal{A}(x) = 0$. Therefore, x is a KKT point. If $d^0 \neq 0$, then d^0 is not guaranteed to be a feasible direction. To obtain a better search direction, we modify (3.3b) by introducing an appropriate right hand side, so we obtain another linear equations as follows:

$$Hd^1 + \nabla\mathcal{A}(x)^T \lambda^1 + \nabla h(x)\mu^1 = -\nabla f(x),$$

$$(\bar{\Lambda} \otimes_s I)\nabla\mathcal{A}(x)d^1 + (I \otimes_s \mathcal{A}(x))\lambda^1 = -\bar{\lambda} \|d^0\|, \tag{3.4}$$

$$\nabla h(x)^T d^1 = -h(x).$$

In order to ensure that SLEs (3.3a)-(3.3c) and (3.4) have a unique solution, respectively, the following assumption is required.

A1 For any $x \in \mathcal{F}$, the matrix

$$B(x) = \begin{pmatrix} \nabla \mathcal{A}(x)^T & \nabla h(x) \\ \mathcal{A}(x) \otimes_s I_m & 0 \end{pmatrix}$$

is full of column rank.

The following lemma gives a sufficient condition of the assumption A1.

Lemma 3.1 For any $x \in \mathcal{F}$, if $\mathcal{A}(x) < 0$ and $\{\nabla h_1(x), \dots, \nabla h_l(x)\}$ is linearly independent, then $B(x)$ is full of column rank, i.e., the assumption A1 holds.

Lemma 3.2 Let H be a positive definite matrix. If the assumption A1 holds, then the coefficient matrix of the SLEs (3.3a)-(3.3c) and (3.4)

$$W(x, H, \bar{\Lambda}) \stackrel{\text{def}}{=} \begin{pmatrix} H & \nabla \mathcal{A}(x)^T & \nabla h(x) \\ (\bar{\Lambda} \otimes_s I_m) \nabla \mathcal{A}(x) & \mathcal{A}(x) \otimes_s I & 0 \\ \nabla h(x)^T & 0 & 0 \end{pmatrix} \tag{3.5}$$

is nonsingular, hence, SLEs (3.3a)-(3.3c) and (3.4) have a unique solution, respectively.

The proof is elementary and it is omitted here.

In our algorithm the following exact penalty function is used as a merit function for line search:

$$P(x; \sigma) = f(x) + \sigma \sum_{j \in \mathcal{E}} |h_j(x)|, \tag{3.6}$$

where $\sigma > 0$ is a penalty parameter. Further, we define a function $\bar{P}(\bullet; d; \sigma) : R^n \times R^n \times [0, +\infty) \rightarrow R$ associated with $P(x; \sigma)$ by

$$\bar{P}(x; d; \sigma) = f(x) + \nabla f(x)^T d + \sigma \sum_{j \in \mathcal{E}} |h_j(x) + \nabla h_j(x)^T d|. \tag{3.7}$$

Now the algorithm is described in detail.

Algorithm A

Parameters. $\alpha \in (0, \frac{1}{2})$, $\beta, \xi \in (0, 1)$, $\lambda^l > 0$, $\sigma_{-1} > 0$, $\rho_1, \rho_2 > 0$.

Initialization. Select an initial iteration point $x^0 \in \mathcal{F}_0$, $H_0 \in S_{++}^n$, $\bar{\Lambda}_0 \in (S_{++}^m)$ satisfying $\lambda^l I_m \leq \bar{\Lambda}_0$ such that $\bar{\Lambda}_0$ and $\mathcal{A}(x^0)$ commute. Let $\bar{\lambda}_0 = \text{svec}(\bar{\Lambda}_0)$, $k := 0$.

Step 1. Let $(d^{k0}, \lambda^{k0}, \mu^{k0})$ be the solution of the SLE (3.3a)-(3.3c) in (d, λ, μ) , i.e.,

$$\begin{cases} H_k d + \nabla \mathcal{A}(x^k)^T \lambda + \sum_{j \in \mathcal{E}} \mu_j \nabla h_j(x^k) = -\nabla f(x^k), \\ (\bar{\Lambda}_k \otimes_s I_m) \nabla \mathcal{A}(x^k) d + (\mathcal{A}(x^k) \otimes_s I_m) \lambda = 0, \\ \nabla h_j(x^k)^T d = -h_j(x^k), \quad j \in \mathcal{E}. \end{cases} \tag{3.8}$$

If $d^{k0} = 0$, then stop, x^k is a KKT point of NLSDP (1.1); else, go to Step 2.

Step 2. Let $(d^{k1}, \lambda^{k1}, \mu^{k1})$ be the solution of the SLE (3.4) in (d, λ, μ) , i.e.,

$$\begin{cases} H_k d + \nabla \mathcal{A}(x^k)^T \lambda + \sum_{j \in \mathcal{E}} \mu_j \nabla h_j(x^k) = -\nabla f(x^k), \\ (\bar{\Lambda}_k \otimes_s I_m) \nabla \mathcal{A}(x^k) d + (\mathcal{A}(x^k) \otimes_s I_m) \lambda = -\bar{\lambda}_k \|d^{k0}\|, \\ \nabla h_j(x^k)^T d = -h_j(x^k), \quad j \in \mathcal{E}. \end{cases} \tag{3.9}$$

Step 3. Compute the search direction d^k and the approximate multiplier vector (λ^k, μ^k) :

$$d^k = (1 - \delta_k) d^{k0} + \delta_k d^{k1}, \tag{3.10}$$

$$\lambda^k = (1 - \delta_k) \lambda^{k0} + \delta_k \lambda^{k1}, \tag{3.11}$$

$$\mu^k = (1 - \delta_k) \mu^{k0} + \delta_k \mu^{k1}, \tag{3.12}$$

where

$$\delta_k = \begin{cases} 1 - \xi, & \text{if } \nabla f(x^k)^T d^{k1} \leq 0; \\ 1, & \text{if } \nabla f(x^k)^T d^{k1} > 0 \text{ and } \nabla f(x^k)^T d^{k1} \leq \nabla f(x^k)^T d^{k0}; \\ \min\{\xi, |(1 - \xi) \frac{\nabla f(x^k)^T d^{k0} + (\mu^{k0})^T h(x^k)}{\nabla f(x^k)^T (d^{k0} - d^{k1})}|\}, & \text{otherwise.} \end{cases} \tag{3.13}$$

Step 4. (Update the penalty parameter) Set $\bar{\sigma}_k = (3 - \xi) \max_{j \in \mathcal{E}} |\mu_j^{k0}| + \rho_1$. The updating rule of σ_k is as follows:

$$\sigma_k = \begin{cases} \max\{\bar{\sigma}_k, \sigma_{k-1} + \rho_2\}, & \text{if } \bar{\sigma}_k > \sigma_{k-1}, \\ \sigma_{k-1}, & \text{otherwise.} \end{cases} \tag{3.14}$$

Step 5. (Line search) Set the step size t_k to be the first number of the sequence $\{1, \beta, \beta^2, \dots\}$ satisfying the following two inequalities:

$$P(x^k + t d^k; \sigma_k) \leq P(x^k; \sigma_k) + t \alpha (\bar{P}(x^k; d^k; \sigma_k) - \bar{P}(x^k; 0; \sigma_k)), \tag{3.15}$$

$$\mathcal{A}(x^k + t d^k) < 0. \tag{3.16}$$

Step 6. Set $x^{k+1} = x^k + t_k d^k$. Using the following methods to generate $\bar{\Lambda}_{k+1}$ commuting with $\mathcal{A}(x^{k+1})$:

Step 6.1. If the search direction d^k does not descend or is not feasible, set $\bar{\Lambda}_{k+1} = I_m$ and go to Step 7.

Step 6.2. Compute the least eigenvalue $\lambda_{\min}(\bar{\Lambda}_k)$ of the matrix $\bar{\Lambda}_k$. If $\lambda_{\min}(\bar{\Lambda}_k) \geq \lambda^l$, then let $\bar{\Lambda}_{k+1} = \bar{\Lambda}_k$; otherwise, let $\bar{\Lambda}_{k+1} = \bar{\Lambda}_k + (\lambda^l - \lambda_{\min}(\bar{\Lambda}_k)) I_m$.

Step 7. Set $\bar{\lambda}_{k+1} = \text{svec}(\bar{\Lambda}_{k+1})$, and update H_k by some method to H_{k+1} such that H_{k+1} is symmetric positive definite. Let $k := k + 1$, return to Step 1.

By (3.8), the following lemma is obvious.

Lemma 3.3 *Suppose that the assumption A1 holds. If $d^{k_0} = 0$, then x^k is a KKT point of NLSDP (1.1).*

Lemma 3.4 *Suppose that the assumption A1 holds. Then the search direction d^k of Algorithm A satisfies the following inequality:*

$$\nabla f(x^k)^T d^k \leq -\xi (d^{k_0})^T H_k d^{k_0} + (3 - \xi) \sum_{j \in \mathcal{E}} |\mu_j^{k_0} h_j(x^k)|. \tag{3.17}$$

Proof First we show that the inequality

$$\nabla f(x^k)^T d^{k_0} \leq -(d^{k_0})^T H_k d^{k_0} + \sum_{j \in \mathcal{E}} |\mu_j^{k_0} h_j(x^k)| \tag{3.18}$$

holds. Premultiplying the first equation of (3.8) by $(d^{k_0})^T$, we obtain

$$(d^{k_0})^T H_k d^{k_0} + \sum_{j \in \mathcal{E}} \mu_j^{k_0} (d^{k_0})^T \nabla h_j(x^k) + (d^{k_0})^T \nabla \mathcal{A}(x^k)^T \lambda^{k_0} = -(d^{k_0})^T \nabla f(x^k). \tag{3.19}$$

According to the second equation of (3.8), we get

$$(d^{k_0})^T \nabla \mathcal{A}(x^k)^T \lambda^{k_0} = -(\lambda^{k_0})^T ((\bar{\Lambda}_k \otimes_s I_m)^{-1} (\mathcal{A}(x^k) \otimes_s I_m))^T \lambda^{k_0}.$$

Substituting the above equality and the third equality of (3.8) into (3.19), we have

$$\begin{aligned} & (d^{k_0})^T \nabla f(x^k) \\ &= -(d^{k_0})^T H_k d^{k_0} + (\lambda^{k_0})^T ((\bar{\Lambda}_k \otimes_s I_m)^{-1} (\mathcal{A}(x^k) \otimes_s I_m))^T \lambda^{k_0} + \sum_{j \in \mathcal{E}} \mu_j^{k_0} h_j(x^k). \end{aligned}$$

In view of Lemma 2.4, the matrix $(\bar{\Lambda}_k \otimes_s I_m)^{-1} (\mathcal{A}(x^k) \otimes_s I_m)$ is negative semidefinite, so it follows from the above equality that

$$(d^{k_0})^T \nabla f(x^k) \leq -(d^{k_0})^T H_k d^{k_0} + \sum_{j \in \mathcal{E}} |\mu_j^{k_0} h_j(x^k)|,$$

i.e., the inequality (3.18) holds.

Next, we will prove the inequality (3.17) is true. The rest of the proof is divided into three cases.

Case A. $\nabla f(x^k)^T d^{k_1} \leq 0$. From (3.13) we have $\delta_k = 1 - \xi$. It follows from (3.10), (3.13), (3.18) and $\xi \in (0, 1)$ that

$$\begin{aligned} \nabla f(x^k)^T d^k &\leq -\xi (d^{k_0})^T H_k d^{k_0} + \xi \sum_{j \in \mathcal{E}} |\mu_j^{k_0} h_j(x^k)| \\ &\leq -\xi (d^{k_0})^T H_k d^{k_0} + (3 - \xi) \sum_{j \in \mathcal{E}} |\mu_j^{k_0} h_j(x^k)|, \end{aligned} \tag{3.20}$$

that is, (3.17) holds.

Case B. $\nabla f(x^k)^T d^{k1} > 0$ and $\nabla f(x^k)^T d^{k1} \leq \nabla f(x^k)^T d^{k0}$. From (3.13), one has $\delta_k = 1$. It follows from (3.10), (3.19) and $\xi \in (0, 1)$ that

$$\begin{aligned} \nabla f(x^k)^T d^k &= \nabla f(x^k)^T d^{k1} \leq \nabla f(x^k)^T d^{k0} \\ &\leq -(d^{k0})^T H_k d^{k0} + \sum_{j \in \mathcal{E}} |\mu_j^{k0} h_j(x^k)|, \end{aligned}$$

which implies (3.17) holds.

Case C. $\nabla f(x^k)^T d^{k1} > 0$ and $\nabla f(x^k)^T d^{k1} > \nabla f(x^k)^T d^{k0}$. It follows from (3.13) and $\xi \in (0, 1)$ that

$$\begin{aligned} \delta_k &= \left| (1 - \xi) \frac{\nabla f(x^k)^T d^{k0} + (\mu^{k0})^T h(x^k)}{\nabla f(x^k)^T (d^{k1} - d^{k0})} \right| \\ &\leq \frac{|(\xi - 1) \nabla f(x^k)^T d^{k0}| + |(\mu^{k0})^T h(x^k)|}{\nabla f(x^k)^T (d^{k1} - d^{k0})}. \end{aligned} \tag{3.21}$$

If $\nabla f(x^k)^T d^{k0} \leq 0$, then we obtain from the above inequality

$$(1 - \delta_k) \nabla f(x^k)^T d^{k0} + \delta_k \nabla f(x^k)^T d^{k1} \leq \xi \nabla f(x^k)^T d^{k0} + |(\mu^{k0})^T h(x^k)|,$$

which together with (3.10) and (3.18) gives

$$\begin{aligned} \nabla f(x^k)^T d^k &\leq -\xi (d^{k0})^T H_k d^{k0} + (1 + \xi) \sum_{j \in \mathcal{E}} |\mu_j^{k0} h_j(x^k)| \\ &\leq -\xi (d^{k0})^T H_k d^{k0} + (3 - \xi) \sum_{j \in \mathcal{E}} |\mu_j^{k0} h_j(x^k)|. \end{aligned} \tag{3.22}$$

If $\nabla f(x^k)^T d^{k0} > 0$, then the inequality (3.21) gives rise to

$$\delta_k \nabla f(x^k)^T d^{k1} - \delta_k \nabla f(x^k)^T d^{k0} \leq (1 - \xi) \nabla f(x^k)^T d^{k0} + |(\mu^{k0})^T h(x^k)|,$$

which together with (3.10) and (3.18) shows

$$\begin{aligned} \nabla f(x^k)^T d^k &\leq -(2 - \xi) (d^{k0})^T H_k d^{k0} + (3 - \xi) \sum_{j \in \mathcal{E}} |\mu_j^{k0} h_j(x^k)| \\ &\leq -\xi (d^{k0})^T H_k d^{k0} + (3 - \xi) \sum_{j \in \mathcal{E}} |\mu_j^{k0} h_j(x^k)|. \end{aligned} \tag{3.23}$$

The inequalities (3.22) and (3.23) indicate that the inequality (3.17) is true. □

Lemma 3.5 *Suppose that the assumption A1 holds. If $x^k (\in \mathcal{F})$ is not a KKT point of NLSDP (1.1), then*

$$\bar{P}(x^k; d^k; \sigma_k) - \bar{P}(x^k; 0; \sigma_k) < 0. \tag{3.24}$$

Proof From (3.8) and (3.9) we know that (d^k, λ^k, μ^k) is the solution of the following SLE:

$$H_k d + \nabla \mathcal{A}(x^k)^T \lambda + \sum_{j \in \mathcal{E}} \mu_j \nabla h_j(x^k) = -\nabla f(x^k), \tag{3.25a}$$

$$(\bar{\Lambda}_k \otimes_s I_m) \nabla \mathcal{A}(x^k) d + (\mathcal{A}(x^k) \otimes_s I_m) \lambda = -\delta_k \bar{\lambda}_k \|d^{k0}\|, \tag{3.25b}$$

$$\nabla h_j(x^k)^\top d = -h_j(x^k), \quad j \in \mathcal{E}. \tag{3.25c}$$

From the definition (3.6) of the function $\bar{P}(x^k; d^k; \sigma_k)$ and (3.25c), we have

$$\begin{aligned} & \bar{P}(x^k; d^k; \sigma_k) - \bar{P}(x^k; 0; \sigma_k) \\ &= \nabla f(x^k)^\top d^k - \sigma_k \sum_{j \in \mathcal{E}} |h_j(x^k)| \\ &\leq -\xi (d^{k0})^\top H_k d^{k0} + (3 - \xi) \sum_{j \in \mathcal{E}} |\mu_j^{k0} h_j(x^k)| - \sigma_k \sum_{j \in \mathcal{E}} |h_j(x^k)| \\ &\leq -\xi (d^{k0})^\top H_k d^{k0} + \left((3 - \xi) \max_{j \in \mathcal{E}} |\mu_j^{k0}| - \sigma_k \right) \sum_{j \in \mathcal{E}} |h_j(x^k)|, \end{aligned} \tag{3.26}$$

the first inequality above is due to (3.17).

Since x^k is not a KKT point of NLSDP (1.1), it implies from Step 1 of Algorithm A that $d^{k0} \neq 0$, so $(d^{k0})^\top H_k d^{k0} > 0$. On the other hand, it follows from the updating rule of σ_k that $\sigma_k > (3 - \xi) \max_{j \in \mathcal{E}} |\mu_j^{k0}|$, therefore, (3.26) gives rise to

$$\bar{P}(x^k; d^k; \sigma_k) - \bar{P}(x^k; 0; \sigma_k) < 0,$$

that is, the inequality (3.24) holds. □

Lemma 3.6 *Suppose that the assumption A1 holds. If Algorithm A does not stop at the current iterate x^k , then (3.15) and (3.16) are satisfied for $t > 0$ small enough, so Algorithm A is well defined.*

Proof It follows from the Taylor expansion and (3.6) that

$$\begin{aligned} & P(x^k + td^k; \sigma_k) - P(x^k; \sigma_k) \\ &= t \nabla f(x^k)^\top d^k + \sigma_k \sum_{j \in \mathcal{E}} (|h_j(x^k) + t \nabla h_j(x^k)^\top d^k| - |h_j(x^k)|) + o(t) \\ &= \bar{P}(x^k; td^k; \sigma_k) - \bar{P}(x^k; 0; \sigma_k) + o(t). \end{aligned} \tag{3.27}$$

The second equality above is due to (3.7). From the convexity of $\bar{P}(x^k; d; \sigma_k)$ for d , we obtain

$$\bar{P}(x^k; td^k; \sigma_k) - \bar{P}(x^k; 0; \sigma_k) \leq t(\bar{P}(x^k; d^k; \sigma_k) - \bar{P}(x^k; 0; \sigma_k)), \tag{3.28}$$

which together with (3.27) and Lemma 3.4 gives for t small enough

$$P(x^k + td^k; \sigma_k) - P(x^k; \sigma_k) \leq t\alpha(\bar{P}(x^k; d^k; \sigma_k) - \bar{P}(x^k; 0; \sigma_k)),$$

where $\alpha \in (0, 1)$. Hence, (3.15) holds for sufficiently small $t > 0$.

In what follows, we prove (3.16) holds for sufficiently small $t > 0$. Since $\mathcal{A}(x)$ is twice continuously differentiable function, it follows from Taylor expansion that

$$\mathcal{A}(x^k + td^k) = \mathcal{A}(x^k) + tD\mathcal{A}(x^k)d^k + o(t) = \mathcal{A}(x^k) + O(t). \tag{3.29}$$

Note that the largest eigenvalue function $\lambda_{\max}(A) = \max_{\|v\|=1} v^T A v$, we deduce from (3.29) and $\mathcal{A}(x^k) < 0$ that

$$\lambda_{\max}(\mathcal{A}(x^k + t d^k)) = \max_{\|v\|=1} \{v^T \mathcal{A}(x^k) v + v^T O(t) v\} < 0$$

for $0 < t < 1$ small enough, which implies (3.16) holds for $0 < t < 1$ small enough.

By summarizing the above discussions, we conclude that Algorithm A is well defined. □

4 Global convergence

If Algorithm A terminates at x^k after a finite number of iterations, we know from Lemma 3.3 that x^k is a KKT point of NLSDP (1.1). In this section, without loss of generality, we assume that the sequence $\{x^k\}$ generated by Algorithm A is infinite. We will prove any accumulation point of $\{x^k\}$ is a stationary point or a KKT point of NLSDP (1.1), i.e., Algorithm A is globally convergent. We first generalize the definition of stationary point for nonlinear programming defined in [16] to nonlinear semidefinite programming.

Definition 4.1 Let $x \in R^n$, if there exist a matrix $\Lambda \in S^m$ and a vector $\mu \in R^l$ such that

$$\nabla_x L(x, \Lambda, \mu) = 0, \tag{4.1}$$

$$\Lambda \mathcal{A}(x) = 0, \quad \mathcal{A}(x) \leq 0, \quad h(x) = 0, \tag{4.2}$$

then x is called a stationary point of NLSDP (1.1).

In order to analyze the global convergence, some additional assumptions are required:

- A2 The sequence $\{x^k\}$ yielded by Algorithm A lies in a nonempty closed and bounded set \mathcal{X} .
- A3 The functions $f(x)$, $h(x)$ and $\mathcal{A}(x)$ are twice continuously differentiable on an open set containing \mathcal{X} .
- A4 There exists a positive constant λ^s such that $\lambda^s > \lambda^l$ and $\lambda^l I_m \leq \bar{\Lambda}_k \leq \lambda^s I_m$ for all k .
- A5 The matrix H_k is uniformly positive definite, i.e., there exist two positive constants a and b such that $a\|y\|^2 \leq y^T H_k y \leq b\|y\|^2$ for all $y \in R^n$.

Let x^* be an accumulation point of $\{x^k\}$, then there exists a subset $\mathcal{K} \subseteq \{1, 2, \dots\}$ such that $\lim_{k \in \mathcal{K}} x^k = x^*$. Without loss of generality, we suppose

$$\begin{aligned} H_k &\xrightarrow{\mathcal{K}} H_*, & \nabla h(x^k) &\xrightarrow{\mathcal{K}} \nabla h(x^*), \\ \bar{\Lambda}_k &\xrightarrow{\mathcal{K}} \bar{\Lambda}_*, & W(x^k, H_k, \bar{\Lambda}_k) &\xrightarrow{\mathcal{K}} W(x^*, H_*, \bar{\Lambda}_*), \end{aligned}$$

where $W(x^k, H_k, \bar{\Lambda}_k)$ is defined by (3.5) and

$$W(x^*, H_*, \bar{\Lambda}_*) \stackrel{\text{def}}{=} \begin{pmatrix} H_* & \nabla \mathcal{A}(x^*)^T & \nabla h(x^*) \\ (\bar{\Lambda}_* \otimes_s I_m) \nabla \mathcal{A}(x^*) & \mathcal{A}(x^*) \otimes_s I_m & 0 \\ \nabla h(x^*)^T & 0 & 0 \end{pmatrix}.$$

From the assumptions A2-A3, we obtain the following conclusions immediately.

Lemma 4.1 *Suppose the assumptions A2-A3 hold. Then there exists a constant $\bar{M} > 1$ such that $|f(y^k)| \leq \bar{M}$, $\|\nabla f(y^k)\| \leq \bar{M}$, $\|\nabla^2 f(y^k)\| \leq \bar{M}$, $\|h(y^k)\| \leq \bar{M}$, $\|\nabla h(y^k)\| \leq \bar{M}$, $\|\mathcal{A}(y^k)\|_F \leq \bar{M}$, $\|D\mathcal{A}(y^k)\|_F \leq \bar{M}$ and $\|D^2\mathcal{A}(y^k)\|_F \leq \bar{M}$, for any $y^k \in \mathcal{N}(x^k)$, where $\mathcal{N}(x^k)$ is a neighborhood of x^k .*

Lemma 4.2 *Suppose the assumptions A1-A5 hold. Then*

- (1) *there exists a constant $c_1 > 0$ such that $\|W(x^k, H_k, \bar{\Lambda}_k)^{-1}\| \leq c_1$ for any $k \in \mathcal{K}$;*
- (2) *there exists a constant $\hat{M} > 1$ such that $\|\lambda^{k_0}\| \leq \hat{M}$, $\|\lambda^{k_1}\| \leq \hat{M}$, $\|\mu^{k_0}\| \leq \hat{M}$, $\|\mu^{k_1}\| \leq \hat{M}$, $\|d^{k_0}\| \leq \hat{M}$ and $\|d^{k_1}\| \leq \hat{M}$ for any $k \in \mathcal{K}$.*

The following result is an important property of the penalty parameter σ_k , which is obtained by the updating rule (3.14).

Lemma 4.3 *Suppose the assumptions A1-A5 hold. Then the penalty parameter σ_k is updated only in a finite number of steps.*

Based on Lemma 4.3, in the rest of the paper, we assume, without loss of generality, that $\sigma_k \equiv \tilde{\sigma}$ for all k , where

$$\tilde{\sigma} > \sup_k \left\{ (3 - \xi) \max_{j \in \mathcal{E}} |\mu_j^{k_0}| \right\}.$$

By using of Lemma 4.2, we obtain the following result.

Lemma 4.4 *Suppose the assumptions A1-A5 hold. Then there exists a constant $c_2 > 0$ such that*

$$\|d^k - d^{k_0}\| \leq c_2 \|d^{k_0}\|. \tag{4.3}$$

For the sake of simplicity, in the rest of this section, let $(d^{*0}, \mu^{*0}, \lambda^{*0})$ be the solution of the following SLE in (d, μ, λ) :

$$\begin{cases} H_* d + \nabla \mathcal{A}(x^*)^T \lambda + \sum_{j \in \mathcal{E}} \mu_j \nabla h_j(x^*) = -\nabla f(x^*), \\ (\bar{\Lambda}_* \otimes_s I_m) \nabla \mathcal{A}(x^*) d + (\mathcal{A}(x^*) \otimes_s I_m) \lambda = 0, \\ \nabla h_j(x^*)^T d = -h_j(x^*), \quad j \in \mathcal{E}. \end{cases} \tag{4.4}$$

Let $(d^{*1}, \mu^{*1}, \lambda^{*1})$ be the solution of the following SLE in (d, μ, λ) :

$$\begin{cases} H_* d + \nabla \mathcal{A}(x^*)^T \lambda + \sum_{j \in \mathcal{E}} \mu_j \nabla h_j(x^*) = -\nabla f(x^*), \\ (\bar{\Lambda}_* \otimes_s I_m) \nabla \mathcal{A}(x^*) d + (\mathcal{A}(x^*) \otimes_s I_m) \lambda = -\bar{\lambda}_* \|d^{*0}\|, \\ \nabla h_j(x^*)^T d = -h_j(x^*), \quad j \in \mathcal{E}. \end{cases} \tag{4.5}$$

From the above equalities and Lemma 4.2, we obtain the following conclusion.

Lemma 4.5 *Suppose the assumptions A1-A5 hold, and $\delta_k \xrightarrow{\mathcal{K}} \delta_*$. Then*

- (i) $d^{k_0} \xrightarrow{\mathcal{K}} d^{*0}$, $\mu^{k_0} \xrightarrow{\mathcal{K}} \mu^{*0}$, $\lambda^{k_0} \xrightarrow{\mathcal{K}} \lambda^{*0}$,

- (ii) $d^{k1} \xrightarrow{\mathcal{K}} d^{*1}, \mu^{k1} \xrightarrow{\mathcal{K}} \mu^{*1}, \lambda^{k1} \xrightarrow{\mathcal{K}} \lambda^{*1},$
- (iii) $d^{*0} = 0$ if and only if $d^* = 0$ where $d^* = (1 - \delta_*)d^{*0} + \delta_*d^{*1}.$

Remark 4.1 By (3.13), we know that $\{\delta_k\}$ is bounded, so in the rest of the paper, we assume, without loss of generality, that $\delta_k \xrightarrow{\mathcal{K}} \delta_*.$

Lemma 4.6 *Suppose the assumptions A1-A5 hold. Let x^* be an accumulation point of the sequence $\{x^k\}$ and $x^k \xrightarrow{\mathcal{K}} x^*.$ If $d^k \xrightarrow{\mathcal{K}} 0,$ then x^* is a KKT point or a stationary point of NLSDP (1.1), and $\lambda^k \xrightarrow{\mathcal{K}} \text{svec}(\Lambda^*), \mu^k \xrightarrow{\mathcal{K}} \mu^*,$ where (Λ^*, μ^*) is the Lagrangian multiplier corresponding to $x^*.$*

Proof It is clear from Lemma 4.2 that $\{\lambda^k\}$ and $\{\mu^k\}$ are bounded. Assume that $\hat{\lambda}, \hat{\mu}$ are accumulation points of $\{\lambda^k\}$ and $\{\mu^k\},$ respectively. Without loss of generality, we assume that $\lambda^k \xrightarrow{\mathcal{K}} \hat{\lambda}$ and $\mu^k \xrightarrow{\mathcal{K}} \hat{\mu}.$

Obviously, (d^k, λ^k, μ^k) satisfies the SLE (3.25a)-(3.25c). By taking the limit on \mathcal{K} in (3.25a)-(3.25c), we obtain

$$\nabla \mathcal{A}(x^*) \hat{\lambda} + \sum_{j \in \mathcal{E}} \hat{\mu}_j \nabla h_j(x^*) = -\nabla f(x^*), \tag{4.6a}$$

$$(\mathcal{A}(x^*) \otimes_s I) \hat{\lambda} = 0, \tag{4.6b}$$

$$h_j(x^*) = 0, \quad j \in \mathcal{E}. \tag{4.6c}$$

If $x^* \in \mathcal{F}_0,$ i.e., $\mathcal{A}(x^*) < 0,$ then we know from Lemma 2.1(4) that $\mathcal{A}(x^*) \otimes_s I$ is nonsingular, so the equation (4.6b) has a unique solution $\hat{\lambda} = 0.$ Let $\hat{\Lambda} := \text{smat}(\hat{\lambda}) = 0,$ so $\hat{\Lambda} \mathcal{A}(x^*) = 0.$ Together with (4.6a) and (4.6c), we conclude that x^* is a KKT point of NLSDP (1.1).

If $x^* \in \Omega \setminus \mathcal{F}_0,$ let $\hat{\Lambda} := \text{smat}(\hat{\lambda}).$ It follows from (4.6b) that $\text{sym}(\hat{\Lambda} \mathcal{A}(x^*)) = 0,$ which means that $\hat{\Lambda} \mathcal{A}(x^*)$ is a skw-symmetric matrix. Hence $\text{Tr}(\hat{\Lambda} \mathcal{A}(x^*)) = 0.$ According to Remark 2.2, we obtain $\hat{\Lambda} \mathcal{A}(x^*) = 0.$ Combining with (4.6a) and (4.6c), x^* is a stationary point of NLSDP (1.1). (λ^*, μ^*) is the Lagrangian multiplier corresponding to $x^*,$ that is,

$$\nabla \mathcal{A}(x^*)^T \lambda^* + \sum_{j \in \mathcal{E}} \mu_j^* \nabla h_j(x^*) = -\nabla f(x^*),$$

$$\Lambda^* \mathcal{A}(x^*) = 0,$$

where $\Lambda^* = \text{smat}(\lambda^*).$ It is not difficult to verify that (λ^*, μ^*) is the solution of the following SLE:

$$\nabla \mathcal{A}(x^*)^T \lambda^* + \sum_{j \in \mathcal{E}} \mu_j^* \nabla h_j(x^*) = -\nabla f(x^*), \tag{4.7a}$$

$$(\mathcal{A}(x^*) \otimes_s I) \lambda^* = 0. \tag{4.7b}$$

From (4.6a)-(4.6c), we know that $(\hat{\lambda}, \hat{\mu})$ is also the solution of (4.7a)-(4.7b). It is clear from the assumption A1 that the solution of (4.7a)-(4.7b) is unique, therefore, $\hat{\lambda} = \lambda^*, \hat{\mu} = \mu^*.$ The proof is completed. □

Based on Lemma 4.6, the following conclusion is obvious.

Lemma 4.7 *Suppose the assumptions A1-A5 hold. Let $x^k \xrightarrow{\mathcal{K}} x^*$. If $d^{k-1} \xrightarrow{\mathcal{K}} 0$, then x^* is a KKT point or a stationary point of NLSDP (1.1).*

Lemma 4.8 *Suppose the assumptions A1-A5 hold, $x^k \xrightarrow{\mathcal{K}} x^*$. If $\inf_{\mathcal{K}} \{\|d^{k-1}\|\} > 0$, then $d^k \xrightarrow{\mathcal{K}} 0$.*

Proof By contradiction, we assume that there exist a subset $\mathcal{K}' \subset \mathcal{K}$ and a constant $\bar{d} > 0$ such that $\|d^k\| \geq \bar{d}, \forall k \in \mathcal{K}'$ large enough. From the assumptions A1-A5, (3.13) and the updating rule of $\bar{\Lambda}_k$, we assume without loss of generality that $H_k \xrightarrow{\mathcal{K}'} H_*, \delta_k \xrightarrow{\mathcal{K}'} \delta_*, \bar{\Lambda}_k \xrightarrow{\mathcal{K}'} \bar{\Lambda}_*$. On the other hand, it follows from the updating rule of $\bar{\Lambda}_k$ and the assumption A4 that $\bar{\Lambda}_*$ is positive definite. According to Lemma 4.5(iii), there exists $\underline{d} > 0$ such that $\|d^{k_0}\| \geq \underline{d}$ for all $k \in \mathcal{K}'$.

Firstly, we show that there exists $\underline{t} > 0$ independent of k such that (3.15) and (3.16) are satisfied for all $t \geq \underline{t}$. For any $k \in \mathcal{K}'$, it is clear from the assumptions A1 and A5 and Lemmas 3.3-3.4 and Lemmas 4.1-4.2 that

$$\bar{P}(x^k; d^k; \bar{\sigma}) - \bar{P}(x^k; 0; \bar{\sigma}) \leq -\xi a \underline{d}^2. \tag{4.8}$$

Together with (3.27)-(3.28), there exists $t_f > 0$ independent of k such that

$$P(x^k + td^k; \bar{\sigma}) - P(x^k; \bar{\sigma}) \leq t\alpha [\bar{P}(x^k; d^k; \bar{\sigma}) - \bar{P}(x^k; 0; \bar{\sigma})] \tag{4.9}$$

for all $k \in \mathcal{K}'$ and $t \in (0, t_f]$, where $\alpha \in (0, 1)$. The above inequality shows the inequality (3.15) holds.

We next prove the inequality (3.16) holds. It follows from (3.8) and Lemma 2.1(4) and Lemma 2.4 that

$$\begin{aligned} & |\nabla f(x^k)^T d^{k_0} + (\mu^{k_0})^T h(x^k)| \\ &= |-(d^{k_0})^T H_k d^{k_0} + (\lambda^{k_0})^T ((\bar{\Lambda}_k \otimes_s I_m)^{-1} (\mathcal{A}(x^k) \otimes_s I_m))^T \lambda^{k_0}| \\ &\geq a \|d^{k_0}\|^2. \end{aligned}$$

Combining with Lemmas 4.1-4.2 and (3.13), there exists a constant $0 < \tilde{\delta} \leq 1$ such that $\delta_k \geq \tilde{\delta}$ for $k \in \mathcal{K}'$. By the mean-value theorem and Lemmas 4.1-4.2, we obtain

$$\begin{aligned} \mathcal{A}(x^k + td^k) &= \mathcal{A}(x^k) + tD\mathcal{A}(x^k)d^k + t^2(D^2\mathcal{A}(x + t\vartheta d^k)(d^k, d^k)) \\ &\leq \mathcal{A}(x^k) + tD\mathcal{A}(x^k)d^k + t^2M^3I_m \end{aligned} \tag{4.10}$$

for any $k \in \mathcal{K}'$, where $\vartheta \in (0, 1), M = \max\{\widehat{M}, \overline{M}\}$. Let $N(t; x^k) = \mathcal{A}(x^k) + tD\mathcal{A}(x^k)d^k + t^2M^3I_m$, the above inequality is rewritten as

$$\mathcal{A}(x^k + td^k) \leq N(t; x^k), \tag{4.11}$$

thus, in order to prove that $\mathcal{A}(x^k + td^k)$ is negative definite, it is sufficient to prove that $N(t; x^k)$ is negative definite. In view of $\bar{\Lambda}_k > 0$, the definition (2.2) of sym and Lemma 2.2,

it is sufficient to show that there exists $t_{\mathcal{A}} > 0$ independent of k such that

$$\text{sym}(\overline{\Lambda}_k N(t; x^k)) < 0, \quad \forall t \in (0, t_{\mathcal{A}}]. \tag{4.12}$$

In view of (2.10), (2.5) and Lemma 2.1(1), we obtain

$$(\overline{\Lambda}_k \otimes_s I_m) \nabla \mathcal{A}(x^k) d^k = \text{svec}(\text{sym}(\overline{\Lambda}_k D \mathcal{A}(x^k) d^k)). \tag{4.13}$$

Let $\Lambda^k = \text{smat}(\lambda^k)$, i.e., $\lambda^k = \text{svec}(\Lambda^k)$, it is obvious from (2.5) that

$$(\mathcal{A}(x^k) \otimes_s I_m) \lambda^k = (\mathcal{A}(x^k) \otimes_s I_m) \text{svec}(\Lambda^k) = \text{svec}(\text{sym}(\Lambda^k \mathcal{A}(x^k))). \tag{4.14}$$

Hence, (4.13), (4.14) and (3.25b) give rise to

$$\begin{aligned} & \text{sym}(\overline{\Lambda}_k D \mathcal{A}(x^k) d^k + \Lambda^k \mathcal{A}(x^k)) \\ &= \text{smat}(\text{svec}(\text{sym}(\overline{\Lambda}_k D \mathcal{A}(x^k) d^k)) + \text{svec}(\text{sym}(\Lambda^k \mathcal{A}(x^k)))) \\ &= \text{smat}(-\delta_k \overline{\lambda}_k \|d^{k0}\|) = -\delta_k \|d^{k0}\| \overline{\Lambda}_k. \end{aligned}$$

Based on the above equality, we have

$$\begin{aligned} \text{sym}(\overline{\Lambda}_k N(t; x^k)) &= \text{sym}(\overline{\Lambda}_k (\mathcal{A}(x^k) + t D \mathcal{A}(x^k) d^k + t^2 M^3 I_m)) \\ &= \text{sym}((\overline{\Lambda}_k - t \Lambda^k) \mathcal{A}(x^k)) + (t^2 M^3 \overline{\Lambda}_k - t \delta_k \|d^{k0}\| \overline{\Lambda}_k) \\ &< \text{sym}((\overline{\Lambda}_k - t \Lambda^k) \mathcal{A}(x^k)) + (2t^2 M^3 - t \delta \underline{d}) \overline{\Lambda}_k; \end{aligned} \tag{4.15}$$

note the positive definiteness of $\overline{\Lambda}_k$, hence, if

$$\max\{v^T((\overline{\Lambda}_k - t \Lambda^k) \mathcal{A}(x^k))v : v \in \mathbb{R}^m, \|v\| = 1\} \leq 0, \quad \text{for any } k \in \mathcal{K}', \tag{4.16}$$

then (4.12) holds for $t \leq \frac{\delta \underline{d}}{2M^3}$.

Since $\overline{\Lambda}_k$ and $\mathcal{A}(x^k)$ are symmetric and commuting, there exists an orthogonal matrix Q_k such that

$$\overline{\Lambda}_k = Q_k^T \overline{D}_\lambda^k Q_k, \quad \mathcal{A}(x^k) = Q_k^T D_{\mathcal{A}}^k Q_k, \tag{4.17}$$

where \overline{D}_λ^k and $D_{\mathcal{A}}^k$ are diagonal matrices. Then $(\overline{\Lambda}_k - t \Lambda^k) \mathcal{A}(x^k) = Q_k^T (\overline{D}_\lambda^k - t Q_k \Lambda^k Q_k^T) \times D_{\mathcal{A}}^k Q_k$. Let $\tilde{\Lambda}^k = Q_k \Lambda^k Q_k^T$, so in order to prove (4.16), it is enough to show that there exists a constant $t_{\mathcal{A}} > 0$ such that

$$v^T((\overline{D}_\lambda^k - t \tilde{\Lambda}^k) D_{\mathcal{A}}^k) v \leq 0, \quad \forall v: \|v\| = 1, \tag{4.18}$$

for any $t \in (0, t_{\mathcal{A}})$ and $k \in \mathcal{K}'$. By Lemma 4.6 and $\Lambda^k = \text{smat}(\lambda^k)$, we know $\{\Lambda^k\}$ is bounded, furthermore, $\{\tilde{\Lambda}^k\}$ is also bounded. Let $\tilde{\Lambda}^*$ be an accumulation point of $\{\tilde{\Lambda}^k\}$. Without loss of generality, we assume that $\tilde{\Lambda}^k \xrightarrow{\mathcal{K}'} \tilde{\Lambda}^*$. Let $B^k = \tilde{\Lambda}^k - \tilde{\Lambda}^*$, obviously, $B^k \xrightarrow{\mathcal{K}'} 0$, thus there exists $\gamma > 0$ such that

$$|v^T(B^k D_{\mathcal{A}}^k)v| < \gamma \tag{4.19}$$

for any $k \in \mathcal{K}'$. Note that

$$v^T (\overline{D}_\lambda^k - t\tilde{\Lambda}^k) D_{\mathcal{A}}^k v = v^T (\overline{D}_\lambda^k - t\tilde{\Lambda}^*) D_{\mathcal{A}}^k v - tv^T (B^k D_{\mathcal{A}}^k) v. \tag{4.20}$$

It follows from the assumption A4 that all eigenvalues of \overline{D}_λ^k are between λ^l and λ^s for all k . According to Weyl's theorem (see [6]), there exists $t_1 > 0$ such that all eigenvalues of $(\overline{D}_\lambda^k - t\tilde{\Lambda}^*)$ are positive for any $t \in (0, t_1]$. We also know from $\mathcal{A}(x^k) < 0$ and the second equality in (4.17) that $D_{\mathcal{A}}$ is negative definite. Therefore, for any v with $\|v\| = 1$ and $t \in (0, t_1]$, it follows from Lemma 2.3 that $(\overline{D}_\lambda^k - t\tilde{\Lambda}^*) D_{\mathcal{A}}^k$ is also negative definite. Combining with (4.19), for any v with $\|v\| = 1$ and any $t \in (0, t_1]$, we obtain

$$v^T ((\overline{D}_\lambda^k - t\tilde{\Lambda}^*) D_{\mathcal{A}}^k) v - tv^T (B^k D_{\mathcal{A}}^k) v \leq 0, \tag{4.21}$$

together with (4.20) shows that (4.18) is satisfied, further, (4.16) and (4.12) hold.

Let $t_{\mathcal{A}} = \min\{t_1, \frac{md}{2M^3}\}$, thus (4.12) holds for any $t \in (0, t_{\mathcal{A}}]$. Hence, we see that $\mathcal{A}(x^k + td^k) < 0$ holds for $t \in (0, t_{\mathcal{A}}]$ and any $k \in \mathcal{K}'$. Let $\underline{t} = \min\{t_f, t_{\mathcal{A}}\}$, for any $\underline{t} \in (0, \underline{t}]$, (3.15) and (3.16) are satisfied for all $t \geq \underline{t}$. Combining with (4.8) and (4.9), we obtain for any $k \in \mathcal{K}'$

$$P(x^{k+1}; \tilde{\sigma}) \leq P(x^k; \tilde{\sigma}) - \underline{t}\alpha\xi\underline{a}\underline{d}^2. \tag{4.22}$$

On the other hand, the sequence $\{P(x^k; \tilde{\sigma})\}$ decreases monotonically and $P(x^k; \tilde{\sigma}) \xrightarrow{\mathcal{K}'} P(x^*; \tilde{\sigma})$, so $\{P(x^k; \tilde{\sigma})\}_{k=1}^\infty$ is convergent. Let $\lim_{k \rightarrow \infty} P(x^k; \tilde{\sigma}) = \varrho$ and taking the limit in the above inequality, we have $-\underline{t}\xi\alpha\underline{a}\underline{d}^2 \geq 0$, which is a contradiction. Hence, $d^k \xrightarrow{\mathcal{K}'} 0$. \square

Based on Lemmas 4.6-4.8, the following global convergence of Algorithm A is immediate.

Theorem 4.1 *Suppose the assumptions A1-A5 hold. Then Algorithm A either terminates in a finite number of iterations at a KKT point of the NLSDP (1.1), or it generates a sequence $\{x^k\}$ whose every accumulation point is a KKT point or a stationary point of the NLSDP (1.1).*

5 Numerical experiments

Algorithm A has been implemented in Matlab 2011b and the codes have been run on a 3.40 GHz Intel(R) Core(TM)i3-3240 machine with a Windows 7 system. We choose H_0 as n -order identical matrix and at each iteration, H_k is updated by the damped BFGS formula in [15] and $\overline{\Lambda}_0$ as m -order identical matrix. In the numerical experiments, we choose the parameters as follows:

$$\begin{aligned} \alpha &= 0.25, & \beta &= 0.5, & \xi &= 0.5, & \lambda^l &= 0.5, \\ \sigma_{-1} &= 0.5, & \rho_1 &= 1, & \rho_2 &= 2. \end{aligned}$$

The stop criterion is $\|d^{k_0}\| \leq 10^{-4}$.

The test problems are described as follows:

I. The first test problem is Rosen-Suzuki problem [29] combined with a negative semidefinite constraint and denoted by *CM*:

$$\begin{aligned} \min f_0(x) &= x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4 \\ \text{s.t. } x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 - x_2 + x_3 - x_4 - 8 &= 0, \\ x_1^2 + 2x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_4 - 9 &= 0, \\ 2x_1^2 + x_2^2 + x_3^2 + 2x_1 - x_2 - x_4 - 5 &= 0, \\ \begin{pmatrix} -x_2 - x_3 & 0 & 0 & 0 \\ 0 & -2x_4 & -x_1 & 0 \\ 0 & -x_1 & -2x_4 & 0 \\ 0 & 0 & 0 & -x_2 - x_3 \end{pmatrix} &\preceq 0. \end{aligned}$$

II. We select some test problems from [7] only with equality constraints and we add a negative semidefinite matrix constraint.

- (1) We select the problems HS6, HS7, HS8, HS9 combined with the following 2×2 order symmetric matrix which comes from [14] and rename them MHS6, MHS7, MHS8 and MHS9, respectively:

$$\begin{pmatrix} -x_1^2 & -\frac{x_1}{2} \\ -\frac{x_1}{2} & -x_2^2 \end{pmatrix} \preceq 0.$$

- (2) Choose the problems HS26, HS27, HS28 and HS61 combined with the following 3×3 order symmetric matrix and rename them MHS26, MHS27, MHS28 and MHS61, respectively:

$$\begin{pmatrix} -x_1^2 & -\frac{x_1}{2} & 0 \\ -\frac{x_1}{2} & -x_2^2 & 0 \\ 0 & 0 & -x_3^4 \end{pmatrix} \preceq 0.$$

- (3) Choose the problems HS40, HS42, HS47, HS48, HS50, HS51, HS77 and HS79, adding the negative semidefinite matrix constraint in the problem *CM* and renaming them MHS40, MHS42, MHS47, MHS48, MHS50, MHS51, MHS77 and MHS79.

III. Nearest correlation matrix problem (NCM for short) (see [23]):

$$\begin{aligned} \min f(X) &= \frac{1}{2} \|X - A\|_F \\ \text{s.t. } X &\succeq \epsilon I, \\ X_{ii} &= 1, \quad i = 1, 2, \dots, m, \end{aligned}$$

where $A \in \mathcal{S}^m$ is given. In NCM problem, eigenvalues of X should not be less than ϵ , and the diagonal elements of X are equal to 1. Elements of the matrix A are uniform random numbers in $[-1, 1]$ with $A_{ii} = 1, i = 1, 2, \dots, m$. Set $\epsilon = 10^{-3}$. In addition, we compare with the results of [23] (Algo. SDPIP for short) and [24] (Algo. YYNY for short), and their results from [24].

Table 1 The numerical results of test problems I and II

Problem	n	l	m	x^0	Iter.	NF	NC	f_{final}	Time (s)
CM	4	3	4	$(2.5, 2.5, 2.5, -2.5)^T$	19	72	72	$-4.400000e+001$	$4.097408e-001$
PHS6	2	1	2	$(-2, -2)^T$	99	128	128	$1.226381e-006$	$3.541575e-001$
PHS7	2	1	2	$(1, 5)^T$	43	169	169	$-1.732051e+000$	$3.551911e-001$
PHS8	2	2	2	$(1, 4)^T$	4	4	4	-1	$2.195229e-001$
PHS9	2	1	2	$(-4, 4)^T$	2	2	2	$-4.999996e-001$	$2.025914e-001$
PHS26	3	1	3	$(1.5, 1.5, 1.5)^T$	28	28	28	$3.726010e-005$	$2.514937e-001$
PHS27	3	1	3	$(-1, 1, 1)^T$	17	17	17	$5.426241e-002$	$2.354974e-001$
PHS28	3	1	3	$(1, -1, -1)^T$	6	6	6	$6.756098e-001$	$1.708627e-001$
PHS40	4	3	4	$(0.5, 0.5, 0.5, 0.5)^T$	8	10	10	$-2.500001e-001$	$2.773717e-001$
PHS42	4	2	4	$(-1, 1, 1, 1)^T$	17	28	28	$1.385766e+001$	$2.415490e-001$
PHS47	5	3	4	$(-1, 1, 1, 1, 1)^T$	31	80	80	$2.910505e-001$	$2.642828e-001$
PHS48	5	2	4	$(3, 3, 3, 3, -3)^T$	49	140	140	$3.060758e-008$	$2.962501e-001$
PHS50	5	3	4	$(-3, 3, 3, 3, 3)^T$	23	84	84	$2.390072e-009$	$3.139633e-001$
PHS51	5	3	4	$(-1, 1, 1, 1, 1)^T$	13	14	14	$4.687353e-008$	$2.302719e-001$
PHS61	3	2	3	$(2.5, 2.5, 2.5)^T$	59	59	59	$-8.191909e+001$	$3.401501e-001$
PHS77	5	2	4	$(1, 1, 1, 1, 1)^T$	23	25	25	$2.415051e-001$	$2.393263e-001$
PHS79	5	3	4	$(-1, 1, 1, 1, 1)^T$	44	50	50	$7.877716e-002$	$3.415668e-001$

Table 2 The numerical results for NCM problem

n	l	m	Algorithm	Iter.	NF	NC
10	5	5	Algo. A	8	15	15
			Algo. YNYN	8	-	-
			Algo. SDPIP	9	-	-
45	10	10	Algo. A	10	19	19
			Algo. YNYN	8	-	-
			Algo. SDPIP	10	-	-
105	15	15	Algo. A	10	20	20
			Algo. YNYN	10	-	-
			Algo. SDPIP	11	-	-
190	20	20	Algo. A	10	18	18
			Algo. YNYN	11	-	-
			Algo. SDPIP	12	-	-
300	25	25	Algo. A	10	25	25
			Algo. YNYN	10	-	-
			Algo. SDPIP	11	-	-
435	30	30	Algo. A	10	19	19
			Algo. YNYN	9	-	-
			Algo. SDPIP	10	-	-
595	35	35	Algo. A	11	25	25
			Algo. YNYN	11	-	-
			Algo. SDPIP	11	-	-
780	40	40	Algo. A	11	24	24
			Algo. YNYN	11	-	-
			Algo. SDPIP	11	-	-
1,225	50	50	Algo. A	12	34	34
			Algo. YNYN	-	-	-
			Algo. SDPIP	-	-	-

The numerical results are listed in Table 1 and Table 2. The meanings of the notations in Table 1 and Table 2 are as follows:

- n : the number of variables;
- l : the number of equality constraints;
- m : the dimensionality of the negative semidefinite matrix;
- Iter.: the number of iterations;

- NF: the number of evaluations for $f(x)$;
 NC: the number of evaluations for all constraint functions;
 f_{final} : the optimal value;
 Time (s): the time of calculation;
 -: means that the result is not given.

6 Concluding remarks

We have presented a globally convergent QP-free algorithm for nonlinear SDP problems. Based on KKT conditions of nonlinear SDP problems and techniques of perturbation, we construct two SLEs skillfully. Under some linear independence condition, the SLEs have unique solution. At each iteration, the search direction is yielded by solving two SLEs with the same coefficient matrix; some penalty function is used as the merit function for line search and the penalty parameter is updated automatically in the algorithm. The preliminary numerical results show that the proposed algorithm is effective and comparable.

Acknowledgements

Project supported by the Natural Science Foundation of China (No. 11561005), the Natural Science Foundation of Guangxi Province (Nos. 2016GXNSFAA380248, 2014GXNSFFA118001).

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors read and approved the final manuscript.

Author details

¹College of Mathematics and Information Science, Guangxi University, Daxue Road 100, Nanning, Guangxi 530004, China.

²School of Mathematics and Statistics, Guangxi Colleges and Universities Key Laboratory of Complex System Optimization and Big Data Processing, Yulin Normal University, Jiaoyu Road, Yulin, Guangxi 537000, China.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 5 March 2017 Accepted: 5 June 2017 Published online: 23 June 2017

References

- Jarre, F: An interior point method for semidefinite programming. *Optim. Eng.* **1**, 347-372 (2000)
- Ben, TA, Jarre, F, Kocvara, M, Nemirovski, A, Zowe, J: Optimization design of trusses under a nonconvex global buckling constraint. *Optim. Eng.* **1**, 189-213 (2000)
- Wolkowicz, H, Saigal, R, Vandenberghe, L (eds.): *Handbook of Semidefinite Programming*. Kluwer Academic, Boston (2000)
- Freund, RW, Jarre, F, Vogelbusch, CH: Nonlinear semidefinite programming: sensitivity, convergence, and an application in passive reduced-order modeling. *Math. Program.* **109**, 581-611 (2007)
- Gao, ZY, He, GP, Wu, F: Sequential systems of linear equation algorithm with arbitrary initial point. *Sci. China Ser. A* **27**, 24-33 (1997) (in Chinese)
- Horn, RA, Johnson, CR: *Matrix Analysis*. Cambridge University Press, Cambridge (1985)
- Hock, W, Schittkowski, K: *Test Examples for Nonlinear Programming Codes*. Lectures Notes in Economics and Mathematical Systems, vol. 187. Springer, Berlin (1981)
- Jian, JB, Quan, R, Cheng, WX: A feasible QP-free algorithm combining the interior point method with active set for constrained optimization. *Comput. Math. Appl.* **58**, 1520-1533 (2009)
- Kanzow, C, Nagel, C, Kato, H, Fukushima, M: Successive linearization methods for nonlinear semidefinite programs. *Comput. Optim. Appl.* **31**, 251-273 (2005)
- Kovara, M, Stingl, M: PENNON: a code for convex nonlinear and semidefinite programming. *Optim. Methods Softw.* **18**, 317-333 (2003)
- Luo, HZ, Wu, HX, Chen, GT: On the convergence of augmented Lagrangian methods for nonlinear semidefinite programming. *J. Glob. Optim.* **54**, 599-618 (2012)
- Li, JL, Lv, J, Jian, JB: A globally and superlinearly convergent primal-dual interior point method for general constrained optimization. *Numer. Math., Theory Methods Appl.* **8**, 313-335 (2015)
- Li, JL, Huang, RS, Jian, JB: A superlinearly convergent QP-free algorithm for mathematical programs with equilibrium constraints. *Appl. Math. Comput.* **269**, 885-903 (2015)
- Noll, D: Local convergence of an augmented Lagrangian method for matrix inequality constrained programming. *Optim. Methods Softw.* **22**, 777-802 (2007)

15. Powell, MJD: A fast algorithm for nonlinearly constrained optimization calculations. In: Numerical Analysis. Lecture Notes in Mathematics, vol. 630, pp. 144-157. Springer, Berlin (1978)
16. Panier, ER, Tits, RL, Herskovits, N: A QP-free globally convergent, locally superlinear convergent algorithm for inequality constrained optimization. *SIAM J. Optim.* **26**, 788-811 (1988)
17. Qi, HD, Qi, LQ: A new QP-free, globally convergent, locally superlinearly convergent algorithm for inequality constrained optimization. *SIAM J. Optim.* **11**, 113-132 (2000)
18. Shapiro, A: First and second order analysis of nonlinear semidefinite programs. *Math. Program.* **77**, 301-320 (1997)
19. Sun, DF, Sun, J, Zhang, LW: The rate of convergence of the augmented Lagrangian method for nonlinear semidefinite programming. *Math. Program.* **114**, 349-391 (2008)
20. Sun, J, Zhang, LW, Wu, Y: Properties of the augmented Lagrangian in nonlinear semidefinite optimization. *J. Optim. Theory Appl.* **129**, 437-456 (2006)
21. Todd, MJ, Toh, KC, Tütüncü, RH: On the Nesterov-Todd direction in semidefinite programming. *SIAM J. Optim.* **8**, 769-796 (1998)
22. Wu, HX, Luo, HZ, Ding, XD, Chen, GT: Global convergence of modified augmented Lagrangian methods for nonlinear semidefinite programming. *Comput. Optim. Appl.* **56**, 531-558 (2013)
23. Yamashita, H, Yabe, H, Harada, K: A primal-dual interior point method for nonlinear semidefinite programming. *Math. Program., Ser. A* **135**, 89-121 (2012)
24. Yamakawa, Y, Yamashita, N, Yabe, H: A differentiable merit function for the shifted perturbed Karush-Kuhn-Tucker conditions of the nonlinear semidefinite programming. *Pac. J. Optim.* **11**, 557-579 (2015)
25. Zhu, ZB, Zhu, HL: A filter method for nonlinear semidefinite programming with global convergence. *Acta Math. Sin.* **30**, 1810-1826 (2014)
26. Correa, R, Ramirez, H: A global algorithm for nonlinear semidefinite programming. *SIAM J. Optim.* **15**, 303-318 (2004)
27. Fares, B, Noll, D, Apkarian, P: Robust control via sequential semidefinite programming. *SIAM J. Control Optim.* **40**, 1791-1820 (2002)
28. Aroztegui, M, Herskovits, J, Roche, JR, Bažan, E: A feasible direction interior point algorithm for nonlinear semidefinite programming. *Struct. Multidiscip. Optim.* **50**, 1019-1035 (2014)
29. Chen, ZW, Miao, SC: A penalty-free method with trust region for nonlinear semidefinite programming. *Asia-Pac. J. Oper. Res.* **32**, 1-24 (2015)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
