The Extended Semidefinite Linear Complementarity Problem: A Reformulation Approach

Masahiro Shibata[1], Nobuo Yamashita[2] and Masao Fukushima[2]

Abstract. In this paper we consider the extended semidefinite linear complementarity problem (XSDLCP). We propose merit functions that can constitute minimization problems equivalent to XSDLCP and show some properties of these functions. We also consider the semidefinite nonlinear complementarity problem (SDNCP) and give new sufficient conditions that guarantee the existence of a solution to SDNCP.

Key words. extended semidefinite linear complementarity problem, semidefinite nonlinear complementarity problem, merit function, stationary point, existence of solutions

1. Introduction

The linear complementarity problem (LCP) is to find a vector $x \in \Re^m$ such that

$$\text{[LCP]} \qquad x \geq 0,\ Mx + q \geq 0,\ \langle x, Mx + q\rangle = 0,$$

where M is an $m \times m$ matrix, q is an m-dimensional vector and $\langle \cdot, \cdot \rangle$ denotes the Euclidean inner product. This problem is considered one of the most important problems in operations research, and it has a variety of applications in such problems as linear and quadratic programming problems, bimatrix game problems, market equilibrium problems and network equilibrium problems. See [4, 7]. To deal with a wider class of practical problems, various extensions of LCP have been proposed, and the analysis and solution methods for those problems including the nonlinear complementarity problem (NCP) [9] have been studied extensively [7, 8, 11, 12, 13, 15, 16, 17]. In particular, the extended linear complementarity problem (XLCP) [8, 11, 16, 17] is defined as follows: Find a pair $(x, y) \in \Re^m \times \Re^m$ such that

$$\text{[XLCP]} \qquad Mx - Ny \in C,\ x \geq 0,\ y \geq 0,\ \langle x, y\rangle = 0,$$

where M and N are $n \times m$ matrices and C is a polyhedral set in $\Re^n$. If N is the identity matrix and $C = \{-q\}$, then XLCP reduces to LCP. XLCP itself is also an extension of the horizontal linear complementarity problem which has received increasing attention because of various applications [16, 17]. Another extension of LCP is the semidefinite linear complementarity problem (SDLCP) [1, 6, 10, 12, 15]. Let $\mathcal{S}^m$ be the linear space of $m \times m$ real symmetric matrices. Then SDLCP is defined as follows: Find a matrix $x \in \mathcal{S}^m$ such that

$$\text{[SDLCP]} \qquad y = Mx + q,\ x \in \mathcal{K}^m,\ y \in \mathcal{K}^m, \langle x, y\rangle = 0,$$

where $M : \mathcal{S}^m \to \mathcal{S}^m$ is a linear mapping, $q \in \mathcal{S}^m$, $\mathcal{K}^m \subset \mathcal{S}^m$ denotes the closed convex cone of $m \times m$ positive semidefinite symmetric matrices and $\langle \cdot, \cdot \rangle$ is the inner product in $\mathcal{S}^m$ defined by $\langle x, y\rangle := \mathrm{tr}[xy]$, where $x, y \in \mathcal{S}^m$ and $\mathrm{tr}[\cdot]$ denotes the matrix trace. A further extension of this problem is the semidefinite nonlinear complementarity problem (SDNCP): Find a matrix $x \in \mathcal{S}^m$ such that

$$\text{[SDNCP]} \qquad y = F(x),\ x \in \mathcal{K}^m,\ y \in \mathcal{K}^m,\ \langle x, y\rangle = 0,$$

where $F : \mathcal{S}^m \to \mathcal{S}^m$ is a differentiable mapping. SDLCP and SDNCP are closely related to the optimality conditions for the semidefinite programming problem (SDP) [1, 13] which has recently drawn growing interest in control theory and combinatorial optimization.

[1]Information & Communication System Department, Osaka Gas Co., Ltd, 3-2-27 Minami Chiyozaki, Nishiku, Osaka 550-0023, Japan, masahiro-shibata@osakagas.co.jp

[2]Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan, nobuo@kuamp.kyoto-u.ac.jp, fuku@kuamp.kyoto-u.ac.jp

Putting the above-mentioned extensions together, we consider the extended semidefinite linear complementarity problem (XSDLCP) which is formally stated as follows: Find a pair $(x,y) \in \mathcal{S}^m \times \mathcal{S}^m$ such that

$$[\text{XSDLCP}] \qquad Mx - Ny \in C,\ x \in \mathcal{K}^m,\ y \in \mathcal{K}^m, \langle x, y\rangle = 0,$$

where $M : \mathcal{S}^m \to \mathcal{S}^n$ and $N : \mathcal{S}^m \to \mathcal{S}^n$ are linear mappings, and $C = \{u \in \mathcal{S}^n | Au - b \in \mathcal{K}^k\}$ with $A : \mathcal{S}^n \to \mathcal{S}^k$ being a linear mapping and $b \in \mathcal{S}^k$. Throughout the paper we assume that the feasible set $\{(x,y)|Mx - Ny \in C, x \in \mathcal{K}^m, y \in \mathcal{K}^m\}$ is nonempty. When $\mathcal{S}^m$ and $\mathcal{S}^n$ are restricted to spaces of diagonal matrices, XSDLCP essentially reduces to XLCP. If $m = n$, C is a singleton, and N is the identity mapping, then XSDLCP reduces to SDLCP.

Since XSDLCP is considered an extension of both XLCP and SDLCP, it is natural to expect that XSDLCP can be solved by applying the solution methods developed for XLCP and SDLCP. Recently, various reformulations of XLCP and SDLCP to a minimization problem have been proposed and studied vigorously [10, 11, 12, 15]. The objective function of an equivalent minimization problem is called a merit function. For SDLCP, Tseng [12] first studied such reformulation approaches. He showed that some well-known merit functions for NCP such as the gap function, the regularized gap function, the implicit Lagrangian and the squared Fischer-Burmeister function can be extended to SDLCP. For XLCP, Solodov [11] proposed merit functions which are extensions of the Fischer-Burmeister function and the implicit Lagrangian augmented by a penalty term associated with $Mx - Ny \in C$.

Motivated by the above mentioned development of reformulations for XLCP and SDLCP, we study smooth optimization reformulations of XSDLCP based on the following merit function:

$$f(x,y) := p(x,y) + \psi(x,y), \tag{1.1}$$

where the function $p : \mathcal{S}^m \times \mathcal{S}^m \to \Re$ is a smooth penalization for the set C defined by

$$p(x,y) := \frac{1}{2}\|[-AMx + ANy + b]_+\|^2, \tag{1.2}$$

and $\psi : \mathcal{S}^m \times \mathcal{S}^m \to \Re$ is a nonnegative function satisfying the condition

$$\psi(x,y) = 0 \Leftrightarrow\ \ x \in \mathcal{K}^m,\ y \in \mathcal{K}^m,\ \langle x, y\rangle = 0.$$

Here $[\cdot]_+$ denotes the orthogonal projection onto $\mathcal{K}^k$. A penalization of the form (1.2) has been used by Solodov [11] to construct a merit function for XLCP. As the function ψ, we consider two merit functions studied by Tseng [12] for SDLCP. One is the squared Fischer-Burmeister function

$$\psi_1(x,y) := \frac{1}{2}\|\phi(x,y)\|^2, \tag{1.3}$$

where $\phi : \mathcal{S}^m \times \mathcal{S}^m \to \mathcal{S}^m$ is defined by $\phi(a,b) := (a^2 + b^2)^{\frac{1}{2}} - a - b$, and the other is the implicit Lagrangian involving a scalar parameter $\alpha > 1$

$$\psi_2(x,y) := \max_{\xi,\zeta \in \mathcal{K}^m} \{\langle y, x - \zeta\rangle - \langle \xi, x\rangle - \frac{1}{2\alpha}(\|y - \xi\|^2 + \|x - \zeta\|^2)\}, \tag{1.4}$$

where the norm $\|\cdot\|$ in $\mathcal{S}^m$ is defined by $\|x\| := \sqrt{\langle x, x\rangle} = \sqrt{\mathrm{tr}[x^2]}$. Specifically, we propose the following merit functions for XSDLCP:

$$f(x,y) \ \ = \ \ \frac{1}{2}\|[-AMx + ANy + b]_+\|^2 + \psi_i(x,y) \quad (i = 1,2).$$

It can be shown that the function f is continuously differentiable on $\mathcal{S}^m \times \mathcal{S}^m$. However since f is not convex in general, we may only expect that a decent method for minimizing f yields a stationary point of f. We will give conditions for any stationary point of f to be a solution of the original problem by extending the results obtained by Solodov [11] for XLCP and by Tseng [12] for SDLCP.

This paper is organized as follows. In Section 2, we review some results that will be used in the subsequent analysis. In Section 3, we show conditions under which any stationary point of the merit functions f defined by (1.1) solves XSDLCP. In Section 4, we consider SDNCP and give conditions for the merit functions to be coercive. We also discuss the existence of a solution for SDLCP.

For any matrix $x \in \mathcal{S}^m$, $[x]_+$ and $[x]_-$ denote the orthogonal projections of x onto $\mathcal{K}^m$ and $-\mathcal{K}^m$, respectively. If p is the orthogonal matrix such that $p^T x p = \mathrm{diag}(\lambda_1, \ldots, \lambda_m)$, then $[x]_+ = p\ \mathrm{diag}([\lambda_1]_+, \ldots, [\lambda_m]_+)\ p^T$ and $[x]_- = p\ \mathrm{diag}([\lambda_1]_-, \ldots, [\lambda_m]_-)\ p^T$, where $[\lambda_i]_+ = \max\{0, \lambda_i\}$ and $[\lambda_i]_- = \min\{0, \lambda_i\}$ for $i = 1, \ldots, m$. Therefore, for any $x \in \mathcal{S}^m$, we have $x = [x]_+ + [x]_-$ and $[x]_+[x]_- = 0$. Let $I = \{1, 2, \ldots, m\}$. For any $x \in \mathcal{S}^m$, we denote by x_{ij} the (i,j)th entry of x. For a function $F : \mathcal{S}^m \to \mathcal{S}^m$, $\nabla F(x)$ denotes the Jacobian of F at $x \in \mathcal{S}^m$, which may be viewed as a linear mapping from $\mathcal{S}^m$ to $\mathcal{S}^m$.

2. Basic results

We review some basic results that will play an important role in the subsequent analysis. We basically adopt the notations used in [12].

Definition 2.1 **(a)** *For any $m \times m$ matrix x, $\mathrm{sym}[x]$ is defined by $\mathrm{sym}[x] := x + x^T$, where x^T denotes the transpose of x.*

(b) *For any $c \in \mathcal{K}^m$, the set $\mathcal{S}^m_c \subset \mathcal{S}^m$ denotes the subspace comprising those $x \in \mathcal{S}^m$ whose null space contains the null space of c.*

(c) *For any $c \in \mathcal{K}^m$, the linear mapping $L_c : \mathcal{S}^m_c \to \mathcal{S}^m_c$ is defined by $L_c[x] := cx + xc$.*

(d) *For any subspace $\mathcal{T} \subseteq \mathcal{S}^m$, a mapping $G : \mathcal{T} \to \mathcal{T}$ is said to be positive definite on $\mathcal{T}$ if*

$$\langle x, G[x] \rangle > 0 \quad \text{for any } x \in \mathcal{T} \ \ (x \neq 0).$$

With these definitions, the following properties hold.

Lemma 2.1 *[12]*

(a) *For any $a, b \in \mathcal{S}^m$, let $c := (a^2 + b^2)^{\frac{1}{2}}$. Then $a, b \in \mathcal{S}^m_c$.*

(b) *For any $c \in \mathcal{K}^m$, the linear mapping L_c is positive definite and hence invertible on $\mathcal{S}^m_c$. Specifically, for any $x \in \mathcal{S}^m_c$, $L_c^{-1}[x]$ is the unique matrix $y \in \mathcal{S}^m_c$ satisfying $cy + yc = x$.*

(c) *For any $c \in \mathcal{K}^m$ and $x, y \in \mathcal{S}^m_c$, we have $\langle y, L_c[x] \rangle = \langle L_c[y], x \rangle$, $xL_c^{-1}[c] = L_c^{-1}[c]x = \frac{1}{2}x$ and $xL_c^{-1}[x] = 0 \Rightarrow x = 0$.* □

The merit function f defined by (1.1) consists of the penalty function p defined by (1.2) and either of the two functions ψ_1 defined by (1.3) and ψ_2 defined by (1.4). We show some properties of ψ_i $(i = 1, 2)$ and p, which will be used in the next subsection where we consider conditions under which any stationary point of f is a solution of XSDLCP.

Lemma 2.2 *[12, 15]*

(a) *$\psi_1(a, b) \geq 0$ for all $(a, b) \in \mathcal{S}^m \times \mathcal{S}^m$. $\psi_1(a, b) = 0$ and $(a, b) \in \mathcal{S}^m \times \mathcal{S}^m$ if and only if $a \in \mathcal{K}^m$, $b \in \mathcal{K}^m$ and $\langle a, b \rangle = 0$.*

(b) *ψ_1 is differentiable on $\mathcal{S}^m \times \mathcal{S}^m$ and*

$$\begin{aligned} \nabla_a \psi_1(a, b) &= \mathrm{sym}[L_c^{-1}[c - a - b](a - c)] \in \mathcal{S}^m, \\ \nabla_b \psi_1(a, b) &= \mathrm{sym}[L_c^{-1}[c - a - b](b - c)] \in \mathcal{S}^m, \end{aligned}$$

where $c := (a^2 + b^2)^{\frac{1}{2}} \in \mathcal{K}^m$.

(c) *For every* $(a,b) \in \mathcal{S}^m \times \mathcal{S}^m$, *we have*

$$\langle \nabla_a \psi_1(a,b), \nabla_b \psi_1(a,b) \rangle \geq \|(c-a-b)g\|^2,$$

where $c := (a^2+b^2)^{\frac{1}{2}} \in \mathcal{K}^m$ *and* $g := L_c^{-1}[c-a-b] \in \mathcal{S}^m$.

(d) *For every* $(a,b) \in \mathcal{S}^m \times \mathcal{S}^m$, *we have*

$$\langle a, \nabla_a \psi_1(a,b) \rangle + \langle b, \nabla_b \psi_1(a,b) \rangle = \|c-a-b\|^2,$$

where $c := (a^2+b^2)^{\frac{1}{2}} \in \mathcal{K}^m$. □

Lemma 2.3 *[12]*

(a) $\psi_2(a,b) \geq 0$ *for all* $(a,b) \in \mathcal{S}^m \times \mathcal{S}^m$. $\psi_2(a,b) = 0$ *and* $(a,b) \in \mathcal{S}^m \times \mathcal{S}^m$ *if and only if* $a \in \mathcal{K}^m$, $b \in \mathcal{K}^m$ *and* $\langle a,b \rangle = 0$.

(b) *For any fixed* $\alpha \in (0,\infty)$, ψ_2 *is differentiable on* $\mathcal{S}^m \times \mathcal{S}^m$ *and*

$$\begin{aligned} \nabla_a \psi_2(a,b) &= b - [b-\alpha a]_+ - \frac{1}{\alpha}(a - [a-\alpha b]_+), \\ \nabla_b \psi_2(a,b) &= a - [a-\alpha b]_+ - \frac{1}{\alpha}(b - [b-\alpha a]_+). \end{aligned}$$

(c) *For every* $(a,b) \in \mathcal{S}^m \times \mathcal{S}^m$, *we have* $\langle \nabla_a \psi_2(a,b), \nabla_b \psi_2(a,b) \rangle \geq 0$.

(d) $\nabla_a \psi_2(a,b) = 0$ *and* $\nabla_b \psi_2(a,b) = 0$ *if and only if* $\psi_2(a,b) = 0$. □

Next we give some important properties which will be used when we investigate the properties of the penalty function p .

Lemma 2.4 *For any* $x, y \in \mathcal{S}^m$, *we have*

$$\|[x+y]_+\| \leq \|[x]_+\| + \|[y]_+\|.$$

Proof. Let us consider the following optimization problem:

$$\min_{z \in -\mathcal{K}^m} \|x+y-z\|. \tag{2.1}$$

Since $[x]_- + [y]_- \in -\mathcal{K}^m$, we have

$$\min_{z \in -\mathcal{K}^m} \|x+y-z\| \leq \|x+y-[x]_- - [y]_-\| = \|[x]_+ + [y]_+\| \leq \|[x]_+\| + \|[y]_+\|,$$

where the last inequality follows from the triangle inequality. By the definition of the orthogonal projections, the optimal solution of problem (2.1) is attained at $z = [x+y]_-$. Since $\|[x+y]_+\| = \|x+y-[x+y]_-\|$ for any $x, y \in \mathcal{S}^m$, we have

$$\|[x+y]_+\| = \min_{z \in -\mathcal{K}^m} \|x+y-z\| \leq \|[x]_+\| + \|[y]_+\|.$$

The proof is complete. □

Lemma 2.5 *Let the function* $g : \mathcal{S}^k \to \Re$ *be defined by* $g(z) = \|[z]_+\|^2$. *Then (a) the function* g *is convex on* $\mathcal{S}^k$ *and (b) the function* g *is differentiable on* $\mathcal{S}^k$ *with* $\nabla g(x) = 2[z]_+$.

Proof. (a) Let the functions $g_1 : \mathcal{S}^k \to \Re$ and $g_2 : \Re \to \Re$ be defined by $g_1(u) := \|[u]_+\|$ and $g_2(t) := t^2$. Note that g_1 is a convex function on $\mathcal{S}^k$ by Lemma 2.4. Then, for any $x, y \in \mathcal{S}^m$ and $\alpha \in [0,1]$, we have

$$\begin{aligned} g((1-\alpha)x + \alpha y) &= g_2(g_1((1-\alpha)x + \alpha y)) \\ &\leq g_2((1-\alpha)g_1(x) + \alpha g_1(y)) \\ &\leq (1-\alpha)g_2(g_1(x)) + \alpha g_2(g_1(y)) \\ &= (1-\alpha)g(x) + \alpha g(y), \end{aligned}$$

where the first inequality follows from the convexity and nonnegativity of the function g_1, and the monotonicity of the function g_2 on $[0,\infty)$, and the second inequality follows from the convexity of the function g_2.
(b) For any $x \in \mathcal{S}^k$, we have $g(x) = \|x - [x]_-\|^2 = \min_{y \in -\mathcal{K}^k} \|x - y\|^2$. Let us define the function $h : \mathcal{S}^k \times \mathcal{S}^k \to \Re$ by $h(x,y) = \|x - y\|^2$. Note that the function h is differentiable. Since, by definition,

$$g(x) = \min\{h(x,y) | y \in -\mathcal{K}^k\} \tag{2.2}$$

and the minimum on the right-hand side of (2.2) is uniquely attained at $y = [x]_-$, it follows from [2, Chapter 4, Theorem 1.7] that g is differentiable and $\nabla g(x)$ is given by $\nabla g(x) = \nabla_x h(x, [x]_-)$. Thus we have $\nabla g(x) = 2(x - [x]_-) = 2[x]_+$. □

We define some concepts about linear mappings on $\mathcal{S}^m$.

Definition 2.2 (a) *A linear mapping $M : \mathcal{S}^m \to \mathcal{S}^m$ is positive semidefinite if*

$$\langle y, My \rangle \geq 0 \quad \text{for all } y \in \mathcal{S}^m.$$

(b) *A linear mapping $M : \mathcal{S}^m \to \mathcal{S}^m$ is said to be copositive (respectively, strictly copositive) on the cone $\mathcal{T} \subset \mathcal{S}^m$, if the following holds:*

$$\langle x, Mx \rangle \geq 0 \quad (\text{respectively, } \langle x, Mx \rangle > 0) \quad \text{for all } x \in \mathcal{T}, x \neq 0.$$

(c) *The adjoint of a linear mapping $M : \mathcal{S}^m \to \mathcal{S}^k$ is the linear mapping $M^* : \mathcal{S}^k \to \mathcal{S}^m$ such that $\langle Mu, \mu \rangle = \langle u, M^* \mu \rangle$ for all $u \in \mathcal{S}^m$ and $\mu \in \mathcal{S}^k$.*

Finally, we generalize the recession cone in the m-dimensional real space to the space of $m \times m$ symmetric matrices.

Definition 2.3 *The recession cone 0^+C of the set $C = \{u \in \mathcal{S}^n | Au - b \in \mathcal{K}^k\}$ is given by $0^+C = \{u \in \mathcal{S}^n | Au \in \mathcal{K}^k\}$, and its dual cone $(0^+C)^*$ is represented as*

$$(0^+C)^* := \{v \in \mathcal{S}^n | \langle u, v \rangle \geq 0 \text{ for all } u \in 0^+C\} = \{v \in \mathcal{S}^n | v = A^*\mu, \mu \in \mathcal{K}^k\}.$$

3. Merit functions for XSDLCP

In this section, we show the equivalence between XSDLCP and the minimization problem

$$\text{minimize } f(x,y), \quad (x,y) \in \mathcal{S}^m \times \mathcal{S}^m, \tag{3.1}$$

where f is defined by (1.1). Moreover, we give conditions which guarantee that every stationary point of the unconstrained minimization problem (3.1) is a solution of XSDLCP.

The following theorem implies the equivalence between XSDLCP and the minimization problem (3.1), no matter whether f consists of ψ_1 or ψ_2.

Theorem 3.1 *Let f be defined by (1.1) and ψ be given by either ψ_1 or ψ_2. Then f is nonnegative on $\mathcal{S}^m \times \mathcal{S}^m$. Moreover, $f(x,y) = 0$ and $(x,y) \in \mathcal{S}^m \times \mathcal{S}^m$ if and only if (x,y) is a solution of XSDLCP.*

Proof. The theorem is evident from the definition of f, Lemma 2.2 (a) and Lemma 2.3 (a). □

This theorem indicates that we can solve XSDLCP by finding a global optimum of the minimization problem

$$\min_{(x,y) \in \mathcal{S}^m \times \mathcal{S}^m} f(x,y) = \min_{(x,y) \in \mathcal{S}^m \times \mathcal{S}^m} \frac{1}{2} \|[-AMx + ANy + b]_+\|^2 + \psi(x,y), \tag{3.2}$$

where ψ is either ψ_1 defined by (1.3) or ψ_2 defined by (1.4). However, the function f is not convex in general. So it is important to know when any stationary point of f is a global minimum of f. In the next two theorems, we present sufficient conditions for this property to hold.

First we consider the merit function that uses the squared Fischer-Burmeister function.

Theorem 3.2 *Let f be defined by ψ_1. If the mapping MN^* is copositive on $(0^+C)^*$, then every stationary point of f is a solution of XSDLCP.*

Proof. Let $(x,y) \in \mathcal{S}^m \times \mathcal{S}^m$ be a stationary point of problem (3.2). Then we have

$$-M^*v + \nabla_a \psi_1(x,y) = 0 \text{ and } N^*v + \nabla_b \psi_1(x,y) = 0, \tag{3.3}$$

where $v = A^*[-AMx + ANy + b]_+ \in (0^+C)^*$. It follows from (3.3) that

$$\langle \nabla_a \psi_1(x,y), \nabla_b \psi_1(x,y) \rangle = -\langle M^*v, N^*v \rangle = -\langle MN^*v, v \rangle. \tag{3.4}$$

By Lemma 2.2 (c), we have $\langle \nabla_a \psi_1(x,y), \nabla_b \psi_1(x,y) \rangle = \|(c-a-b)g\|^2 \geq 0$, where $a := x, b := y, c := (a^2+b^2)^{\frac{1}{2}}$ and $g := L_c^{-1}(c-a-b)$. It then follows from (3.4) and the copositiveness of MN^* on $(0^+C)^*$ that $\langle \nabla_a \psi_1(x,y), \nabla_b \psi_1(x,y) \rangle = 0$, which implies $\|(c-a-b)g\|^2 = 0$. Hence we have $(c-a-b)g = 0$. It follows from Lemma 2.1 (c) that $(x^2+y^2)^{\frac{1}{2}} = x+y$. Hence by Lemma 2.2 (a), we obtain $x \in \mathcal{K}^m, y \in \mathcal{K}^m, \langle x,y \rangle = 0$. Also from Lemma 2.3 (b), we have $\nabla_a \psi_1(x,y) = \nabla_b \psi_1(x,y) = 0$. Therefore the stationarity conditions (3.3) for (3.2) imply $-M^*v = 0$ and $-N^*v = 0$, which in turn imply $\nabla p(x,y) = 0$, i.e., (x,y) is a stationary point of the function p. Since p is a convex function by Lemma 2.5 (a), (x,y) is in fact a global minimum of p. Moreover, since the set $\{(x,y) | Mx - Ny \in C, x \in \mathcal{K}^m, y \in \mathcal{K}^m\}$ is nonempty, the global minimum (x,y) actually satisfies $Mx - Ny \in C$. Thus (x,y) solves XSDLCP. □

Next, we consider the merit function f that uses the implicit Lagrangian function ψ_2.

Theorem 3.3 *Let f be defined with ψ_2. Suppose that MN^* is strictly copositive on $(0^+C)^*$. Then any stationary point of f is a solution of XSDLCP.*

Proof. Let (x,y) be a stationary point of f. Then, in a way similar to the proof of Theorem 3.2, we obtain $\langle \nabla_a \psi_2(x,y), \nabla_b \psi_2(x,y) \rangle = -\langle MN^*v, v \rangle$. Since MN^* is strictly copositive on $(0^+C)^*$, it follows from Lemma 2.3 (c) that $v = A^*[-AMx + ANy + b]_+ = 0$. From this, we can deduce $\nabla \psi_2(x,y) = 0$. By applying Lemma 2.3 (d), we have $\psi_2(x,y) = 0$. It then follows from Lemma 2.3 (a) that $x \in \mathcal{K}^m, y \in \mathcal{K}^m, \langle x,y \rangle = 0$. Moreover, by an argument similar to the proof of Theorem 3.2, we obtain $[-AMx + ANy + b]_+ = 0$. Hence (x,y) solves XSDLCP. □

Note that Theorem 3.3 assumes the strict copositivity of MN^* on $(0^+C)^*$, which is stronger than the copositivity assumption in Theorem 3.2. This is because the equality

$$\langle \nabla_a \psi_2(x,y), \nabla_b \psi_2(x,y) \rangle = 0$$

dose not necessarily mean that the pair (x,y) satisfies the complementarity condition.

4. Existence results for SDNCP

In this section, we consider the existence of solutions for SDNCP. To establish the existence results, we will show that a merit function for SDNCP has at least one stationary point and that such a stationary point is a solution of SDNCP. By this approach, Yamashita and Fukushima [15] showed that SDNCP has a solution if F is monotone and SDNCP is strictly feasible. This result was based on the fact that $\nabla F(x)$ being positive semidefinite guarantees that a stationary point of the squared Fischer-Burmeister function for SDNCP is a solution of SDNCP [12]. Recently, this condition has been improved by Qi and Chen [10]. Specifically, they showed that the positive semidefiniteness of $\nabla F(x)$ can be weakened to the condition that $\nabla F(x)$ is a P_0-matrix, whose definition is stated as follows.

Definition 4.1 *A linear mapping $M : \mathcal{S}^m \to \mathcal{S}^m$ is called a P_0-matrix (respectively, P-matrix) if for any $x \in \mathcal{S}^m$ with $x \neq 0$ and any nonsingular matrix $p \in \Re^{m\times m}$, there exists an index $i \in I$ such that $(px)_i \neq 0$ and $[(px)M(px)^T]_{ii} \geq 0$ (respectively, $[(px)M(px)^T]_{ii} > 0$), where $(px)_i$ is the ith row of matrix px.*

Moreover, the existence of solutions for NCP is guaranteed under the condition that F is a uniform P-function. Naturally we may expect that this property is preserved for SDNCP. To see this, we need to introduce the notions of P-function and uniform P-function for functions from $\mathcal{S}^m$ into $\mathcal{S}^m$.

Definition 4.2 *A function $F : \mathcal{S}^m \to \mathcal{S}^m$ is called*

(i) *a P-function if, for any $x, y \in \mathcal{S}^m$ with $x \neq y$ and any nonsingular matrix p, there exists an index $i \in I$ such that $[p(x - y)(F(x) - F(y))p^T]_{ii} > 0$,*

(ii) *a uniform P-function if, for any $x, y \in \mathcal{S}^m$ and any nonsingular matrix p, there exists an index $i \in I$ such that $[p(x - y)(F(x) - F(y))p^T]_{ii} \geq \mu\|p(x - y)\|^2$, where μ is a positive constant.*

The following lemma, which reveals a relationship between P-function and P_0-matrix, will play an important role in studying conditions for the existence of a solution of SDNCP.

Lemma 4.1 *Suppose that $F : \mathcal{S}^m \to \mathcal{S}^m$ is differentiable. If F is a P-function, then $\nabla F(x)$ is a P_0-matrix. Moreover, if F is a uniform P-function, then for any $d \in \mathcal{S}^m$ with $d \neq 0$ there exists an index $i \in I$ such that*

$$[(pd)\nabla F(y)(pd)^T]_{ii} \geq \mu\|pd\|^2,$$

where μ is a positive constant.

Proof. For any $d \in \mathcal{S}^m$ with $d \neq 0$, let p be an orthogonal matrix p such that $p^T dp = \mathrm{diag}(\lambda_1, \cdots, \lambda_m)$, where λ_i are eigenvalues of d. From the differentiability of F we have for any $x \in \mathcal{S}^m$

$$F(y) - F(x) = \nabla F(x)(td) + o(t)\|d\|, \tag{4.1}$$

where $y = x + td$ and $t > 0$. Since F is a P-function, there exists an index i such that

$$[p(y - x)(F(y) - F(x))p^T]_{ii} > 0. \tag{4.2}$$

Note that (4.2) implies $[p(y - x)]_i = t(pd)_i \neq 0$. Thus (4.1) and (4.2) yield

$$[p(td)\{t\nabla F(x)d + o(t)\|d\|\}p^T]_{ii} = [t^2(pd)\nabla F(x)(pd)^T + t(pd)o(t)\|d\|p^T]_{ii} \ > \ 0, \tag{4.3}$$

where the equality follows from the symmetry of d. Dividing the inequality (4.3) by t^2 and letting $t \to 0$, we have $[(pd)\nabla F(x)(pd)^T]_{ii} \geq 0$. (Note that the index i in (4.3) depends on t. Nevertheless, we may assume without loss of generality that the index i is fixed because there are only finitely many indices.) Thus $\nabla F(x)$ is a P_0-matrix. In a similar way, we can show the last half of the lemma. □

Note that every uniform P-function is a P-function, so if F is a differentiable uniform P-function, then $\nabla F(x)$ is a P_0-matrix for any x.

Remark 4.1 Suppose that $F(x) = Mx + q$. If F is a P-function, then, by Definition 4.2 (ii), for any $x, y \in \mathcal{S}^m$ with $x \neq y$, there exists an index $i \in I$ such that $[(p(x - y))M(p(x - y))^T]_{ii} > 0$. This implies that M is a P-matrix.

To show that F being a uniform P-function guarantees the existence of a solution to SDNCP, we use, as a theoretical tool, the merit function $\Psi : \mathcal{S}^m \to \Re$ defined by

$$\Psi(x) = \max\{0, \langle x, F(x)\rangle\}^2 + \Psi_1(x), \tag{4.4}$$

where $\Psi_1 : \mathcal{S}^m \to \Re$ is given by $\Psi_1(x) = \frac{1}{2}\|\phi(x, F(x))\|^2$. The function Ψ was proposed by Yamashita and Fukushima [15].

The following proposition constitutes the main step in proving the existence of a solution of SDNCP. Before proceeding to the proposition, we introduce some notation. For any $J, L \subseteq I$, we denote by x_{JL} the submatrix of x with rows $j \notin J$ and columns $l \notin L$ removed. Also, we will use "$\omega(t)$" as a shorthand to denote an element of $\mathcal{S}^m$ which depends on t and whose norm tends to ∞ slower than t^2, i.e., $\lim_{t\to\infty} \|\omega(t)\|/t^2 = 0$.

Proposition 4.1 *Let $\lambda_{max}(x)$ denote the maximum eigenvalue of $x \in \mathcal{S}^m$. Suppose that $F : \mathcal{S}^m \to \mathcal{S}^m$ is such that $F(x) = \omega(\lambda_{max}(x))$ and F is a uniform P-function. Then the function $\Psi : \mathcal{S}^m \to \Re$ defined by (4.4) is coercive, i.e., $\lim_{\|x\|\to\infty} \Psi(x) = \infty$.*

Proof. Suppose that the proposition is false. Then we can find an unbounded sequence of $\{x^k\}$ such that $\{\Psi(x^k)\}$ is bounded. It can be seen that the sequences of the minimum eigenvalues of $\{x^k\}$ and $\{F(x^k)\}$ are bounded below. In fact, if not, it follows from [15, Lemma 2.6] that $\Psi_1(x^k) \to \infty$, which implies $\Psi(x^k) \to \infty$. This contradicts the boundedness of the sequence $\{\Psi(x^k)\}$. Hence the unboundedness of $\{x^k\}$ implies that the sequence of the maximum eigenvalues of $\{x^k\}$ tends to ∞. By the assumption of the proposition, for any k, there exists an index i_k such that

$$[p^k x^k (F(x^k) - F(0))(p^k)^T]_{i_k i_k} \geq \mu \parallel p^k x^k \parallel^2,$$

where μ is a positive constant and p^k is an orthogonal matrix $p^k \in \Re^{m\times m}$ such that $(p^k)^T x^k p^k = \mathrm{diag}(\lambda_1^k, \cdots, \lambda_m^k)$ with λ_i^k being the eigenvalues of x^k. Let λ_{max}^k be the maximum eigenvalue of x^k. Then we have

$$\mu(\lambda_{max}^k)^2 \leq \mu \parallel p^k x^k \parallel^2 \leq [p^k\{x^k F(x^k) - x^k F(0)\}(p^k)^T]_{i_k i_k}, \tag{4.5}$$

where the first inequality follows from the definition of the norm on $\mathcal{S}^m$. Since $x^k F(0) = \omega(\lambda_{max}^k)$, (4.5) implies

$$[p^k x^k F(x^k)(p^k)^T]_{i_k i_k} \geq c(\lambda_{max}^k)^2, \tag{4.6}$$

where c is a positive constant.

Now, for the purpose of contradiction, we will prove that $\Psi(x^k) \to \infty$ as $\lambda_{max}^k \to \infty$. To this end, it is sufficient to show that $\langle x^k, F(x^k)\rangle \to \infty$ because Ψ_1 is nonnegative on $\mathcal{S}^m$. Note that

$$\begin{aligned} \langle x^k, F(x^k)\rangle &= \mathrm{tr}[p^k x^k F(x^k)(p^k)^T] \\ &= \langle p^k x^k (p^k)^T, p^k F(x^k)(p^k)^T\rangle \\ &= \langle \mathrm{diag}(\lambda_1^k, \cdots, \lambda_m^k), p^k F(x^k)(p^k)^T\rangle. \end{aligned} \tag{4.7}$$

We may assume, without loss of generality, that $i_k \equiv 1$ in (4.6). Furthermore, for the sake of convenience, we denote $\alpha^k := \lambda_1^k [p^k F(x^k)(p^k)^T]_{11}$. Since

$$\begin{aligned} [p^k x^k F(x^k)(p^k)^T]_{11} &= [p^k x^k (p^k)^T p^k F(x^k)(p^k)^T]_{11} \\ &= [\mathrm{diag}(\lambda_1^k, \ldots, \lambda_m^k) p^k F(x^k)(p^k)^T]_{11} \\ &= \lambda_1^k [p^k F(x^k)(p^k)^T]_{11}, \end{aligned}$$

it follows from (4.6) and the definition of α^k that

$$\alpha^k \geq c(\lambda_{max}^k)^2. \tag{4.8}$$

Let $J = \{2, 3, \cdots, n\}$. For simplicity, we denote $\Lambda^k = \mathrm{diag}(\lambda_2^k, \cdots, \lambda_n^k)$ and $\hat{F}(x^k) = [p^k F(x^k)(p^k)^T]_{JJ}$. Then (4.7) yields

$$\begin{aligned} \langle x^k, F(x^k)\rangle &= \alpha^k + \langle \Lambda^k, \hat{F}(x^k)\rangle \\ &= \alpha^k + \langle [\Lambda]_+, [\hat{F}(x^k)]_+\rangle + \langle [\Lambda^k]_+, [\hat{F}(x^k)]_-\rangle \\ &\quad + \langle [\Lambda^k]_-, [\hat{F}(x^k)]_+\rangle + \langle [\Lambda^k]_-, [\hat{F}(x^k)]_-\rangle. \end{aligned} \tag{4.9}$$

Note that $\langle [\Lambda^k]_+, [\hat{F}(x^k)]_+ \rangle \geq 0$ and $\langle [\Lambda^k]_-, [\hat{F}(x^k)]_- \rangle \geq 0$. Thus we compare the other two terms $\langle [\Lambda^k]_+, [\hat{F}(x^k)]_- \rangle$ and $\langle [\Lambda^k]_-, [\hat{F}(x^k)]_+ \rangle$ with α^k. First we consider $\langle [\Lambda^k]_-, [\hat{F}(x^k)]_+ \rangle$. Since the eigenvalues of $\{x^k\}$ are bounded below, those of $\{\Lambda^k\}$ are also bounded below. Thus, by the assumption that $F(x) = \omega(\lambda_{max}(x))$, we have

$$\langle [\Lambda^k]_-, [\hat{F}(x^k)]_+ \rangle = \omega(\lambda^k_{max}). \tag{4.10}$$

Next we consider $\langle [\Lambda^k]_+, [\hat{F}(x^k)]_- \rangle$. Since the eigenvalues of $\{F(x^k)\}$ are bounded below and $\hat{F}(x^k)$ is the principal submatrix of $p^k F(x^k)(p^k)^T$, it follows from [5, Corollary 8.1.4] that the eigenvalues of $\{\hat{F}(x^k)\}$ are also bounded below. So we have

$$\langle [\Lambda^k]_+, [\hat{F}(x^k)]_- \rangle = \omega(\lambda^k_{max}). \tag{4.11}$$

Thus (4.8), (4.9), (4.10) and (4.11) yield $\langle x^k, F(x^k) \rangle \geq \bar{c}(\lambda^k_{max})^2$ for some constant $\bar{c} > 0$. Hence $\langle x^k, F(x^k) \rangle \to \infty$ as $\lambda^k_{max} \to \infty$. This implies $\Psi(x^k) \to \infty$, which contradicts the assumption that $\{\Psi(x^k)\}$ is bounded. □

In the case where $F(x) = Mx + q$, it is clear that $F(x) = \omega(\lambda_{max}(x))$. Furthermore if M is a P-matrix, F is a uniform P-function. Hence this proposition implies that, in SDLCP, if M is a P-matrix, then the merit function Ψ is coercive.

Note that Proposition 4.1 says that the level sets $\mathcal{L}(\gamma) := \{x \in \mathcal{S} | \Psi(x) \leq \gamma\}$ are compact for all $\gamma \in \Re$. We are now in a position to prove the main result in this section.

Theorem 4.1 *Let $\lambda_{max}(x)$ denote the maximum eigenvalue of x. Suppose that $F : \mathcal{S}^m \to \mathcal{S}^m$ is differentiable and $F(x) = \omega(\lambda_{max}(x))$. If F is a uniform P-function, then SDNCP has a solution.*
Proof. Let $x^0 \in \mathcal{S}$ be arbitrary and let $\gamma := \Psi(x^0)$. By Proposition 4.1, the level set $\mathcal{L}(\gamma)$ is nonempty and compact. Hence the function Ψ attains its global minimum $\bar{x}$ on $\mathcal{L}(\gamma)$. In particular, $\bar{x}$ is a stationary point of Ψ. Since F is a uniform P-function, $\nabla F(\bar{x})$ is a P_0-matrix by Lemma 4.1. Hence, from [15, Theorem3.3], SDNCP has a solution. □

As a direct consequence of Theorem 4.1, we obtain the following corollary to SDLCP.

Corollary 4.1 *Let $F : \mathcal{S}^m \to \mathcal{S}^m$ be defined by $F(x) = Mx + q$, where $M : \mathcal{S}^m \to \mathcal{S}^m$ is a linear mapping and $q \in \mathcal{S}^m$. If M is a P-matrix, then SDLCP has a solution.* □

5. Concluding remark

In this paper, we have proposed a class of merit functions for the extended semidefinite linear complementarity problem (XSDLCP). The proposed merit functions consist of a smooth exterior penalty function and either the squared Fischer-Burmeister function or the implicit Lagrangian. By using those merit functions, we reformulated XSDLCP as unconstrained minimization problems. Furthermore we established sufficient conditions which guarantee that any stationary point of the merit functions is a solution of XSDLCP. By using another merit function, we also gave conditions for the existence of a solution of SDNCP.

References

[1] F. Alizadeh, "Interior point methods in semidefinite programming with applications to combinatorial optimization," *SIAM Journal on Optimization* 5: 13-51, 1995.

[2] A. Auslender, *Optimisation: Mèthodes Numèriques*, Masson, Paris, 1976.

[3] S, Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, 1994.

[4] R.W. Cottle, J.S. Pang and R.E. Stone, *The Linear Complementarity Problem*, Academic Press, San Diego, 1992.

[5] G.H. Golub and C.F. Van Loan, "The symmetric eigenvalue problem," in *Matrix Computation*, A.S. Householder and J.M. Wilkinson, eds., The Johns Hopkins University Press, Baltimore, 1989, pp.409-474.

[6] M. Kojima, S. Shindo and S. Hara, "Interior-point methods for monotone semidefinite linear complementarity problems," *SIAM Journal on Optimization* 7: 86-125, 1997.

[7] O.L. Mangasarian and M.V. Solodov, "Nonlinear complementarity as unconstrained and constrained minimization," *Mathematical Programming* 62: 277-297, 1993.

[8] O.L. Mangasarian and J.-S. Pang, "The extended linear complementarity problem," *SIAM Journal on Matrix Analysis and Applications* 16: 359-368, 1995.

[9] J.-S. Pang, "Complementarity problems," in *Handbook of Global Optimization*, R. Horst and P. Pardalos, eds., Kluwer Academic Publishers, Boston, 1994, pp.271-338.

[10] H.D. Qi and X. Chen, "On stationary points of merit functions for semi-definite complementarity problems," Working paper, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, Beijing, China, 1997.

[11] M.V. Solodov, "Some optimization reformulations for the extended linear complementarity problem," Working paper, Instituto de Matemática Pure e Aplicada, Rio de Janeiro, Brazil, 1997.

[12] P. Tseng, "Merit functions for semi-definite complementarity problems," *Mathematical Programming*, to appear.

[13] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review* 38: 49-95, 1996.

[14] L. Vandenberghe, B. De Moor and J. Vandewalle, "The generalized linear complementarity problem applied to the complete analysis of resistive piecewise linear circuits," *IEEE Transactions on Circuits and Systems* 36: 1382-1391, 1989.

[15] N. Yamashita and M. Fukushima, "A new merit function and a descent method for semidefinite complementarity problems", in *Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, M. Fukushima and L. Qi, eds., Kluwer Academic Publishers, Boston, 1998, to appear.

[16] Y. Ye, "A fully polynomial-time approximation algorithm for computing a stationary point of the general linear complementarity problem," *Mathematics of Operations Research* 18: 334-345, 1993.

[17] Y. Zhang, "On the convergence of a class of infeasible interior-point algorithms for the horizontal linear complementarity problem," *SIAM Journal on Optimization* 4: 208-227, 1994.