Asia-Pacific Journal of Operational Research
Vol. 33, No. 4 (2016) 1650025 (26 pages)

DOI: 10.1142/S0217595916500251

A Sequential Convex Program Approach to an Inverse Linear Semidefinite Programming Problem

Jia Wu*

Institute of Operations Research and Control Theory
School of Mathematical Sciences
Dalian University of Technology
Dalian 116024, P. R. China
wujia@dlut.edu.cn

Yi Zhang

Department of Mathematics, School of Science
East China University of Science and Technology
Shanghai 200237, P. R. China
yizhang@ecust.edu.cn

Liwei Zhang

Institute of Operations Research and Control Theory
School of Mathematical Sciences
Dalian University of Technology
Dalian 116024, P. R. China
lwzhang@dlut.edu.cn

Yue Lu

School of Mathematical Sciences
Tianjin Normal University
Tianjin 300387, P. R. China
jinjin403@sina.com

Received 8 January 2015
Revised 26 July 2015
Accepted 3 December 2015
Published 18 July 2016

This paper is devoted to the study of solving method for a type of inverse linear semidefinite programming problem in which both the objective parameter and the right-hand side parameter of the linear semidefinite programs are required to adjust. Since such kind of inverse problem is equivalent to a mathematical program with semidefinite cone complementarity constraints which is a rather difficult problem, we reformulate it as a nonconvex semi-definte programming problem by introducing a nonsmooth partial penalty function to penalize the complementarity constraint. The penalized problem is

*Corresponding author.

actually a nonsmooth DC programming problem which can be solved by a sequential convex program approach. Convergence analysis of the penalty models and the sequential convex program approach are shown. Numerical results are reported to demonstrate the efficiency of our approach.

Keywords: Inverse linear semidefinite programming problems; mathematical program with semidefinite cone complementarity constraints; penalty methods; sequential convex program.

1. Introduction

Consider the following linear semidefinite programming problem:

$$\begin{aligned} \mathrm{LSDP(c, \mathcal{A}, B)} \ \min_{x\in\Re^m} \ & c^T x \\ \text{s.t.} \ \ & \mathcal{A}x - B \in S^n_-, \end{aligned} \tag{1}$$

where $c \in \Re^m, B \in S^n$ and $\mathcal{A} : \Re^m \to S^n$ is a linear operator defined by

$$\mathcal{A}x = \sum_{i=1}^{m} x_i A_i, \quad x \in \Re^n,$$

with $A_i \in S^n$ for $i = 1, \ldots, m$. Here $S^n_+(S^n_-)$ is the cone of $n \times n$ positive(negative) semidefinite matrices in S^n(the space of $n \times n$ real symmetric matrices). It can be easily obtained that the adjoint of $\mathcal{A}$, denoted by $\mathcal{A}^*$, has the form

$$\mathcal{A}^*(X) = \begin{bmatrix} \langle A_1, X\rangle \\ \vdots \\ \langle A_m, X\rangle \end{bmatrix}, \quad X \in S^n.$$

For simplicity of notations, we introduce "SOL" as a mapping whose variables are problems and denote "SOL(P)" as the set of optimal solutions to a problem (P).

Given a feasible point $x^0 \in \Re^m$, which should be the optimal solution to Problem LSDP(c, $\mathcal{A}$, B) and points $c^0 \in \Re^m, B^0 \in S^n$ which are estimates to c and B, respectively. The inverse linear semidefinite programming problem considered in this paper is to find points $c \in \Re^m$ and $B \in S^n$ to solve

$$\begin{aligned} \text{Inverse LSDP} \ \min_{c,B} \ & \frac{1}{2}\{\|c - c^0\|^2 + \|B - B^0\|_F^2\} \\ \text{s.t.} \ \ & x^0 \in \mathrm{SOL(LSDP(c, \mathcal{A}, B))}, \end{aligned} \tag{2}$$

where $\|\cdot\|_F$ is the Frobenius norm which is defined by $\|X\|_F := \sqrt{\langle X, X\rangle}$ for $X \in S^n$.

In many optimization models, we usually assume that the parameters associated with decision variables in the objective function or in the constraint set are known and we need to find an optimal solution to it. Such problems are referred to forward problems. An inverse optimization problem is to find values of parameters which make the known solutions optimal and which differ from the given estimates as little as possible. As described above, the inverse model considered here is to give the values of "c" and "B" in $\mathrm{LSDP}(\mathrm{c}, \mathcal{A}, \mathrm{B})$ such that a given feasible solution x^0 becomes an optimal solution under the new parameter values and the adjustments of the values of the parameters in the model are as little as possible from the estimates.

The interest in inverse optimization problems was initiated by the paper (Burton and Toint, 1992) dealing with an inverse shortest path problem. Since then there are many contributions to inverse optimization, and a large number of inverse combinatorial optimization problems have been studied, see the survey paper by Heuberger (2004), Ahuja and Orlin (2001; 2002) Burkard *et al.* (2004), Cai *et al.* (1999), Zhang and Ma (1999), etc. For inverse continuous optimization, there are just a few papers, among which are Zhang and Liu (1996; 1999) and Jiang *et al.* (2011) for inverse linear programming; Zhang and Zhang (2010) for inverse quadratic programming. Especially, for inverse conic programming problems, Iyengar and Kang (2005) considered a general inverse conic programming and Xiao *et al.* (2009) discussed an inverse semidefinite constrained quadratic programming problem. However, the parameters required to approximate in both of the two inverse problems are in the objective functions. In this paper, our concern is a type of inverse linear semidefinite programming problem, in which the right-hand side parameter of linear semidefinite constraint is required to adjust, too.

Problem (2) can be reformulated to a famous optimization problem, called MPEC (i.e. mathematical programs with equilibrium constraints). Indeed, under a suitable constraint qualification such as the slater condition, we know that $x^0 \in \mathrm{SOL}(\mathrm{LSDP}(\mathrm{c}, \mathcal{A}, \mathrm{B}))$ if and only if there exists $\Omega \in S^n$ such that

$$c + \mathcal{A}^*(\Omega) = 0, \quad \mathcal{A}x^0 - B \in S^n_-, \quad \Omega \in S^n_+, \quad \langle \Omega, \mathcal{A}x^0 - B \rangle = 0.$$

Let $Z := B - \mathcal{A}x^0, U^0 := \mathcal{A}x^0 - B^0$. Then (2) is equivalent to

$$\begin{aligned} &(\mathrm{P}) \ \min_{\Omega, Z} \ f(\Omega, Z) := \frac{1}{2}\{\|\mathcal{A}^*(\Omega) + c^0\|^2 + \|Z + U^0\|_F^2\} \\ &\qquad \text{s.t.} \ \ \Omega \in S^n_+, Z \in S^n_+, \langle \Omega, Z \rangle = 0, \end{aligned} \tag{3}$$

which is a mathematical program with semidefinite cone complementarity constraints (SDCMPCC). SDCMPCC can be considered as a matrix analogue of the mathematical program with complementarity constraints(MPCC) since when the semidefinite cone complementarity constraint in (3) is replaced by the vector complementarity constraint, it becomes a MPCC. MPCC, or more generally, MPEC, is intensively studied and used in modern optimization theory and its numerous applications, such as engineering design and economic modeling. For SDCMPCC, there are only a few papers discussing on the optimality conditions, see Ding *et al.* (2014a), Wu *et al.* (2014) and Li and Qi (2011), and few solving methods can be found.

In this paper, we introduce a nonsmooth penalty function method to solve the SDCMPCC (3), in which the complementarity constraint is partially added as a nonsmooth penalty term in the objective function. Since the penalized problem is a nonconvex nonlinear semidefinite program, which is not easy to solve, we then propose a sequential convex program approach to solve the problem by noticing that the penalized problem is a DC programming problem. In Hong *et al.* (2011) and Xiao *et al.* (2012), the authors first proposed sequential convex program approaches to solve DC programming problems. The main idea of such approach is to linearize the second function of the delta convex or difference convex (DC) function in the model. Such idea is also used by Gao and Sun (2010) in the algorithm design of penalty function method for rank constrained optimization problem.

The organization of this paper is as follows. In Sec. 2, we give some notions and useful lemmas. Section 3 is devoted to construct a nonsmooth penalty function method for solving the reformulation of the inverse problem, i.e. SDCMPCC (3). In Sec. 4, a sequential convex program approach is introduced to solve the penalized problem. Some numerical results are shown in Sec. 5 and conclusions are given in Sec. 6.

Throughout this paper the following notations will be used. Let $\mathcal{O}^n$ be the set of all $n \times n$ orthogonal matrices and S^n_+ (S^n_-) be the cone of $n \times n$ positive(negative) semidefinite matrices in S^n(the space of $n \times n$ real symmetric matrices). For any $Z \in \Re^{m\times n}$, we denote by Z_{ij} the (i, j)th entry of Z, and for index sets $\alpha \subseteq \{1, \ldots, m\}$, $\beta \subseteq \{1, \ldots, n\}$, $Z_{\alpha\beta}$ is a submatrix of Z with entries Z_{ij} such that $i \in \alpha, j \in \beta$. For a matrix $X \in S^n$, $\Pi_{S^n_+}(X)$ denotes the metric projection of X to the positive semidefinite cone. And we use $\lambda(X)$ to denote the eigenvalue vector of X and $\Lambda(X)$ to denote $\operatorname{diag}\{\lambda(X)\}$, where we sometimes omit X in $\Lambda(X)$ if there is no confusion. For a diagonal matrix $\Lambda \in \Re^{n\times n}$, Λ_α is a submatrix of Λ with diagonal entries Λ_{ii} such that $i \in \alpha$. We denote by E the matrix whose entries are all ones, and by I the identity matrix, respectively, in the dimension considered.

2. Preliminaries

Throughout the paper, we extensively use the following notions and results.

For a single-valued Lipschitz continuous mapping $F : \Re^n \to \Re^m$, the B-subdifferential of F at x, denoted by $\partial_B F(x)$ is defined by

$$\partial_B F(x) := \left\{ \lim_{k\to\infty} \mathcal{J}F(x^k) \,\middle|\, x^k \to x, F \text{ is differentiable at } x^k \right\}.$$

The convex hull of $\partial_B F(x)$ is the Clarke's generalized Jacobian of F at x, denoted by $\partial F(x)$, (Clarke, 1983).

For any given $A \in S^n$, let A have the eigenvalue decomposition being arranged in nonincreasing order

$$A = P \begin{bmatrix} \Lambda_\alpha & & \\ & 0_\beta & \\ & & \Lambda_\gamma \end{bmatrix} P^T, \tag{4}$$

where $\alpha := \{i \,|\, \lambda_i(A) > 0\}$, $\beta := \{i \,|\, \lambda_i(A) = 0\}$, $\gamma := \{i \,|\, \lambda_i(A) < 0\}$. Denote all such matrices P in the eigenvalue decomposition (4) by $\mathcal{O}^n(A)$. Suppose $X \in S^n_+$,

$Y \in S^n_+$, $\langle X, Y\rangle = 0$. Let $A = X - Y$ have the eigenvalue decomposition (4), then we can easily get that

$$X = P \begin{bmatrix} \Lambda_\alpha & & \\ & 0_\beta & \\ & & 0_\gamma \end{bmatrix} P^T, \quad -Y = P \begin{bmatrix} 0_\alpha & & \\ & 0_\beta & \\ & & \Lambda_\gamma \end{bmatrix} P^T.$$

For A which admits the eigenvalue decomposition (4), define the matrix $\Sigma \in S^n$ with entries

$$\begin{cases} \Sigma_{ij} = \dfrac{\max\{\lambda_i(A), 0\} + \max\{\lambda_j(A), 0\}}{|\lambda_i(A)| + |\lambda_j(A)|} & \text{if } (i,j) \notin \beta \times \beta, \\ \Sigma_{ij} \in [0,1] & \text{if } (i,j) \in \beta \times \beta. \end{cases} \tag{5}$$

The following lemma on $\partial_B \Pi_{S^n_+}(A)$ and $\partial \Pi_{S^n_+}(A)$ is from Sun, (2006, Proposition 2.2) which is based on Pang *et al.* (2003, Lemma 11).

Lemma 2.1. *Suppose that $A \in S^n$ has the eigenvalue decomposition as in* (4). *Then $V \in \partial_B \Pi_{S^n_+}(A)$ (respectively, $\partial \Pi_{S^n_+}(A)$) if and only if there exists a $W \in \partial_B \Pi_{S^{|\beta|}_+}(0)$ (respectively, $\partial \Pi_{S^{|\beta|}_+}(0)$) such that*

$$V(H) = P \begin{bmatrix} \tilde{H}_{\alpha\alpha} & \tilde{H}_{\alpha\beta} & \Sigma_{\alpha\gamma} \circ \tilde{H}_{\alpha\gamma} \\ \tilde{H}_{\beta\alpha} & W(\tilde{H}_{\beta\beta}) & 0 \\ \Sigma_{\gamma\alpha} \circ \tilde{H}_{\gamma\alpha} & 0 & 0 \end{bmatrix} P^T, \quad \forall\, H \in S^n,$$

where $\tilde{H} = P^T H P$ and $\circ$ denotes the Hadamard product.

From Lemma 2.1 and Pang *et al.* (2003, Lemma 11) it can be easily verified that $V \in \partial \Pi_{S^n_+}(A)$ if and only if there exists Σ satisfying (5) such that

$$V(H) = P(\Sigma \circ (P^T H P))P^T, \quad \forall\, H \in S^n. \tag{6}$$

The following lemma can be obtained by Ding (2012, Proposition 2.5) which is useful in the subsequence.

Lemma 2.2. *Let $A \in S^n$ be given and have the eigenvalue decomposition* (4). *For any $H \in S^n$, let $U \in \mathcal{O}^n$ be an orthogonal matrix such that $A + H = U$ $\mathrm{diag}(\lambda(A + H))U^T$. Then for any $S^n \ni H \to 0$, we have*

$$U = PQ + \mathrm{O}(\|H\|),$$

where $Q = \begin{bmatrix} Q_\alpha & & \\ & Q_\beta & \\ & & Q_\gamma \end{bmatrix} \in \mathcal{O}^n$ *with* $Q_\alpha \in \mathcal{O}^{|\alpha|}$, $Q_\beta \in \mathcal{O}^{|\beta|}$, $Q_\gamma \in \mathcal{O}^{|\gamma|}$.

Since Problem (P) is a mathematical programming problem with semidefinite cone complementarity constraint, in what follows, we give the definitions of stationary points of Problem (P), which will be discussed in the subsequence, in the view of mathematical programs with equilibrium constraints, see Ding *et al.* (2014a) for details.

Definition 2.1. Let $(\overline{\Omega}, \overline{Z})$ be a feasible solution of Problem (P). Let $A = \overline{\Omega} - \overline{Z}$ have the eigenvalue decomposition (4).

(a) We say that $(\overline{\Omega}, \overline{Z})$ is a W-stationary point of (P) if there exists $(\Gamma^\Omega, \Gamma^Z) \in S^n \times S^n$ such that

$$\nabla f(\overline{\Omega}, \overline{Z}) + (\Gamma^\Omega, \Gamma^Z) = 0, \tag{7}$$

$$\tilde{\Gamma}^\Omega_{\alpha\alpha} = 0, \quad \tilde{\Gamma}^\Omega_{\alpha\beta} = 0, \quad \tilde{\Gamma}^\Omega_{\beta\alpha} = 0, \tag{8}$$

$$\tilde{\Gamma}^Z_{\gamma\gamma} = 0, \quad \tilde{\Gamma}^Z_{\gamma\beta} = 0, \quad \tilde{\Gamma}^Z_{\beta\gamma} = 0, \tag{9}$$

$$\Sigma_{\alpha\gamma} \circ \tilde{\Gamma}^\Omega_{\alpha\gamma} - (E_{\alpha\gamma} - \Sigma_{\alpha\gamma}) \circ \tilde{\Gamma}^Z_{\alpha\gamma} = 0, \tag{10}$$

where $\Sigma \in S^n$ is defined by (5) and $\tilde{\Gamma}^\Omega = P^T \Gamma^\Omega P$, $\tilde{\Gamma}^Z = P^T \Gamma^Z P$.

(b) We say that $(\overline{\Omega}, \overline{Z})$ is a C-stationary point of (P) if there exists $(\Gamma^\Omega, \Gamma^Z) \in S^n \times S^n$ such that (7)–(10) hold and

$$\langle \tilde{\Gamma}^\Omega_{\beta\beta}, \tilde{\Gamma}^Z_{\beta\beta} \rangle \geq 0. \tag{11}$$

(c) We say that $(\overline{\Omega}, \overline{Z})$ is an S-stationary point of (P) if there exists $(\Gamma^\Omega, \Gamma^Z) \in S^n \times S^n$ such that (7)–(10) hold and

$$\tilde{\Gamma}^\Omega_{\beta\beta} \in S^{|\beta|}_-, \quad \tilde{\Gamma}^Z_{\beta\beta} \in S^{|\beta|}_-. \tag{12}$$

From (7) in Definition 2.1, we know that the multiplier of any type of stationary point is unique. Also, we can easily get that SDCMPCC LICQ (Ding *et al.*, 2014a) holds at any type of stationary point defined in Definition 2.1.

Definition 2.2. A W-stationary point $(\overline{\Omega}, \overline{Z})$ of (P) is said to satisfy the upper level strict complementarity (SDCMPCC ULSC) condition if there exists $(\Gamma^G, \Gamma^H) \in S^n \times S^n$ satisfying (7)–(10) and

$$\tilde{\Gamma}^\Omega_{\beta\beta}, \tilde{\Gamma}^Z_{\beta\beta} \in S^{|\beta|}_{++} \cup S^{|\beta|}_{--},$$

where $S^{|\beta|}_{++} (S^{|\beta|}_{--})$ is the cone of $|\beta| \times |\beta|$ positive(negative) definite matrices in $S^{|\beta|}$.

3. The Nonsmooth Penalty Function Method for Mathematical Programs with Semidefinite Cone Complementarity Constraints

In this section, we introduce a penalty method to solve the mathematical programs with semidefinite cone complementarity constraints, i.e. Problem (P).

Let $\rho > 0$ be a non-negative parameter. Consider the following nonsmooth penalty formulation of Problem (P):

$$\begin{aligned} (\mathrm{P}_\rho)\ \min_{\Omega, Z}\ & f(\Omega, Z) + \rho\, \Phi(\Omega, Z) \\ \text{s.t. }\ & \Omega \in S^n_+, Z \in S^n_+, \end{aligned} \tag{13}$$

where $\Phi(\Omega, Z) = \langle I, \Omega - \Pi_{S^n_+}(\Omega - Z)\rangle$. Define the feasible sets of Problem (P) and Problem (P_ρ) by $\mathcal{F}$ and $\mathcal{F}_0$, respectively, i.e.

$$\mathcal{F} := \{(\Omega, Z) \,|\, \Omega \in S^n_+, Z \in S^n_+, \langle \Omega, Z\rangle = 0\}, \quad \mathcal{F}_0 := \{(\Omega, Z) \,|\, \Omega \in S^n_+, Z \in S^n_+\}.$$

The next lemma gives the properties of the penalty function Φ.

Lemma 3.1. *Let $\Phi(\Omega, Z) = \langle I, \Omega - \Pi_{S^n_+}(\Omega - Z)\rangle$. For any $(\Omega, Z) \in \mathcal{F}_0$, the following statements hold*:

(a) $\Phi(\Omega, Z) \geq 0$;
(b) $\Phi(\Omega, Z) = 0$ *if and only if* $(\Omega, Z) \in \mathcal{F}$.

Proof. (a) Suppose that the eigenvalues of $\Omega - Z$, Ω, $-Z$ are arranged in nonincreasing orders and $A = \Omega - Z$ has the eigenvalue decomposition (4). Then

$$\Phi(\Omega, Z) = \mathrm{tr}(\Omega) - \mathrm{tr}(\Pi_{S^n_+}(\Omega - Z)) = \sum_{i=1}^{n} \lambda_i(\Omega) - \sum_{i\in\alpha} \lambda_i(\Omega - Z).$$

It follows from Horn and Johnson (1985, Lemma 7.4.50) that for any $(\Omega, Z) \in \mathcal{F}_0$,

$$\sum_{i\in\alpha} [\lambda_i(\Omega) - \lambda_i(\Omega - Z)] \geq \sum_{i\in\alpha} -\lambda_i(-Z) \geq 0,$$

which implies $\Phi(\Omega, Z) \geq \sum_{i\in\beta\cup\gamma} \lambda_i(\Omega) \geq 0$.

(b) For any $(\Omega, Z) \in \mathcal{F}$, suppose that $A = \Omega - Z$ has the eigenvalue decomposition (4). Then

$$\Omega = P \begin{bmatrix} \Lambda_\alpha & & \\ & 0_\beta & \\ & & 0_\gamma \end{bmatrix} P^T, \quad Z = P \begin{bmatrix} 0_\alpha & & \\ & 0_\beta & \\ & & -\Lambda_\gamma \end{bmatrix} P^T.$$

So we have $\Phi(\Omega, Z) = \mathrm{tr}(\Lambda_\alpha) - \mathrm{tr}(\Lambda_\alpha) = 0$.

Conversely, if $\Phi(\Omega, Z) = 0$ and $(\Omega, Z) \in \mathcal{F}_0$, we need to prove that $\langle \Omega, Z\rangle = 0$. Suppose that the eigenvalues of $\Omega - Z$, Ω, $-Z$ are arranged in nonincreasing orders and $A = \Omega - Z$ has the eigenvalue decomposition (4). Let $\Lambda(Z) = -\Lambda(-Z)$ and $\Lambda(Z - \Omega) = -\Lambda(\Omega - Z)$, then the eigenvalues of Z and $Z - \Omega$ are arranged in nondecreasing orders. Since $\Phi(\Omega, Z) = 0$, it follows from the proof of (a) that

$$\sum_{i\in\alpha} \lambda_i(\Omega) = \sum_{i\in\alpha} \lambda_i(\Omega - Z) \quad \text{and} \quad \sum_{i\in\beta\cup\gamma} \lambda_i(\Omega) = 0.$$

Since for every i, $\lambda_i(\Omega) \geq \lambda_i(\Omega - Z)$ and $\lambda_i(\Omega) \geq 0$, we have $\lambda_i(\Omega) = \lambda_i(\Omega - Z)$, $\forall\, i \in \alpha$ and $\lambda_i(\Omega) = 0$, $\forall\, i \in \beta \cup \gamma$. So we obtain that

$$\Lambda(\Omega) = \begin{bmatrix} \Lambda_\alpha & & \\ & 0 & \\ & & 0 \end{bmatrix}.$$

Similarly, since $\Omega - \Pi_{S^n_+}(\Omega - Z) = Z - \Pi_{S^n_+}(Z - \Omega)$, we have that

$$\begin{aligned}\Phi(\Omega, Z) = \mathrm{tr}(Z) - \mathrm{tr}(\Pi_{S^n_+}(Z - \Omega)) &= \sum_{i=1}^{n} \lambda_i(Z) - \sum_{i\in\gamma} \lambda_i(Z - \Omega)\\ &= \sum_{i\in\gamma}[\lambda_i(Z) - \lambda_i(Z - \Omega)] + \sum_{i\in\alpha\cup\beta} \lambda_i(Z)\\ &\geq \sum_{i\in\gamma} \lambda_i(\Omega) + \sum_{i\in\alpha\cup\beta} \lambda_i(Z) \geq 0.\end{aligned}$$

Then $\Phi(\Omega, Z) = 0$ implies $\lambda_i(Z) = \lambda_i(Z - \Omega)$, $\forall\, i \in \gamma$ and $\lambda_i(Z) = 0$, $\forall\, i \in \alpha \cup \beta$. So we obtain that

$$\Lambda(-Z) = \begin{bmatrix} 0 & & \\ & 0 & \\ & & \Lambda_\gamma \end{bmatrix}.$$

It can be easily verified that

$$\begin{aligned}\|\Omega - Z\|_F^2 &= \sum_{i=1}^{n} \lambda_i^2(\Omega - Z) = \sum_{i\in\alpha} \lambda_i^2(\Omega) + \sum_{i\in\gamma} \lambda_i^2(-Z) = \langle \Lambda_\alpha, \Lambda_\alpha\rangle + \langle \Lambda_\gamma, \Lambda_\gamma\rangle\\ &= \|\Omega\|_F^2 + \|-Z\|_F^2.\end{aligned}$$

Then we have $\langle \Omega, Z\rangle = 0$, which together with $(\Omega, Z) \in \mathcal{F}_0$ implies $(\Omega, Z) \in \mathcal{F}$. $\square$

Now we discuss the differential of the nonsmooth part of the penalty function Φ. Define $\Phi_2(\Omega, Z) := \langle I, \Pi_{S^n_+}(\Omega - Z)\rangle$, $\forall\, \Omega \in S^n, Z \in S^n$. The following lemma gives the subdifferential of Φ_2 at any point $(\Omega^*, Z^*) \in \mathcal{F}_0$.

Lemma 3.2. *For any $(\Omega^*, Z^*) \in \mathcal{F}_0$, suppose that $A = \Omega^* - Z^*$ has the eigenvalue decomposition (4), $P = [p_1\ p_2\ \ldots\ p_n]$. Define the vector $\theta \in \Re^n$ with entries*

$$\theta_i \begin{cases} = 1 & \text{if } i \in \alpha,\\ \in [0, 1] & \text{if } i \in \beta,\\ = 0 & \text{if } i \in \gamma. \end{cases} \tag{14}$$

Then

$$\partial\Phi_2(\Omega^*, Z^*) = \left\{(V, -V) \in S^n \times S^n \mid V = \sum_{i=1}^{n} \theta_i p_i p_i^T, \theta \in \Re^n \text{ satisfies } (14)\right\}. \tag{15}$$

Proof. It follows from Sun *et al.* (2008, Lemma 2) that for any $(H_1, H_2) \in S^n \times S^n$,

$$\partial_B\Phi_2(\Omega^*, Z^*)(H_1, H_2) = \langle I, \partial_B\Pi_{S^n_+}(\Omega^* - Z^*)(H_1 - H_2)\rangle.$$

It follows from Meng *et al.* (2005, Proposition 1(i)) and (6) that

$$\begin{aligned}
&\partial\Phi_2(\Omega^*, Z^*)(H_1, H_2)\\
&\quad = \{\langle P(\Theta \circ I)P^T, H_1 - H_2\rangle \mid \Theta \text{ satisfies (5) with } \Sigma = \Theta\}\\
&\quad = \left\{\left\langle \sum_{i=1}^{n} \Theta_{ii} p_i p_i^T, H_1 - H_2 \right\rangle \middle| \Theta \text{ satisfies (5) with } \Sigma = \Theta\right\}\\
&\quad = \left\{\left\langle \left(\sum_{i=1}^{n} \Theta_{ii} p_i p_i^T, -\sum_{i=1}^{n} \Theta_{ii} p_i p_i^T\right), (H_1, H_2) \right\rangle \middle| \Theta \text{ satisfies (5) with } \Sigma = \Theta\right\}\\
&\quad = \left\{\left\langle \left(\sum_{i=1}^{n} \theta_i p_i p_i^T, -\sum_{i=1}^{n} \theta_i p_i p_i^T\right), (H_1, H_2) \right\rangle \middle| \theta \in \Re^n \text{ satisfies (14)}\right\}.
\end{aligned}$$

So we obtain that

$$\partial\Phi_2(\Omega^*, Z^*) = \left\{\left(\sum_{i=1}^{n} \theta_i p_i p_i^T, -\sum_{i=1}^{n} \theta_i p_i p_i^T\right) \middle| \theta \in \Re^n \text{ satisfies (14)}\right\}.$$

The proof is completed. □

Based on the nonsmooth optimization theory and Lemma 3.2, we get the following definition of stationary point of (P_ρ).

Definition 3.1. A feasible point $(\overline{\Omega}, \overline{Z})$ of (P_ρ) is called a stationary point, if

$$0 \in \nabla f(\overline{\Omega}, \overline{Z}) + \rho(I, 0) - \rho\partial\Phi_2(\overline{\Omega}, \overline{Z}) + \mathcal{N}_{S_+^n}(\overline{\Omega}) \times \mathcal{N}_{S_+^n}(\overline{Z}),$$

where $\partial\Phi_2(\overline{\Omega}, \overline{Z})$ has the expression (15).

Since in the next section, we will construct a numerical method for solving the penalty model, i.e. Problem (P_ρ), other than Problem (P). In the remain of the section, we will discuss the relationship between the stationary points of (P_ρ) and (P).

First, we discuss the optimal properties between (P_ρ) and (P) for fixed ρ which is large enough if needed.

Theorem 3.1. *Let $\rho > 0$ be given. If the global optimal solution $(\overline{\Omega}, \overline{Z})$ of (P_ρ) satisfies $\langle\overline{\Omega}, \overline{Z}\rangle = 0$, then $(\overline{\Omega}, \overline{Z})$ is a global optimal solution of (P).*

Proof. Suppose that $(\overline{\Omega}, \overline{Z})$ is a global solution of (P_ρ) such that $\langle\overline{\Omega}, \overline{Z}\rangle = 0$, then $(\overline{\Omega}, \overline{Z})$ is a feasible solution to (P) and $\Phi(\overline{\Omega}, \overline{Z}) = 0$ from Lemma 3.1. Also, for any feasible solution (Ω, Z) of (P), one has $\Phi(\Omega, Z) = 0$ from Lemma 3.1, which implies that

$$f(\overline{\Omega}, \overline{Z}) = f(\overline{\Omega}, \overline{Z}) + \rho\Phi(\overline{\Omega}, \overline{Z}) \le f(\Omega, Z) + \rho\Phi(\Omega, Z) = f(\Omega, Z).$$

Thus, $(\overline{\Omega}, \overline{Z})$ is a global solution of (P). □

The following theorem tells us that an ε-optimal solution to (P) in the sense of (17) can be guaranteed provided that the penalty parameter ρ is larger than some ε-dependent number. We omit the proof of the theorem as similar results can be found in Gao and Sun (2010) and Wang *et al.* (2014). Define (P_0) by

$$\begin{aligned} (\mathrm{P}_0)\ \min_{\Omega,Z}\ & f(\Omega,Z) := \frac{1}{2}\left\{\|\mathcal{A}^*(\Omega)+c^0\|^2 + \|Z+U^0\|_F^2\right\} \\ \text{s.t. }\ & \Omega \in S_+^n, Z \in S_+^n, \end{aligned} \tag{16}$$

Theorem 3.2. *Let $\varepsilon > 0$ be given and (Ω, Z) be a feasible solution to* (P)*. Assume that $\rho > 0$ is chosen such that $(f(\Omega,Z) - f(\Omega_0,Z_0))/\rho \le \varepsilon$, where (Ω_0, Z_0) is an optimal solution to* (P_0)*. Then we have*

$$0 \le \Phi(\Omega_\rho, Z_\rho) \le \varepsilon, \quad f(\Omega_\rho, Z_\rho) \le \bar{f} - \rho\Phi(\Omega_\rho, Z_\rho) \le \bar{f}, \tag{17}$$

where (Ω_ρ, Z_ρ) is the global solution to (P_ρ)*, $\bar{f}$ is the global optimal value of* (P).

Theorem 3.3. *For some $\rho > 0$, if $(\overline{\Omega}, \overline{Z})$ is a stationary point of* (P_ρ) *and feasible to Problem* (P)*, then $(\overline{\Omega}, \overline{Z})$ is a W-stationary point of* (P)*. Moreover, if $(\overline{\Omega}, \overline{Z})$ satisfies*

$$0 \in \nabla f(\overline{\Omega}, \overline{Z}) + \rho(I,0) - \rho\,\mathrm{rint}\,\partial\Phi_2(\overline{\Omega}, \overline{Z}) + \mathcal{N}_{S_+^n}(\overline{\Omega}) \times \mathcal{N}_{S_+^n}(\overline{Z}), \tag{18}$$

(*i.e. we replace $\partial\Phi_2(\overline{\Omega}, \overline{Z})$ by its relative interior* $\mathrm{rint}\,\partial\Phi_2(\overline{\Omega}, \overline{Z})$ *in the stationary condition of Problem* (P_ρ)) *when ρ is large enough, then $(\overline{\Omega}, \overline{Z})$ satisfies the SDCM-PCC ULSC defined in Definition* 2.2 *and is a C-stationary point of* (P).

Proof. If $(\overline{\Omega}, \overline{Z})$ is a stationary point of (P_ρ), then there exist $(H^\Omega, H^Z) \in \mathcal{N}_{S_+^n}(\overline{\Omega}) \times \mathcal{N}_{S_+^n}(\overline{Z})$ and $(M^\Omega, M^Z) \in \partial\Phi_2(\overline{\Omega}, \overline{Z})$ such that

$$\nabla f(\overline{\Omega}, \overline{Z}) + \rho(I,0) - \rho(M^\Omega, M^Z) + (H^\Omega, H^Z) = 0. \tag{19}$$

Suppose that $A = \overline{\Omega} - \overline{Z}$ have the eigenvalue decomposition (4). Then it follows from the proof of Lemma 2.2 that there exists Σ satisfying (5) such that $M^\Omega = P(\Sigma \circ I)P^T$, $M^Z = -M^\Omega$. From the definition of Σ we know that $\Sigma_{\alpha\alpha} = E_{\alpha\alpha}$ and $\Sigma_{\gamma\gamma} = 0$.

Let $\Gamma^\Omega = \rho[I - P(\Sigma \circ I)P^T] + H^\Omega$, $\Gamma^Z = \rho P(\Sigma \circ Is)P^T + H^Z$. If we can verify that

$$\tilde{\Gamma}^\Omega_{\alpha\alpha} = 0, \quad \tilde{\Gamma}^\Omega_{\alpha\beta} = 0, \quad \tilde{\Gamma}^Z_{\gamma\gamma} = 0, \quad \tilde{\Gamma}^Z_{\gamma\beta} = 0, \quad \tilde{\Gamma}^\Omega_{\alpha\gamma} = \tilde{\Gamma}^Z_{\alpha\gamma} = 0, \tag{20}$$

where $\tilde{\Gamma}^\Omega = P^T\Gamma^\Omega P$, $\tilde{\Gamma}^Z = P^T\Gamma^Z P$, then $(\overline{\Omega}, \overline{Z})$ is a W-stationary point. Now we prove (20).

Since $H^\Omega \in \mathcal{N}_{S_+^n}(\overline{\Omega})$, $H^Z \in \mathcal{N}_{S_+^n}(\overline{Z})$, we know that H^Ω and $\overline{\Omega}$(H^Z and $\overline{Z}$) admit a simultaneous ordered eigenvalue decomposition, i.e. there exist two orthogonal matrices $P_1, P_2 \in \mathcal{O}^n$ such that

$$\overline{\Omega} = P_1 \begin{bmatrix} \Lambda_\alpha & 0 \\ 0 & 0_{\beta\cup\gamma} \end{bmatrix} P_1^T, \quad H^\Omega = P_1 \begin{bmatrix} 0_\alpha & 0 \\ 0 & \Lambda(H^\Omega)_{\beta\cup\gamma} \end{bmatrix} P_1^T$$

and

$$-\overline{Z} = P_2 \begin{bmatrix} 0_{\alpha\cup\beta} & 0 \\ 0 & \Lambda_\gamma \end{bmatrix} P_2^T, \quad H^Z = P_2 \begin{bmatrix} \Lambda(H^Z)_{\alpha\cup\beta} & 0 \\ 0 & 0_\gamma \end{bmatrix} P_2^T,$$

where $\Lambda(H^\Omega)_{\beta\cup\gamma} \in S_-^{|\beta\cup\gamma|}$, $\Lambda(H^Z)_{\alpha\cup\beta} \in S_-^{|\alpha\cup\beta|}$.

Therefore, it is easy to check that there exist two orthogonal matrices $S, T \in \mathcal{O}^n$ such that

$$P = P_1 S, \quad P = P_2 T,$$

with

$$S = \begin{bmatrix} S_\alpha & 0 \\ 0 & S_{\beta\cup\gamma} \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} T_{\alpha\cup\beta} & 0 \\ 0 & T_\gamma \end{bmatrix},$$

where $S_\alpha \in \mathcal{O}^{|\alpha|}$, $S_{\beta\cup\gamma} \in \mathcal{O}^{|\beta\cup\gamma|}$ and $T_{\alpha\cup\beta} \in \mathcal{O}^{|\alpha\cup\beta|}$, $T_\gamma \in \mathcal{O}^{|\gamma|}$. Let

$$S_{\beta\cup\gamma} = [S_1 \ S_2], \quad T_{\alpha\cup\beta} = [T_1 \ T_2]$$

with $S_1 \in \Re^{|\beta\cup\gamma|\times|\beta|}$, $S_2 \in \Re^{|\beta\cup\gamma|\times|\gamma|}$, $T_1 \in \Re^{|\alpha\cup\beta|\times|\alpha|}$ and $T_2 \in \Re^{|\alpha\cup\beta|\times|\beta|}$. Then

$$\begin{aligned}
\tilde{\Gamma}^\Omega &= P^T(\rho[I - P(\Sigma \circ I)P^T] + H^\Omega)P = S^T P_1^T H^\Omega P_1 S + \rho(E - \Sigma) \circ I \\
&= \begin{bmatrix} S_\alpha^T & 0 \\ 0 & S_{\beta\cup\gamma}^T \end{bmatrix} \begin{bmatrix} 0_\alpha & 0 \\ 0 & \Lambda(H^\Omega)_{\beta\cup\gamma} \end{bmatrix} \begin{bmatrix} S_\alpha & 0 \\ 0 & S_{\beta\cup\gamma} \end{bmatrix} + \rho \begin{bmatrix} 0 & 0 & 0 \\ 0 & (E - \Sigma_{\beta\beta}) \circ I & 0 \\ 0 & 0 & I \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 \\ 0 & S_1^T \Lambda(H^\Omega)_{\beta\cup\gamma} S_1 + \rho(E - \Sigma_{\beta\beta}) \circ I & S_1^T \Lambda(H^\Omega)_{\beta\cup\gamma} S_2 \\ 0 & S_2^T \Lambda(H^\Omega)_{\beta\cup\gamma} S_1 & S_2^T \Lambda(H^\Omega)_{\beta\cup\gamma} S_2 + \rho I \end{bmatrix},
\end{aligned}$$

from which we obtain that

$$\tilde{\Gamma}^\Omega_{\alpha\alpha} = 0, \tilde{\Gamma}^\Omega_{\alpha\beta} = 0, \tilde{\Gamma}^\Omega_{\alpha\gamma} = 0, \tilde{\Gamma}^\Omega_{\beta\beta} = S_1^T \Lambda(H^\Omega)_{\beta\cup\gamma} S_1 + \rho(E_{\beta\beta} - \Sigma_{\beta\beta}) \circ I. \quad (21)$$

Similarly, we can get that

$$\tilde{\Gamma}^Z_{\gamma\gamma} = 0, \tilde{\Gamma}^Z_{\gamma\beta} = 0, \tilde{\Gamma}^Z_{\alpha\gamma} = 0, \tilde{\Gamma}^Z_{\beta\beta} = T_2^T \Lambda(H^Z)_{\alpha\cup\beta} T_2 + \rho\Sigma_{\beta\beta} \circ I. \quad (22)$$

Therefore, $(\overline{\Omega}, \overline{Z})$ is a W-stationary point of (P).

If (18) holds, then in (19), $(M^\Omega, M^Z) \in \operatorname{rint} \partial\Phi_2(\overline{\Omega}, \overline{Z})$, which implies that $\Sigma_{ii} \in (0, 1), \forall\, i \in \beta$. It follows from (21) and (22) that $\tilde{\Gamma}^\Omega_{\beta\beta}$ and $\tilde{\Gamma}^Z_{\beta\beta}$ are positive definite matrices when ρ is large enough. So we obtain that $(\overline{\Omega}, \overline{Z})$ satisfies the SDCMPCC ULSC and is a C-stationary point of (P). $\square$

To further discuss the relationship between the stationary points of (P_ρ) and (P), we give a partial converse of Theorem 3.3 showing that S-stationary points of (P) are stationary points of (P_ρ) under some assumptions. This together with second order information will helps us study the exact penalty properties of our nonsmooth penalty function in the future.

Theorem 3.4. *Suppose that* $(\overline{\Omega}, \overline{Z})$ *is an S-stationary point of* (P) *with multiplier* $(\Gamma^\Omega, \Gamma^Z)$. *If there exists* $P \in \mathcal{O}^n(\overline{\Omega} - \overline{Z})$ *such that* $(P^T\Gamma^G P)_{\alpha\gamma} = 0$ *or* $(P^T\Gamma^H P)_{\alpha\gamma} = 0$, *then for all* ρ *sufficiently large,* $(\overline{\Omega}, \overline{Z})$ *is a stationary point of* (P_ρ).

Proof. Suppose that $A = \overline{\Omega} - \overline{Z}$ have the eigenvalue decomposition (4), where P satisfies $(P^T\Gamma^G P)_{\alpha\gamma} = 0$ or $(P^T\Gamma^H P)_{\alpha\gamma} = 0$. Let $\tilde{\Gamma}^\Omega = P^T\Gamma^\Omega P$, $\tilde{\Gamma}^Z = P^T\Gamma^Z P$. We know from the assumptions and the definition of S-stationary point that

$$\tilde{\Gamma}^\Omega_{\alpha\alpha} = 0,\ \tilde{\Gamma}^\Omega_{\alpha\beta} = 0, \tilde{\Gamma}^Z_{\gamma\gamma} = 0,\ \tilde{\Gamma}^Z_{\gamma\beta} = 0, \tilde{\Gamma}^\Omega_{\alpha\gamma} = \tilde{\Gamma}^Z_{\alpha\gamma} = 0, \tilde{\Gamma}^\Omega_{\beta\beta} \in S^{|\beta|}_-, \tilde{\Gamma}^Z_{\beta\beta} \in S^{|\beta|}_-,$$

and

$$\nabla f(\overline{\Omega}, \overline{Z}) + (\Gamma^\Omega, \Gamma^Z) = 0.$$

Let $H^\Omega = \Gamma^\Omega - \rho[I - P(\Sigma \circ I)P^T]$ and $H^Z = \Gamma^Z - \rho P(\Sigma \circ I)P^T$, where $\rho > 0$, Σ satisfies (5). Then

$$\nabla f(\overline{\Omega}, \overline{Z}) + \rho(I, 0) - (\rho P(\Sigma \circ I)P^T, -\rho P(\Sigma \circ I)P^T) + (H^\Omega, H^Z) = 0$$

and

$$\langle H^\Omega, \overline{\Omega}\rangle = 0, \langle H^Z, \overline{Z}\rangle = 0.$$

To prove that $(\overline{\Omega}, \overline{Z})$ is a stationary point of (P_ρ), we only need to prove that $H^\Omega \in S^n_-$ and $H^Z \in S^n_-$ after choosing proper $\rho > 0$. Since

$$P^T H^\Omega P = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \tilde{\Gamma}^\Omega_{\beta\beta} - \rho(E_{\beta\beta} - \Sigma_{\beta\beta}) \circ I & \tilde{\Gamma}^\Omega_{\beta\gamma} \\ 0 & \tilde{\Gamma}^\Omega_{\gamma\beta} & \tilde{\Gamma}^\Omega_{\gamma\gamma} - \rho I \end{bmatrix},$$

$$P^T H^Z P = \begin{bmatrix} \tilde{\Gamma}^Z_{\alpha\alpha} - \rho I & \tilde{\Gamma}^Z_{\alpha\beta} & 0 \\ \tilde{\Gamma}^Z_{\beta\alpha} & \tilde{\Gamma}^Z_{\beta\beta} - \rho\Sigma_{\beta\beta} \circ I & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where we choose $\Sigma_{\beta\beta} = \frac{1}{2}E_{\beta\beta}$, a sufficient condition for H^Ω and H^Z to be negative semidefinite is that the Schur-complement of $\tilde{\Gamma}^\Omega_{\beta\beta} - \frac{\rho}{2}I$ in the matrix

$$\begin{bmatrix} \tilde{\Gamma}^\Omega_{\beta\beta} - \frac{\rho}{2}I & \tilde{\Gamma}^\Omega_{\beta\gamma} \\ \tilde{\Gamma}^\Omega_{\gamma\beta} & \tilde{\Gamma}^\Omega_{\gamma\gamma} - \rho I \end{bmatrix}$$

is negative semidefinite, and the Schur-complement of $\tilde{\Gamma}^Z_{\beta\beta} - \frac{\rho}{2}I$ in the matrix

$$\begin{bmatrix} \tilde{\Gamma}^Z_{\alpha\alpha} - \rho I & \tilde{\Gamma}^Z_{\alpha\beta} \\ \tilde{\Gamma}^Z_{\beta\alpha} & \tilde{\Gamma}^Z_{\beta\beta} - \frac{\rho}{2}I \end{bmatrix}$$

is negative semidefinite when $\rho > 2\max\{\lambda_{\max}(\tilde{\Gamma}^\Omega_{\beta\beta}), \lambda_{\max}(\tilde{\Gamma}^Z_{\beta\beta})\}$.

When $\rho > 4\max\{\lambda_{\max}(\tilde{\Gamma}^{\Omega}_{\beta\beta}), \lambda_{\max}(\tilde{\Gamma}^{Z}_{\beta\beta})\}$, we have $\tilde{\Gamma}^{\Omega}_{\beta\beta} - \frac{\rho}{2}I + \frac{\rho}{4}I \in S^{|\beta|}_{-}$ and $\tilde{\Gamma}^{Z}_{\beta\beta} - \frac{\rho}{2}I + \frac{\rho}{4}I \in S^{|\beta|}_{-}$. Then $[\tilde{\Gamma}^{\Omega}_{\beta\beta} - \frac{\rho}{2}I]^{-1} + \frac{4}{\rho}I \in S^{|\beta|}_{+}$ and $[\tilde{\Gamma}^{\Omega}_{\beta\beta} - \frac{\rho}{2}I]^{-1} + \frac{4}{\rho}I \in S^{|\beta|}_{+}$. If we can verify when ρ is large enough that

$$\tilde{\Gamma}^{\Omega}_{\gamma\gamma} - \rho I + \frac{4}{\rho}\tilde{\Gamma}^{\Omega}_{\gamma\beta}\tilde{\Gamma}^{\Omega}_{\beta\gamma} \in S^{|\gamma|}_{-}, \quad \tilde{\Gamma}^{Z}_{\alpha\alpha} - \rho I + \frac{4}{\rho}\tilde{\Gamma}^{Z}_{\alpha\beta}\tilde{\Gamma}^{Z}_{\beta\alpha} \in S^{|\alpha|}_{-}, \tag{23}$$

then both of the Schur-complements mentioned above are negative semidefinite. Equation (23) holds obviously when $\rho > \max\{4, \lambda_{\max}(\tilde{\Gamma}^{\Omega}_{\gamma\gamma} + \tilde{\Gamma}^{\Omega}_{\gamma\beta}\tilde{\Gamma}^{\Omega}_{\beta\gamma}), \lambda_{\max}(\tilde{\Gamma}^{Z}_{\alpha\alpha} + \tilde{\Gamma}^{Z}_{\alpha\beta}\tilde{\Gamma}^{Z}_{\beta\alpha})\}$. So when

$$\begin{aligned} \rho > \bar{\rho} := \max\{&4, 4\lambda_{\max}(\tilde{\Gamma}^{\Omega}_{\beta\beta}), 4\lambda_{\max}(\tilde{\Gamma}^{Z}_{\beta\beta}), \\ &\lambda_{\max}(\tilde{\Gamma}^{\Omega}_{\gamma\gamma} + \tilde{\Gamma}^{\Omega}_{\gamma\beta}\tilde{\Gamma}^{\Omega}_{\beta\gamma}), \lambda_{\max}(\tilde{\Gamma}^{Z}_{\alpha\alpha} + \tilde{\Gamma}^{Z}_{\alpha\beta}\tilde{\Gamma}^{Z}_{\beta\alpha})\}, \end{aligned}$$

we have $H^{\Omega} \in S^n_{-}$ and $H^Z \in S^n_{-}$, which implies that for any $\rho > \bar{\rho}$, $(\overline{\Omega}, \overline{Z})$ is a stationary point of (P_ρ). □

The next theorem gives the asymptotic properties of $\rho \to +\infty$ about the stationary points.

Theorem 3.5. *Let* (Ω^k, Z^k) *be a stationary point of* (P_ρ) *for each* $\rho = \rho_k$, *where* $\rho_k \uparrow \infty$. *Suppose that there exists a subsequence* $\{(\Omega^{k_i}, Z^{k_i})\} \to (\overline{\Omega}, \overline{Z})$ *with* $(\overline{\Omega}, \overline{Z})$ *feasible to* (P) *satisfying*

$$\rho^{k_i}\|\Omega^{k_i} - \overline{\Omega}\| \to 0, \quad \rho^{k_i}\|Z^{k_i} - \overline{Z}\| \to 0, \tag{24}$$

then $(\overline{\Omega}, \overline{Z})$ *is a W-stationary point. Moreover, if there is a nonempty set* $B \in \operatorname{rint}\partial\Phi_2(\Omega^{k_i}, Z^{k_i})$, $\forall\, i$, *satisfying*

$$0 \in \nabla f(\Omega^{k_i}, Z^{k_i}) + \rho^{k_i}(I, 0) - \rho^{k_i}B + \mathcal{N}_{S^n_+}(\Omega^{k_i}) \times \mathcal{N}_{S^n_+}(Z^{k_i}). \tag{25}$$

Then $(\overline{\Omega}, \overline{Z})$ *satisfies the SDCMPCC ULSC defined in Definition* 2.2 *and is a C-stationary point of* (P).

Proof. Without loss of sgenerality, we assume that $(\Omega^k, Z^k) \to (\overline{\Omega}, \overline{Z})$. Suppose that $A = \overline{\Omega} - \overline{Z}$ has the eigenvalue decomposition (4), $A^k = \Omega^k - Z^k$ has the eigenvalue decomposition being arranged in nonincreasing order $A^k = P_k\Lambda(A^k)P_k^T$, and $\Sigma^k \in S^n$ is defined with entries

$$\begin{cases} \Sigma^k_{ij} \in [0, 1] & \text{if } \lambda_i(A^k) = \lambda_j(A^k) = 0, \\ \Sigma^k_{ij} = \dfrac{\max\{\lambda_i(A^k), 0\} + \max\{\lambda_j(A^k), 0\}}{|\lambda_i(A^k)| + |\lambda_j(A^k)|} & \text{otherwise.} \end{cases} \tag{26}$$

Then it follows from the definition of Σ^k that $\Sigma^k_{\alpha\alpha} = E_{\alpha\alpha}$ and $\Sigma^k_{\gamma\gamma} = 0$ when k is large enough.

Since (Ω^k, Z^k) is a stationary point of (P_{ρ^k}), there exist Σ^k satisfying (26) and (25), and Lagrange multipliers $H_k^\Omega \in \mathcal{N}_{S_+^n}(\Omega^k)$, $H_k^Z \in \mathcal{N}_{S_+^n}(Z^k)$, such that

$$\nabla f(\Omega^k, Z^k) + \rho_k(I,0) - (\rho_k P_k(\Sigma^k \circ I)P_k^T, -\rho_k P_k(\Sigma^k \circ I)P_k^T) + (H_k^\Omega, H_k^Z) = 0. \tag{27}$$

Let $\Gamma_k^\Omega = \rho_k[I - P_k(\Sigma^k \circ I)P_k^T] + H_k^\Omega$, $\Gamma_k^Z = \rho_k P_k(\Sigma^k \circ I)P_k^T + H_k^Z$. Then

$$\nabla f(\Omega^k, Z^k) + (\Gamma_k^\Omega, \Gamma_k^Z) = 0.$$

It follows from the continuous differentiability of f that the sequence $\{(\Gamma_k^\Omega, \Gamma_k^Z)\}$ is convergent. Assume that $\Gamma_k^\Omega \to \Gamma^\Omega$ and $\Gamma_k^Z \to \Gamma^Z$. Firstly, we prove that $(\overline{\Omega}, \overline{Z})$ is a W-stationary point. We only need to prove that Γ^Ω and Γ^Z satisfy (8)–(10). Since $(\Omega^k, Z^k) \to (\overline{\Omega}, \overline{Z}) \in \mathcal{F}$, $H_k^\Omega \in \mathcal{N}_{S_+^n}(\Omega^k)$, it follows from Lemma 2.2 that there exist $Q_k^1 \in \mathcal{O}^n$, $W = \begin{bmatrix} W_\alpha & \\ & W_{\beta\cup\gamma} \end{bmatrix} \in \mathcal{O}^n$ with $W_\alpha \in \mathcal{O}^{|\alpha|}$, $W_{\beta\cup\gamma} \in \mathcal{O}^{|\beta\cup\gamma|}$ and $Q_k^1 W \to P$ such that Ω^k has the following ordered eigenvalue decomposition

$$\Omega^k = Q_k^1 \begin{bmatrix} \Lambda_\alpha^k & & \\ & \Lambda_\beta^k & \\ & & \Lambda_\gamma^k \end{bmatrix} (Q_k^1)^T,$$

where $S_{++}^{|\alpha|} \ni \Lambda_\alpha^k \to \Lambda_\alpha$, $S_+^{|\beta|} \ni \Lambda_\beta^k \to 0_\beta$, $S_+^{|\gamma|} \ni \Lambda_\gamma^k \to 0_\gamma$ when k is large enough, and

$$H_k^\Omega = Q_k^1 \begin{bmatrix} 0_\alpha & \\ & \Lambda_{\beta\cup\gamma}(H_k^\Omega) \end{bmatrix} (Q_k^1)^T, \quad S_-^{|\beta\cup\gamma|} \ni \Lambda_{\beta\cup\gamma}(H_k^\Omega) \perp \begin{bmatrix} \Lambda_\beta^k & \\ & \Lambda_\gamma^k \end{bmatrix} \in S_+^{|\beta\cup\gamma|}.$$

Since $\Omega^k - Z^k \to \overline{\Omega} - \overline{Z}$, there exist $L = \begin{bmatrix} L_\alpha & & \\ & L_\beta & \\ & & L_\gamma \end{bmatrix} \in \mathcal{O}^n$ with $L_\alpha \in \mathcal{O}^{|\alpha|}$, $L_\beta \in \mathcal{O}^{|\beta|}$, $L_\gamma \in \mathcal{O}^{|\gamma|}$, such that $P_k L \to P$. It follows from Lemma 2.2 that

$$\begin{aligned} \|Q_k^1 W - P_k L\| &\le \|Q_k^1 W - P\| + \|P_k L - P\| \\ &= \mathrm{O}(\|\Omega^k - \overline{\Omega}\|) + \mathrm{O}(\|\Omega^k - Z^k - (\overline{\Omega} - \overline{Z})\|) \\ &= \mathrm{O}(\|\Omega^k - \overline{\Omega}\|) + \mathrm{O}(\|Z^k - \overline{Z}\|), \end{aligned}$$

which together with (24) implies

$$\lim_{k\to\infty} \rho_k L^T (P_k)^T H_k (Q_k^1 W - P_k L) + \rho_k (Q_k^1 W - P_k L)^T H_k Q_k^1 W = 0,$$

where $H_k = I - P_k(\Sigma^k \circ I)P_k^T$ is bounded. Then

$$\begin{aligned} \lim_{k\to\infty} \tilde{\Gamma}_k^\Omega &:= \lim_{k\to\infty} P^T \Gamma_k^\Omega P \\ &= \lim_{k\to\infty} W^T (Q_k^1)^T (\rho_k[I - P_k(\Sigma^k \circ I)P_k^T] + H_k^\Omega) Q_k^1 W \\ &= \lim_{k\to\infty} W^T (Q_k^1)^T H_k^\Omega Q_k^1 W + \rho_k L^T [I - (\Sigma^k \circ I)] L \end{aligned}$$

$$
\begin{aligned}
&+\rho_k L^T(P_k)^T H_k(Q_k^1 W - P_k L) + \rho_k (Q_k^1 W - P_k L)^T H_k Q_k^1 W \\
&= \lim_{k\to\infty} \begin{bmatrix} W_\alpha^T & \\ & W_{\beta\cup\gamma}^T \end{bmatrix} \begin{bmatrix} 0_\alpha & \\ & \Lambda_{\beta\cup\gamma}(H_k^\Omega) \end{bmatrix} \begin{bmatrix} W_\alpha & \\ & W_{\beta\cup\gamma} \end{bmatrix} \\
&\quad + \rho_k \begin{bmatrix} 0_\alpha & & \\ & L_\beta^T[(E_{\beta\beta} - \Sigma_{\beta\beta}^k)\circ I]L_\beta & \\ & & I_\gamma \end{bmatrix} \\
&= \lim_{k\to\infty} \begin{bmatrix} 0_\alpha & \\ & W_{\beta\cup\gamma}^T \Lambda_{\beta\cup\gamma}(H_k^\Omega) W_{\beta\cup\gamma} \end{bmatrix} \\
&\quad + \rho_k \begin{bmatrix} 0_\alpha & & \\ & L_\beta^T[(E_{\beta\beta} - \Sigma_{\beta\beta}^k)\circ I]L_\beta & \\ & & I_\gamma \end{bmatrix}.
\end{aligned}
$$

Similarly, there exist $Q_k^2 \in \mathcal{O}^n$, $T = \begin{bmatrix} T_{\alpha\cup\beta} & \\ & T_\gamma \end{bmatrix} \in \mathcal{O}^n$ with $T_{\alpha\cup\beta} \in \mathcal{O}^{|\alpha\cup\beta|}$, $T_\gamma \in \mathcal{O}^{|\gamma|}$ and $Q_k^2 T \to P$ such that

$$
\begin{aligned}
\lim_{k\to\infty} \tilde{\Gamma}_k^Z := \lim_{k\to\infty} P^T \Gamma_k^Z P = \lim_{k\to\infty} &\begin{bmatrix} T_{\alpha\cup\beta}^T \Lambda_{\alpha\cup\beta}(H_k^Z) T_{\alpha\cup\beta} & \\ & 0_\gamma \end{bmatrix} \\
&+ \rho_k \begin{bmatrix} I_\alpha & & \\ & L_\beta^T[\Sigma_{\beta\beta}^k \circ I]L_\beta & \\ & & 0_\gamma \end{bmatrix},
\end{aligned}
$$

where $\Lambda_{\alpha\cup\beta}(H_k^Z) \in S_-^{|\alpha\cup\beta|}$.

Thus it can be easily verified that Γ^Ω and Γ^Z satisfy (8)–(10) and $(\overline{\Omega}, \overline{Z})$ is a W-stationary point of (P). Moreover, if (25) holds, from the expression of $\partial\Phi_2$ in (15), we have that for every $i \in \beta$, there exists a nonempty set $B_i \subseteq (0,1)$ such that $\lim_{k\to\infty} \Sigma_{ii}^k \in B_i$(one may take a subsequence if necessary), which implies that $\tilde{\Gamma}_{\beta\beta}^\Omega$ and $\tilde{\Gamma}_{\beta\beta}^Z$ are positive definite matrices when ρ is large enough. So we obtain that $(\overline{\Omega}, \overline{Z})$ satisfies the SDCMPCC ULSC defined in Definition 2.2 and is a C-stationary point of (P). □

4. The Sequential Convex Program Approach for the Penalty Model

Since f is a convex function and $\Phi(\Omega, Z) = \langle I, \Omega\rangle - \langle I, \Pi_{S_+^n}(\Omega - Z)\rangle$, which is the difference of two convex functions, i.e. Φ is a nonsmooth DC function, Problem (P_ρ) can be considered as a DC programming problem

$$
\begin{aligned}
(\mathrm{DC}_\rho)\ \min_{\Omega, Z}\ & f_\rho(\Omega, Z) := \Phi_1(\Omega, Z) - \rho\Phi_2(\Omega, Z) \\
\text{s.t. }\ & (\Omega, Z) \in \mathcal{F}_0,
\end{aligned} \tag{28}
$$

where $\Phi_1(\Omega, Z) = \frac{1}{2}\{\|\mathcal{A}^*(\Omega)+c^0\|^2+\|Z+U^0\|_F^2\}+\rho\langle I, \Omega\rangle$, $\Phi_2(\Omega, Z) = \langle I, \Pi_{S_+^n}(\Omega - Z)\rangle$, $\mathcal{F}_0 = \{(\Omega, Z) | \, \Omega \in S_+^n, Z \in S_+^n\}$. It can be easily seen that Φ_1 is a twice differentiable mapping and is a strict convex function if $\mathcal{A}$ is onto, Φ_2 is a semismooth convex mapping (Chang and Chen, 2013; Auslender, 2003). In this section, we introduce a sequential convex program approach to solve this nonsmooth DC programming problem.

Definition 4.1. $(\Omega^*, Z^*) \in \mathcal{F}_0$ is called a stationary point of Problem (DC_ρ), if

$$0 \in \nabla\Phi_1(\Omega^*, Z^*) - \rho\partial\Phi_2(\Omega^*, Z^*) + \mathcal{N}_{S_+^n}(\Omega^*) \times \mathcal{N}_{S_+^n}(Z^*). \tag{29}$$

For any $\overline{\Omega} \in S^n, \overline{Z} \in S^n$, let $A = \overline{\Omega} - \overline{Z}$ have the eigenvalue decomposition (4), $P = [p_1 \ p_2 \ \ldots \ p_n]$. Define $W_{\overline{\Omega},\overline{Z}} \in S^n \times S^n$ by

$$W_{\overline{\Omega},\overline{Z}} := (V_{\overline{\Omega},\overline{Z}}, -V_{\overline{\Omega},\overline{Z}}), \tag{30}$$

where $V_{\overline{\Omega},\overline{Z}} = \sum_{i\in\alpha} p_i p_i^T + \frac{1}{2}\sum_{i\in\beta} p_i p_i^T$. Then by Lemma 3.2, $W_{\overline{\Omega},\overline{Z}} \in \operatorname{rint} \partial\Phi_2(\overline{\Omega}, \overline{Z})$. We will see in (33) that the subdifferential is chosen in the way of (30) at each iteration. The reason is to ensure that the accumulate point (Ω^l, Z^l) of the sequence generated by Algorithm 1 corresponding to penalty parameter ρ^l forms a sequence $\{(\Omega^l, Z^l)\}$ which always has subsequence satisfying (25) in Theorem 3.5.

From the definition of subdifferential for the convex mapping, we obtain that

$$\Phi_2(\Omega, Z) \geq \Phi_2(\overline{\Omega}, \overline{Z}) + \langle V_{\overline{\Omega},\overline{Z}}, \Omega - \overline{\Omega}\rangle - \langle V_{\overline{\Omega},\overline{Z}}, Z - \overline{Z}\rangle, \quad \forall\, \Omega, Z \in S^n. \tag{31}$$

Define the following convex problem:

$$\begin{aligned} CP(\overline{\Omega}, \overline{Z}, V_{\overline{\Omega},\overline{Z}}) \ \min_{D_\Omega, D_Z} \ & \Phi_1(\overline{\Omega} + D_\Omega, \overline{Z} + D_Z) - \rho[\Phi_2(\overline{\Omega}, \overline{Z}) + \langle V_{\overline{\Omega},\overline{Z}}, D_\Omega - D_Z\rangle] \\ \text{s.t.} \ & (\overline{\Omega} + D_\Omega, \overline{Z} + D_Z) \in \mathcal{F}_0. \end{aligned} \tag{32}$$

Define $S_k(\Omega, Z) := \nabla\Phi_1(\Omega, Z) - \rho(V_{\Omega^k,Z^k}, -V_{\Omega^k,Z^k}) + \mathcal{N}_{\mathcal{F}_0}(\Omega, Z)$. Then (D_Ω^*, D_Z^*) is an optimal solution of $CP(\Omega^k, Z^k, V_{\Omega^k,Z^k})$ if and only if $0 \in S_k(\Omega^k + D_\Omega^*, Z^k + D_Z^*)$.

The sequential convex program approach for (DC_ρ) is based on solving a sequel of convex problems formulated as the above.

Algorithm 1 (Sequential convex program approach).

Step 0. Choose $\Omega^0 \in S_+^n$, $Z^0 \in S_+^n$, $\delta, \sigma \in (0, 1)$. Choose $\varepsilon_k > 0$. Set $k := 0$.
Step 1. Let $A = \Omega^k - Z^k$ have the eigenvalue decomposition (4). Let

$$V_{\Omega^k,Z^k} = \sum_{i\in\alpha} p_i p_i^T + \frac{1}{2}\sum_{i\in\beta} p_i p_i^T. \tag{33}$$

Step 2. Find an inexact optimal solution (D_Ω^k, D_Z^k) of Problem $CP(\Omega^k, Z^k, V_{\Omega^k,Z^k})$ such that

$$\operatorname{dist}(0, S_k(\Omega^k + D_\Omega^k, Z^k + D_Z^k)) \leq \varepsilon_k \|(D_\Omega^k, D_Z^k)\|. \tag{34}$$

Step 3. (Armijo line search) Let l_k be the smallest non-negative integer l satisfying

$$f_\rho((\Omega^k, Z^k) + \delta^l(D^k_\Omega, D^k_Z)) \le f_\rho(\Omega^k, Z^k) + \sigma\delta^l(\langle \nabla\Phi_1(\Omega^k, Z^k), (D^k_\Omega, D^k_Z)\rangle - \rho\langle V_{\Omega^k,Z^k}, D^k_\Omega - D^k_Z\rangle). \tag{35}$$

Define $\alpha_k = \delta^{l_k}, \Omega^{k+1} = \Omega^k + \alpha_k D^k_\Omega, Z^{k+1} = Z^k + \alpha_k D^k_Z$.

Step 4. If $\Omega^{k+1} = \Omega^k, Z^{k+1} = Z^k$, stop. Otherwise, replace k by $k+1$, go to Step 1.

The following theorem tells us that l_k in Armijo line search always exists.

Theorem 4.1. *Suppose that* $\{(\Omega^k, Z^k)\}$, $\{(D^k_\Omega, D^k_Z)\}$ *are generated by Algorithm* 1, *for every* k, V_{Ω^k,Z^k} *is defined by* (33) *and* $\mathcal{A}$ *is onto. Then there exists* $\bar\varepsilon > 0$ *such that when* $\varepsilon_k \le \bar\varepsilon$, *we have that*

$$f_\rho((\Omega^k, Z^k) + \alpha(D^k_\Omega, D^k_Z)) \le f_\rho(\Omega^k, Z^k) + \sigma\alpha[\langle \nabla\Phi_1(\Omega^k, Z^k), (D^k_\Omega, D^k_Z)\rangle - \rho\langle V_{\Omega^k,Z^k}, D^k_\Omega - D^k_Z\rangle],$$

for all $\alpha \in [0, 1-\sigma]$.

Proof. Define $X^k := (\Omega^k, Z^k)$, $D^k := (D^k_\Omega, D^k_Z)$ and $W^k := (V_{\Omega^k,Z^k}, -V_{\Omega^k,Z^k})$. It follows from (34) that there exists $H^k \in S_k(\Omega^k + D^k_\Omega, Z^k + D^k_Z)$ such that $\|H^k\| \le \varepsilon_k\|D^k\|$, which implies

$$H^k \in \nabla\Phi_1(X^k + D^k) - \rho W^k + \mathcal{N}_{\mathcal{F}_0}(X^k + D^k).$$

Then there exists $M^k \in \mathcal{N}_{\mathcal{F}_0}(X^k + D^k)$, from which we can obtain $\langle M^k, -D^k\rangle \le 0$, such that $H^k = \nabla\Phi_1(X^k + D^k) - \rho W^k + M^k$. Since $\mathcal{A}$ is onto, we can prove by contradiction that there exists $a > 0$ such that for any $H \in S^n$, $\|\mathcal{A}^*(H)\| \ge a\|H\|$. Since for any $H = (H_1, H_2) \in S^n \times S^n$,

$$\langle (H_1, H_2), \nabla^2\Phi_1(X^k)(H_1, H_2)\rangle = \|\mathcal{A}^*(H_1)\|^2 + \|H_2\|^2 \ge a^2\|H_1\|^2 + \|H_2\|^2 \ge \min(a^2, 1)\|H\|^2,$$

we have that

$$\langle D^k, \nabla^2\Phi_1(X^k)D^k\rangle \ge 2\bar\varepsilon\|D^k\|^2, \quad \forall\, k = 1, 2, \ldots, \tag{36}$$

where $\bar\varepsilon = \frac{1}{2}\min(a^2, 1)$. Actually,

$$\bar\varepsilon = \frac{1}{2}\min\left(\inf_{H\neq 0}\frac{\|\mathcal{A}^*(H)\|^2}{\|H\|^2}, 1\right). \tag{37}$$

Because Φ_1 is quadratic, we have that

$$\begin{aligned}
&\langle \nabla\Phi_1(X^k), D^k\rangle - \rho\langle W^k, D^k\rangle \\
&\quad = \langle \nabla\Phi_1(X^k + D^k), D^k\rangle - \langle D^k, \nabla^2\Phi_1(X^k)D^k\rangle - \rho\langle W^k, D^k\rangle \\
&\quad = \langle H^k - M^k, D^k\rangle - \langle D^k, \nabla^2\Phi_1(X^k)D^k\rangle
\end{aligned}$$

$$\leq \langle H^k, D^k\rangle - \langle D^k, \nabla^2\Phi_1(X^k)D^k\rangle$$
$$\leq \varepsilon_k\|D^k\|^2 - \langle D^k, \nabla^2\Phi_1(X^k)D^k\rangle, \tag{38}$$

which implies

$$\langle D^k, \nabla^2\Phi_1(X^k)D^k\rangle \leq \varepsilon_k\|D^k\|^2 - [\langle\nabla\Phi_1(X^k), D^k\rangle - \rho\langle W^k, D^k\rangle].$$

This together with (36) implies that when $\varepsilon_k \leq \bar{\varepsilon}$,

$$\varepsilon_k\|D^k\|^2 \leq -[\langle\nabla\Phi_1(X^k), D^k\rangle - \rho\langle W^k, D^k\rangle].$$

Again from (38), we obtain that

$$\langle D^k, \nabla^2\Phi_1(X^k)D^k\rangle \leq -2[\langle\nabla\Phi_1(X^k), D^k\rangle - \rho\langle W^k, D^k\rangle].$$

So we have that

$$\begin{aligned} f_\rho(X^k + \alpha D^k) - f_\rho(X^k) &= \Phi_1(X^k + \alpha D^k) - \rho\Phi_2(X^k + \alpha D^k) \\ &\quad - \Phi_1(X^k) + \rho\Phi_2(X^k) \\ &\leq \Phi_1(X^k + \alpha D^k) - \Phi_1(X^k) - \alpha\rho\langle W^k, D^k\rangle \\ &= \alpha\langle\nabla\Phi_1(X^k), D^k\rangle + \frac{\alpha^2}{2}\langle D^k, \nabla^2\Phi_1(X^k)D^k\rangle \\ &\quad - \alpha\rho\langle W^k, D^k\rangle \\ &= \alpha[\langle\nabla\Phi_1(X^k), D^k\rangle - \rho\langle W^k, D^k\rangle] \\ &\quad + \frac{\alpha^2}{2}\langle D^k, \nabla^2\Phi_1(X^k)D^k\rangle \\ &\leq \alpha(1-\alpha)[\langle\nabla\Phi_1(X^k), D^k\rangle - \rho\langle W^k, D^k\rangle]. \end{aligned}$$

Then for any $\alpha \in [0, 1-\sigma]$, $f_\rho(X^k + \alpha D^k) - f_\rho(X^k) \leq \alpha\sigma[\langle\nabla\Phi_1(X^k), D^k\rangle - \rho\langle W^k, D^k\rangle]$. We complete the proof. □

Now we give the main result of this section, i.e. the convergence property of Algorithm 1.

Theorem 4.2. *Suppose that $\{(\Omega^k, Z^k)\}$, $\{D_\Omega^k, D_Z^k\}$ are generated by Algorithm 1, $\mathcal{A}$ is onto and $\varepsilon_k \leq \bar{\varepsilon}$ with $\bar{\varepsilon}$ being defined by (37). Then $\{f_\rho(\Omega^k, Z^k)\}$ is a nonincreasing sequence and satisfies*

$$\begin{aligned} f_\rho(\Omega^{k+1}, Z^{k+1}) - f_\rho(\Omega^k, Z^k) &\leq \alpha_k\varepsilon_k\|(D_\Omega^k, D_Z^k)\|^2 - \frac{1}{2}\|\mathcal{A}^*(\Omega^{k+1} - \Omega^k)\|^2 \\ &\quad - \frac{1}{2}\|Z^{k+1} - Z^k\|^2. \end{aligned} \tag{39}$$

If $(\Omega^{k+1}, Z^{k+1}) = (\Omega^k, Z^k)$ for some integer $k \geq 0$, then (Ω^k, Z^k) is a stationary point of Problem (DC_ρ). Otherwise, any accumulation point of the infinite sequence

$\{(\Omega^k, Z^k)\}$ is a stationary point of Problem (DC_ρ), provided that Problem (DC_ρ) has a finite optimal value.

Proof. From the line search rule (35), we obtain that $\{f_\rho(\Omega^k, Z^k)\}$ is a non-increasing sequence. Define $X := (\Omega, Z)$, $X^k := (\Omega^k, Z^k)$, $D^k := (D_\Omega^k, D_Z^k)$ and $W^k := (V_{\Omega^k,Z^k}, -V_{\Omega^k,Z^k})$. Define $f^k(X) := \Phi_1(X) - \rho[\Phi_2(X^k) + \rho\langle W^k, X - X^k\rangle]$, $\forall\, X \in S^n \times S^n$, $k = 1, 2, \ldots$. Then for each k, f^k is a convex quadratic function, and $f^k(X^k) = f_\rho(X^k)$, $\nabla f^k(X) = \nabla\Phi_1(X) - \rho W^k$, $\nabla^2 f^k(X) = \nabla^2\Phi_1(X)$,

$$\begin{aligned} f^k(X^k) &= f^k(X^k + \alpha_k D^k) - \alpha_k\langle\nabla f^k(X^k + \alpha_k D^k), D^k\rangle \\ &\quad + \frac{\alpha_k^2}{2}\langle D^k, \nabla^2\Phi_1(X^k + \alpha_k D^k)D^k\rangle. \end{aligned} \tag{40}$$

It follows from (34) that there exists $H^k \in S_k(\Omega^k + D_\Omega^k, Z^k + D_Z^k)$ such that $\|H^k\| \le \varepsilon_k\|D^k\|$, which implies

$$H^k \in \nabla f^k(X^k + D^k) + \mathcal{N}_{\mathcal{F}_0}(X^k + D^k),$$

which, by the definition of normal cone, implies

$$\langle\nabla f^k(X^k + D^k) - H^k, D^k\rangle \le 0. \tag{41}$$

Combing (40) and (41), noticing the relationship (31), we obtain that

$$\begin{aligned} f_\rho(X^k + \alpha_k D^k) - f_\rho(X^k) &= \Phi_1(X^k + \alpha_k D^k) - \rho\Phi_2(X^k + \alpha_k D^k) - f^k(X^k) \\ &\le \Phi_1(X^k + \alpha_k D^k) - \rho\Phi_2(X^k) - \alpha_k\rho\langle W^k, D^k\rangle - f^k(X^k) \\ &= f^k(X^k + \alpha_k D^k) - f^k(X^k) \\ &= \alpha_k\langle\nabla f^k(X^k + \alpha_k D^k), D^k\rangle \\ &\quad - \frac{\alpha_k^2}{2}\langle D^k, \nabla^2\Phi_1(X^k + \alpha_k D^k)D^k\rangle \\ &\le \alpha_k\langle H^k, D^k\rangle - \frac{\alpha_k^2}{2}\langle D^k, \nabla^2\Phi_1(X^k + \alpha_k D^k)D^k\rangle \\ &\le \alpha_k\varepsilon_k\|D^k\|^2 - \frac{\alpha_k^2}{2}\langle D^k, \nabla^2\Phi_1(X^k + \alpha_k D^k)D^k\rangle. \end{aligned}$$

So

$$\begin{aligned} &f_\rho(\Omega^{k+1}, Z^{k+1}) - f_\rho(\Omega^k, Z^k) \\ &\quad\le \alpha_k\varepsilon_k\|(D_\Omega^k, D_Z^k)\|^2 - \frac{1}{2}\langle(X^{k+1} - X^k), \nabla^2\Phi_1(X^{k+1})(X^{k+1} - X^k)\rangle \\ &\quad= \alpha_k\varepsilon_k\|(D_\Omega^k, D_Z^k)\|^2 - \frac{1}{2}\|\mathcal{A}^*(\Omega^{k+1} - \Omega^k)\|^2 - \frac{1}{2}\|Z^{k+1} - Z^k\|^2. \end{aligned}$$

We next consider the case $(\Omega^{k+1}, Z^{k+1}) = (\Omega^k, Z^k)$ for some integer $k \ge 0$. In this case, $D^k = 0$ and $0 \in S_k(\Omega^k + D_\Omega^k, Z^k + D_Z^k)$ for some integer $k \ge 0$. For such

k one has

$$0 \in \nabla\Phi_1(\Omega^k, Z^k) - \rho(V_{\Omega^k,Z^k}, -V_{\Omega^k,Z^k}) + \mathcal{N}_{\mathcal{F}_0}(\Omega^k, Z^k),$$

which implies

$$0 \in \nabla\Phi_1(\Omega^k, Z^k) - \rho\partial\Phi_2(\Omega^k, Z^k) + \mathcal{N}_{S^n_+}(\Omega^k) \times \mathcal{N}_{S^n_+}(Z^k).$$

So (Ω^k, Z^k) is a stationary point of Problem (DC_ρ).

Now assume that $(\Omega^{k+1}, Z^{k+1}) \neq (\Omega^k, Z^k)$ for any integer $k \geq 0$. Assume that $(\Omega^{k_i}, Z^{k_i}) \to (\Omega^*, Z^*)$. Since $\{f_\rho(\Omega^k, Z^k)\}$ is nonincreasing and is bounded from below (this is because Problem (DC_ρ) has a finite optimal value), we have $\lim_{k\to\infty} f_\rho(\Omega^k, Z^k) = f_\rho(\Omega^*, Z^*)$. It follows from (35) and the proof of Theorem 4.1 that

$$\begin{aligned} f_\rho(X^{k+1}) - f_\rho(X^k) &\leq \sigma\alpha_k[\langle\nabla\Phi_1(X^k), D^k\rangle - \rho\langle W^k, D^k\rangle] \leq 0, \\ f_\rho(X^k + \alpha D^k) - f_\rho(X^k) &\leq \Phi_1(X^k + \alpha D^k) - \Phi_1(X^k) - \alpha\rho\langle W^k, D^k\rangle, \quad \forall\, \alpha, \end{aligned} \tag{42}$$

and

$$\langle\nabla\Phi_1(X^k), D^k\rangle - \rho\langle W^k, D^k\rangle \leq -\frac{1}{2}\langle D^k, \nabla^2\Phi_1(X^k)D^k\rangle \leq -\bar{\varepsilon}\|D^k\|^2. \tag{43}$$

So we obtain that $\lim_{k\to\infty} \alpha_k\|D^k\| = 0$. Now we prove that $\lim_{i\to\infty} D^{k_i} = 0$. If this is not true, without loss of generality, we assume that there exists $a > 0$ such that $\|D^{k_i}\| \geq a$, $\forall\, i$. Then $\lim_{i\to\infty} \alpha_{k_i} = 0$. Again, from the line search rule, we have

$$f_\rho(X^k + (\alpha_k/\delta)D^k) - f_\rho(X^k) > \sigma(\alpha_k/\delta)[\langle\nabla\Phi_1(X^k), D^k\rangle - \rho\langle W^k, D^k\rangle],$$

from which and (42) we obtain that

$$\sigma[\langle\nabla\Phi_1(X^k), D^k\rangle - \rho\langle W^k, D^k\rangle] < \frac{\Phi_1(X^k + (\alpha_k/\delta)D^k) - \Phi_1(X^k)}{\alpha_k/\delta} - \rho\langle W^k, D^k\rangle.$$

Then

$$\begin{aligned} (1-\sigma)\bar{\varepsilon}\|D^k\|^2 &\leq (\sigma - 1)[\langle\nabla\Phi_1(X^k), D^k\rangle - \rho\langle W^k, D^k\rangle] \\ &< \frac{\Phi_1(X^k + (\alpha_k/\delta)D^k) - \Phi_1(X^k)}{\alpha_k/\delta} - \langle\nabla\Phi_1(X^k), D^k\rangle. \end{aligned}$$

Consequently,

$$\begin{aligned} (1-\sigma)\bar{\varepsilon}a \leq (1-\sigma)\bar{\varepsilon}\|D^{k_i}\| &< \frac{\Phi_1(X^{k_i} + (\alpha_{k_i}\|D^{k_i}\|/\delta)\frac{D^{k_i}}{\|D^{k_i}\|}) - \Phi_1(X^{k_i})}{\alpha_{k_i}\|D^{k_i}\|/\delta} \\ &\quad - \left\langle \nabla\Phi_1(X^{k_i}), \frac{D^{k_i}}{\|D^{k_i}\|} \right\rangle. \end{aligned}$$

When $i \to \infty$, without loss of generality, let $\frac{D^{k_i}}{\|D^{k_i}\|} \to \bar{D}$, $\|\bar{D}\| = 1$. Let $i \to \infty$, we obtain that

$$(1-\sigma)\bar{\varepsilon}a \leq \langle \nabla\Phi_1(\Omega^*, Z^*), \bar{D}\rangle - \langle \nabla\Phi_1(\Omega^*, Z^*), \bar{D}\rangle = 0.$$

This is a contradiction, which implies $\lim_{i\to\infty} D^{k_i} = 0$.

It follows from (34) that there exists H^{k_i} such that $\|H^{k_i}\| \leq \varepsilon_{k_i}\|D^{k_i}\|$ and

$$H^{k_i} \in \nabla\Phi_1(X^{k_i} + D^{k_i}) - \rho W^{k_i} + \mathcal{N}_{\mathcal{F}_0}(X^{k_i} + D^{k_i}),$$

where $W^{k_i} = (V_{\Omega^{k_i}, Z^{k_i}}, -V_{\Omega^{k_i}, Z^{k_i}}) \in \partial\Phi_2(\Omega^{k_i}, Z^{k_i})$. It follows from the outer semi-continuity of $\partial\Phi_2$ that $W^{k_i} \to W^* = (V_{\Omega^*, Z^*}, -V_{\Omega^*, Z^*}) \in \partial\Phi_2(\Omega^*, Z^*)$. Again, from the outer semi-continuity of $\mathcal{N}_{\mathcal{F}_0}$, we know that

$$0 \in \nabla\Phi_1(\Omega^*, Z^*) - \rho\partial\Phi_2(\Omega^*, Z^*) + \mathcal{N}_{\mathcal{F}_0}(\Omega^*, Z^*)$$

This means that (Ω^*, Z^*) is a stationary point of Problem (DC_ρ). □

At the end of this section, we discuss some practical issues in the implementation of the sequential convex program approach toghether with penalty method for solving Problem (P). The main framework of the algorithm is described as follows, whose convergence follows the classic results of the penalty methods (Luenberger and Ye, 2008, Chap. 13.1) together with Theorem 4.2.

Algorithm 2 (Penalty method).

Step 0. Given $\varepsilon > 0$, $\rho^1 > 0$ and a sequence $\{\mu^i\}$ with $\mu^i > 1$, for $i = 1, 2, \ldots$. Let $(\Omega^0, Z^0) = (0, 0)$. Set $k := 1$.
Step 1. Apply Algorithm 1 with (Ω^{k-1}, Z^{k-1}) as the starting point to solve (DC_{ρ^k}) to find (Ω^k, Z^k), i.e.

$$(\Omega^k, Z^k) = \text{argmin}_{(\Omega,Z)\in\mathcal{F}_0} f_{\rho^k}(\Omega, Z).$$

Step 2. If $\Phi(\Omega^k, Z^k) \leq \varepsilon$, stop; otherwise, go to Step 3.
Step 3. Set $\rho^{k+1} = \mu^k\rho^k$ and $k := k+1$, go to Step 1.

Now we discuss some main details in Algorithms 1 and 2.

(i) The main difficulty of Algorithm 1 is that it needs to solve a series of convex subproblems. Fortunately, the solution of Problem $CP(\Omega^k, Z^k, V_{\Omega^k, Z^k})$ has the following separable expression:

$$D_\Omega^k = \text{argmin}_X \delta_{S_+^n}(X) + \frac{1}{2}\|\mathcal{A}^*(X) + c^0\|^2 + \langle \rho(\mathcal{I} - V_{\Omega^k, Z^k}), X\rangle - \Omega^k,$$

$$D_Z^k = \Pi_{S_+^n}(-U^0 - \rho V_{\Omega^k, Z^k}) - Z^k,$$

where the notation $\delta_{S_+^n}(\cdot)$ is the indicator function of S_+^n and the minimization problem can be solved by the accelerated proximal gradient algorithm (Jiang *et al.*, 2012). Let $P(X) = \delta_{S^n_+}(X)$ and $f(X) = \frac{1}{2}\|\mathcal{A}^*(X) + c^0\|^2 + \langle M, X\rangle$. Consider the minimization problem

$$\min_X F(X) = f(X) + P(X), \tag{44}$$

the accelerated proximal gradient algorithm for solving (44) can be described as follows.

Algorithm 3 (Accelerated proximal gradient algorithm). Choose $X^0 = X^{-1} = 0 \in S^n$, $t^0 = t^{-1} = 1$. For $k = 0, 1, 2, \ldots$, generate X^{k+1} from X^k according the following steps.

Step 1. Set $Y^k = X^k + \frac{t^{k-1}-1}{t^k}(X^k - X^{k-1})$.
Step 2. Set $G^k = Y^k - (\tau^k)^{-1}\nabla f(Y^k)$, where $\tau^k > 0$.
Step 3. Set $X^{k+1} = \Pi_{S^n_+}(G^k)$.
Step 4. Compute $t^{k+1} = \frac{1+\sqrt{1+4(t^k)^2}}{2}$.

In practice, we always choose $\tau^k = L_f$ which is the Lipschitz constant of ∇f and similar to Toh and Yun (2010), the stopping condition is

$$\frac{\|S^{k+1}\|}{\tau^k \max\{1, \|X^{k+1}\|\}} < 10^{-4},$$

where $S^{k+1} := \tau^k(Y^k - X^{k+1}) + \mathcal{A}\mathcal{A}^*(X^{k+1} - Y^k)$.

(ii) In Algorithm 2, the initial penalty parameter ρ^1 is set to be

$$\rho^1 := \min\left\{1, \frac{0.25|f(\Omega^0, Z^0) - f(\Omega^*, Z^*)|}{\max\{1, \Phi(\Omega^*, Z^*)\}}\right\},$$

where (Ω^*, Z^*) is the optimal solution of the convex problem (P_0) which can be calculated by

$$\Omega^* = \mathrm{argmin}_{\Omega}\delta_{S^n_+}(\Omega) + \frac{1}{2}\|\mathcal{A}^*(\Omega) + c^0\|^2,$$

$$Z^* = \Pi_{S^n_+}(-U^0).$$

The sequence $\{\mu^k\}$ is defined by $\mu^k := \begin{cases} 4 & \text{if } \Phi(\Omega^k, Z^k) \geq 5 \\ 2 & \text{if } 0.1 < \Phi(\Omega^k, Z^k) < 5 \\ 1.4 & \text{otherwise.} \end{cases}$

(iii) In Algorithm 2, the stopping criterion is chosen by

$$|\Phi(\Omega^k, Z^k)| \leq 10^{-5} \quad \text{and} \quad \frac{\left|\sqrt{f_{\rho^k}(\Omega^k, Z^k)} - \sqrt{f_{\rho^k}(\Omega^{k-1}, Z^{k-1})}\right|}{\max\left\{1, \sqrt{f_{\rho^k}(\Omega^k, Z^k)}\right\}} \leq 10^{-5}.$$

5. Numerical Results

In this section, we report some numerical results of the sequential convex program approach for solving Problem (P). Our numerical experiments are carried out on a Laptop of Intel Core i5-3210M CPU 2.5 GHz with 4 GB RAM memory, running Windows 7 and MATLAB R2014a. We test on some randomly generated problems described as follows.

Example 5.1. Consider the linear semidefinite programming problem LSDP $\times$ $(\mathrm{c}, \mathcal{A}, \mathrm{B})$ whose parameters are randomly generated: the linear operator $\mathcal{A}$ is defined by $\mathrm{svec}(\mathcal{A}x) := Ax$ for any $x \in \Re^m$, where the random matrix $A \in \Re^{na\times m}$ with $na = 0.5 * n * (n+1)$ is:

$$\mathrm{A} = 2 * \mathrm{rand}(\mathrm{na}, \mathrm{m}) - \mathrm{ones}(\mathrm{na}, \mathrm{m}).$$

$B \in S^n$ is randomly generated and by randomly generating a matrix $M \in S^n_+$, we define $c := \mathcal{A}^*(M)$. Let x^0 be the solution of $\mathrm{LSDP}(\mathrm{c}, \mathcal{A}, \mathrm{B})$ and (c^0, B^0) be the perturbation of (c, B):

$$\mathrm{c}^0 = \mathrm{c} + 1.0\mathrm{e} - 1 * \mathrm{rand}(\mathrm{m}, 1), \quad \mathrm{B}^0 = \mathrm{B} + 1.0\mathrm{e} - 1 * \mathrm{rB},$$

where $rB \in S^n$ is randomly generated.

Suppose that $(\hat{\Omega}, \hat{Z})$ is the final iterative value of our main algorithm. Then $\hat{c} = -\mathcal{A}^*(\hat{\Omega}), \hat{B} = \hat{Z} + \mathcal{A}x^0$ should be the solution of (2). To verify the efficiency of our approach, we report the following results of problems with different dimensions in Table 1, where $\hat{x}$ is the solution of $\mathrm{LSDP}(\hat{\mathrm{c}}, \mathcal{A}, \hat{\mathrm{B}})$.

- time: the CPU time required by the algorithm.
- iter: the number of iterations in the penalty framework.
- infeas: The infeasibility of Problem(P) at $(\hat{\Omega}, \hat{Z})$,

$$\mathrm{infeas} = \Phi(\hat{\Omega}, \hat{Z}).$$

- error_x: the difference between x^0 and $\hat{x}$, i.e.

$$\mathrm{error}_x := \|x^0 - \hat{x}\|.$$

Table 1. Numerical results of Example 5.1.

n	m	time(s)	iter	infeas	error_{cx}	error_x
50	100	5.68e+01	14	9.43e-06	1.40e-06	5.90e-07
50	200	1.16e+02	19	5.11e-06	1.42e-05	6.00e-06
50	1000	4.95e+02	37	2.50e-10	5.71e-07	4.51e-05
100	500	8.83e+02	20	7.78e-06	3.22e-06	5.79e-07
100	1000	1.62e+03	21	1.67e-06	2.03e-05	3.96e-06
100	5000	7.63e+03	38	6.54e-10	2.31e-07	4.92e-05
200	1000	5.80e+03	22	1.42e-06	3.36e-06	5.72e-07
200	3000	1.87e+04	22	3.32e-06	7.15e-05	6.87e-06
200	4000	3.18e+04	28	2.73e-07	1.52e-04	1.48e-05
300	1000	1.24e+04	20	5.57e-06	5.34e-07	1.27e-07
300	2000	2.62e+04	20	5.93e-06	6.53e-06	8.36e-07
400	1000	2.87e+04	21	6.84e-06	4.38e-07	7.39e-08

- error_{cx}: the difference of objective function values of $\mathrm{LSDP}(\hat{\mathrm{c}}, \mathcal{A}, \hat{\mathrm{B}})$ at x^0 and $\hat{x}$, i.e.

$$\mathrm{error}_{cx} := |\langle \hat{c}, x^0 - \hat{x} \rangle|.$$

6. Conclusions

In this paper, we have proposed the penalty framework together with the sequential convex program approach for solving an inverse linear semidefinite programming problem where the parameters in both the objective function and the right-hand side term of the constraint need to be adjusted. Since the inverse problem is actually a mathematical program with semidefinite cone complementarity constraints which is a nonconvex problem and the usual constraint qualifications such as Robinson's CQ fails to hold at each feasible point, we use a nonsmooth partial penalty method as the first step to convert the problem into a nonconvex semidefinite programming problem. Convergence analysis is shown in considerable detail. Then, we introduce the sequential convex program approach as the main algorithm to solve each penalized problems. Fortunately, each substep of the sequential convex program approach which is actually solving two convex problems can be implemented via projection operator and the classical accelerated proximal gradient algorithm. Numerical results of randomly generated inverse linear semidefinite programming problems demonstrate the efficiency of our approach.

Acknowledgments

We would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of this paper. The research is supported by the National Natural Science Foundation of China (Nos. 11301049, 11401210, 11571059 and 91330206), the Fundamental Research Funds for the Central Universities (No. DUT15LK32), and the Doctoral Foundation of Tianjin Normal University (No. 52XB1513).

References

Ahuja, RK and JB Orlin (2001). Inverse optimization. *Operations Research*, 49, 771–783.

Ahuja, RK and JB Orlin (2002). Combinatorial algorithms for inverse network flow problems. *Networks*, 40, 181–187.

Auslender, A (2003). Variational inequalities over the cone of semidefinite positive symmetric matrices and over the Lorentz cone. *Optimization Methods and Software*, 18, 359–376.

Burkard, RE, Y Lin and J Zhang (2004). Weight reduction problems with certain bottleneck objectives. *European Journal of Operational Research*, 153, 191–199.

Burton, WR and PHL Toint (1992). On an instance of the inverse shortest paths problem. *Mathematical Programming*, 53, 45–61.

Cai, MC, XG Yang and J Zhang (1999). The complexity analysis of the inverse center location problem. *Journal of Global Optimization*, 15, 213–218.

Chang, YL and JS Chen (2013). Convexity of symmetric cone trace functions in euclidean Jordan algebras. *Journal of Nonlinear and Convex Analysis*, 14, 53–61.
Clarke, FH (1983). *Optimization and Nonsmooth Analysis*. New York: John Wiley and Sons.
Ding, C (2012). An introduction to a class of matrix optimization problems. Ph.D. Thesis, National University of Singapore.
Ding, C, DF Sun and JJ Ye (2014). First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints. *Mathematical Programming*, 147, 539–579.
Ding, C, DF Sun and KC Toh (2014). An introduction to a class of matrix cone programming. *Mathematical Programming*, 144, 141–179.
Gao, Y and DF Sun (2010). A majorized penalty approach for calibrating rank constrained correlation matrix problems. Technical report, NUS.
Heuberger, C (2004). Inverse combinatorial optimization: A survey on problems, methods and results. *Journal of Combinatorial Optimization*, 8, 329–361.
Hong, J, Y Yang and LW Zhang (2011). Sequential convex approximations to joint chance constrained programs: A Monte Carlo approach. *Operations Research*, 59, 617–630.
Horn, RA and CR Johnson (1985). *Matrix Analysis*. Cambridge University Press.
Iyengar, G and W Kang (2005). Inverse conic programming and applications. *Operations Research Letters*, 33, 319–330.
Jiang, KF, DF Sun and KC Toh (2012). An inexact accelerated proximal gradient method for large scale linearly constrainted convex SDP. *SIAM Journal on Optimization*, 22, 1042–1064.
Jiang, Y, XT Xiao, LW Zhang and JZ Zhang (2011). A perturbation approach for a type of inverse linear programming problems. *International Journal of Computer Mathematics*, 88, 508–516.
Li, Q and HD Qi (2011). A sequential semismooth Newton method for the nearest low-rank correlation matrix problem. *SIAM Journal on Optimization*, 21, 1641–1666.
Luenberger, DG and YY Ye (2008). *Linear and Nonlinear Programming*, 2nd Edition, Springer.
Meng, F, D Sun and G Zhao (2005). Semismoothness of solutions to generalized equations and the Moreau Yosida regularization. *Mathematical Programming Series B*, 104, 561–581.
Mordukhovich, BS (2006) *Variational Analysis and Generalized Differentiation, I: Basic Theory*. Berlin: Springer.
Pang, JS, D Sun and J Sun (2003). Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems. *Mathematics of Operations Research*, 28, 39–63.
Sun, DF (2006). The strong second order sufficient condition and constraint nondegeneracy in nonlinear semidefinite programming and their implications. *Mathematics of Operations Research*, 31, 761–776.
Sun, DF, J Sun and LW Zhang (2008). The rate of convergence of the augmented Lagrangian method for nonlinear semidefinite programming. *Mathematical Programming*, 114, 349–391.
Toh, KC and S Yun (2010). An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems. *Pacific Journal of Optimization*, 6, 615–640.
Wang, SY, YJ Liu and Y Jiang (2014). A majorized penalty approach to inverse linear second order cone programming problems. *Journal of Industrial Management and Optimization*, 10, 965–976.

Wu, J, LW Zhang and Y Zhang (2014). Mathematical programs with semidefinite cone complementarity constraints: Constraint qualifications and optimality conditions. *Set-Valued and Variational Analysis*, 22, 155–187.

Xiao, XT, LW Zhang and JZ Zhang (2009). On convergence of augmented Lagrange method for inverse semi-definite quadratic programming problems. *Journal of Industrial and Management Optimization*, 5, 319–339.

Xiao, XT, J Gu, LW Zhang and SW Zhang (2012). A sequential convex program method to DC program with joint chance constraints. *Journal of Industrial Management and Optimization*, 8, 733–747.

Zhang, J and Z Liu (1996). Calculating some inverse linear programming problems. *Journal of Computational and Applied Mathematics*, 72, 261–273.

Zhang, J and Z Liu (1999). A further study on inverse linear programming problems. *Journal of Computational and Applied Mathematics*, 106, 345–359.

Zhang, J and Z Ma (1999). Solution structure of some inverse combinatorial optimization problems. *Journal of Combinatorial Optimization*, 3, 127–139.

Zhang, JZ and LW Zhang (2010). An augument Lagrangian method for a type of inverse quadratic programming problems. *Applied Mathematics and Optimization*, 61, 57–83.

Biography

Jia Wu received her BS Degree in Applied Mathematics and the MS and PhD degrees in Operational Research from School of Mathematical Sciences, Dalian University of Technology in 2006, 2008 and 2012, respectively. She is currently an Associate Professor in School of Mathematical Sciences, Dalian University of Technology. Her current research interests include nonconvex programming problems and matrix optimization.

Yi Zhang received her BS Degree in Applied Mathematics and the MS and PhD Degrees in Operational Research from School of Mathematical Sciences, Dalian University of Technology in 2006, 2008 and 2012, respectively. She is currently an Assistant Professor in School of Science, East China University of Science and Technology. Her current research interests include nonlinear programming and conic optimization.

Liwei Zhang is a Professor at Dalian University of Technology in the School of Mathematics Science, Dalian, China. He received his MSc in operations research in 1992, PhD in computational mathematics in 1998 from Dalian University of Technology. His main research areas of interest are nonlinear conic constrained optimization, equilibrium constrained optimization and stochastic optimization.

Yue Lu received his BSc Degree, MSc Degree in Mathematics from Suzhou University, Suzhou, China, in 2007, 2010 and the PhD Degree in Optimization from Dalian University of Technology in 2015. He is a Lecturer in School of Mathematical Sciences, Tianjin Normal University, Tianjin, China. His recent research interests include matrix optimization, inverse problems and numerical algorithms.