AMO - Advanced Modeling and Optimization, Volume 13, Number 2, 2011

An R-Linearly Convergent Nonmonotone Derivative-Free Method for Symmetric Cone Complementarity Problems

Shaohua Pan [1]
Department of Mathematics
South China University of Technology
Guangzhou 510640, China
E-mail: shhpan@scut.edu.cn.

Jein-Shan Chen [2]
Department of Mathematics
National Taiwan Normal University
Taipei, Taiwan 11677
E-mail: jschen@math.ntnu.edu.tw

Abstract. This paper extends the derivative-free descent method [18] for the nonlinear complementarity problem to the symmetric cone complementarity problem (SCCP). The algorithm is based on the unconstrained implicit Lagrangian reformulation of the SCCP, but uses a convex combination of the negative partial gradients of the implicit Lagrangian function ψ_α, i.e. the vector of the form $-\theta\nabla_x\psi_\alpha - (1-\theta)\nabla_y\psi_\alpha$ for $\theta \in [0,1]$, as the search direction, and a nonmonotone line search rule to seek a desirable stepsize. We show that the derivative-free algorithm converges in terms of the implicit Lagrangian value for a large class of SCCPs that may even not be monotone. If θ is restricted to be less than a threshold $\bar{\theta} \in (0,1)$ and the SCCP is strongly monotone, the sequence generated converges globally to the solution of SCCP at a R-linear rate.

Key words: Symmetric cone complementarity problem, implicit Lagrangian, derivative-free methods, nonmonotone, linear convergence.

[1]The author's work is supported by Guangdong Natural Science Foundation (No. 9251802902000001) and the Fundamental Research Funds for the Central Universities (SCUT).
[2]Corresponding author. Member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. The author's work is partially supported by National Science Council of Taiwan.

1 Introduction

Let $\mathbb{V}$ be a finite-dimensional vector space over the real field $\mathbb{R}$, $\mathbb{A} \equiv (\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ be a Euclidean Jordan algebra (see Section 2 for the definition), and $\mathcal{K}$ be a symmetric cone in $\mathbb{A}$. Given a continuously differentiable mapping $F : \mathbb{V} \to \mathbb{V}$, we are interested in the following symmetric cone complementarity problem (SCCP): to *find a* $\zeta \in \mathbb{V}$ *such that*

$$\zeta \in \mathcal{K}, \quad F(\zeta) \in \mathcal{K}, \quad \langle \zeta, F(\zeta)\rangle = 0. \tag{1}$$

This class of problem provides a unified framework for the classical nonlinear complementarity problem (NCP), the second-order cone complementarity problem (SOCCP), and the semidefinite complementarity problem (SDCP), as well as arises from the KKT system of a nonlinear symmetric cone optimization problem. When $F(\zeta) = \mathcal{L}(\zeta) + q$ with $\mathcal{L} : \mathbb{V} \to \mathbb{V}$ being a linear transformation and $q \in \mathbb{V}$, the problem (1) reduces to the linear complementarity problem over symmetric cones (LSCCP):

$$\zeta \in \mathcal{K}, \quad \mathcal{L}(\zeta) + q \in \mathcal{K}, \quad \langle \zeta, \mathcal{L}(\zeta) + q\rangle = 0. \tag{2}$$

Recently, there is active research for the solution of the symmetric cone optimization and complementarity problems, and have been proposed various solution methods. They include the interior-point methods [3, 22, 29], the merit function methods [14, 16], the regularized smoothing method [13], and the smoothing Newton method [8]. This paper is concerned with a derivative-free method based on the implicit Lagrangian reformulation (5) of the SCCP (1). An attractive feature of this method is that no derivatives of $F(\cdot)$ need to be computed, which makes the method suitable for large-scale problems, as well as for applications where the derivatives of $F(\cdot)$ are not available or are costly to compute.

The implicit Lagrangian function was first introduced by Mangasarian and Solodov [17] as a smooth merit function for the NCP, and further studied under the setting of nonnegative orthant cones by [10, 15, 18, 25, 27] and other literature. Recently, Kong, Tuncel and Xiu [14] utilized the Jordan-algebraic technique to extend the implicit Lagrangian to the symmetric cone $\mathcal{K}$. The corresponding function is defined as

$$\psi_\alpha(x,y) := \langle x,y\rangle + \frac{1}{2\alpha}\Big\{\|(x-\alpha y)_+\|^2 - \|x\|^2 + \|(y-\alpha x)_+\|^2 - \|y\|^2\Big\}, \quad \forall x,y \in \mathbb{V} \tag{3}$$

where $\alpha > 1$ is a parameter, $\|\cdot\|$ is the norm induced by the inner product $\langle\cdot,\cdot\rangle$, and $(\cdot)_+$ denotes the metric projection onto the symmetric cone $\mathcal{K}$. They have showed that ψ_α is a continuously differentiable merit function associated with

$\mathcal{K}$, that is,

$$\psi_\alpha(x,y)=0 \iff x\in\mathcal{K}, \quad y\in\mathcal{K}, \quad \langle x,y\rangle=0, \tag{4}$$

and thus the SCCP can be formulated as an unconstrained smooth minimization problem

$$\min_{\zeta\in\mathbb{V}} \ \Psi_\alpha(\zeta):=\psi_\alpha(\zeta,F(\zeta)) \tag{5}$$

in the sense that the minimizer of (5) with zero objective value is a solution of (1). For the unconstrained reformulation, they particularly gave a sufficient and necessary condition for each stationary point of Ψ_α to be a solution of (1), and established that Ψ_α offers a global error bound for the SCCP (1) when F has the uniform Cartesian P-property.

Although the literature on derivative-free algorithms for the NCPs is vast (see, e.g., [1, 7, 9, 12, 18, 26, 27]), to our best knowledge, there are few papers to consider the ones for the nonpolyhedral symmetric cone complementarity problems except [20, 28]. In these two papers, the derivative-free methods are developed for the SDCP and the SOCCP, respectively, by the Fischer-Burmeister type merit function. Moreover, the rate of convergence result is not established in [28] and the one in [20] is only shown to be Q-linear. We also note that almost all derivative-free methods mentioned above are descent ones with the monotone Armijo-type line search.

To the contrast, in this work we develop a nonmonotone derivative-free method for the SCCP by using the vector of the form $d(\zeta)\equiv-\theta\nabla_x\psi_\alpha(\zeta,F(\zeta))-(1-\theta)\nabla_y\psi_\alpha(\zeta,F(\zeta))$ with $\theta\in[0,1]$ as the search direction. As shown in Prop.3.4, when θ is sufficiently small, $d(\zeta)$ is a descent direction, and it reduces to the one adopted in [18]. However, for a general $\theta\in[0,1]$, $d(\zeta)$ is not necessarily descent, and we adopt a nonmonotone line search rule to seek a desirable stepsize. We show that the method converges in terms of the implicit Lagrangian value for a large class of SCCPs, and if θ is restricted to be less than a threshold $\bar\theta\in(0,1)$ and the SCCP is strongly monotone, the sequence generated converges to the solution at a R-linear rate. Numerical tests verify the theoretical results, and show that the method with a smaller θ does not have better performance than the method with a θ close to 1, though it may have a R-linear rate of convergence when θ is sufficiently small.

Throughout this paper, **int** $\mathcal{K}$ denotes the interior of the cone $\mathcal{K}$ and $\|\cdot\|$ represents the norm induced by the inner product $\langle\cdot,\cdot\rangle$, i.e., $\|\cdot\|:=\sqrt{\langle\cdot,\cdot\rangle}$. For any $x\in\mathbb{V}$, we write $(x)_+$ and $(x)_-$ as the metric projection of x onto $\mathcal{K}$ and $-\mathcal{K}$, respectively, i.e.,

$$(x)_+:=\mathrm{argmin}_{y\in\mathcal{K}}\{\|x-y\|\}.$$

For a differentiable mapping $F : \mathbb{V} \to \mathbb{V}$, we denote its transposed Jacobian at $x \in \mathbb{V}$ by $\nabla F(x)$. Unless otherwise stated, the parameter α in the sequel always satisfies $\alpha > 1$.

2 Preliminaries

A *Euclidean Jordan algebra* is a triple $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle_{\mathbb{V}})$, where $(\mathbb{V}, \langle\cdot,\cdot\rangle_{\mathbb{V}})$ is a finite-dimensional inner product space over $\mathbb{R}$ and $(x, y) \mapsto x \circ y : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ is a bilinear mapping satisfying:

(i) $x \circ y = y \circ x$ for all $x, y \in \mathbb{V}$;

(ii) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ for all $x, y \in \mathbb{V}$, where $x^2 := x \circ x$;

(iii) $\langle x \circ y, z\rangle_{\mathbb{V}} = \langle y, x \circ z\rangle_{\mathbb{V}}$ for all $x, y, z \in \mathbb{V}$.

We call $x \circ y$ the *Jordan product* of x and y. We assume that there is an element $e \in \mathbb{V}$ such that $x \circ e = x$ for all $x \in \mathbb{V}$, and call such e the *unit element*. Let

$$\zeta(x) := \min\left\{k : \{e, x, x^2, \ldots, x^k\} \text{ are linearly dependent}\right\}.$$

Since $\zeta(x)$ is bounded by the dimension of $\mathbb{V}$, denoted by $\dim(\mathbb{V})$, the *rank* of $(\mathbb{V}, \circ)$ is well defined by $r := \max\{\zeta(x) : x \in \mathbb{V}\}$. Define *the set of squares* as $\mathcal{K} := \left\{x^2 : x \in \mathbb{V}\right\}$. Then, from [11, Theorem III.2.1], it follows that $\mathcal{K}$ is a symmetric cone. This means that $\mathcal{K}$ is a self-dual closed convex cone with nonempty interior $\textbf{int}\ \mathcal{K}$, and for any $x, y \in \textbf{int}\ \mathcal{K}$, there exists an invertible linear transformation $\mathcal{T} : \mathbb{V} \to \mathbb{V}$ such that $\mathcal{T}(\mathcal{K}) = \mathcal{K}$.

Recall that an element $c \in \mathbb{V}$ is *idempotent* if $c^2 = c$, and two idempotents c and d are *orthogonal* if $c \circ d = 0$. A nonzero idempotent is *primitive* if it cannot be written as the sum of two other nonzero idempotents. A complete system of orthogonal idempotents is a finite set $\{c_1, c_2, \ldots, c_k\}$ of idempotents with $c_i \circ c_j = 0$ $(i \neq j)$ and $\sum_{i=1}^k c_i = e$. We call a complete system of orthogonal primitive idempotents a *Jordan frame*.

Theorem 2.1 *[11, Theorem III.1.2] Suppose that* $\mathbb{A} = (\mathbb{V}, \circ, \langle\cdot,\cdot\rangle_{\mathbb{V}})$ *is a Euclidean Jordan algebra with rank* r*. Then for each* $x \in \mathbb{V}$*, there exist a Jordan frame* $\{c_1, c_2, \ldots, c_r\}$ *and real numbers* $\lambda_1(x), \lambda_2(x), \ldots, \lambda_r(x)$ *such that*

$$x = \lambda_1(x)c_1 + \lambda_2(x)c_2 + \cdots + \lambda_r(x)c_r.$$

The numbers $\lambda_1(x), \ldots, \lambda_r(x)$ *(counting multiplicities) are called the eigenvalues of* x*. Furthermore, the trace of* x*, denoted by* $\mathrm{tr}(x)$*, is defined as* $\mathrm{tr}(x) = \sum_{j=1}^r \lambda_j(x)$*.*

Since, by [11, Prop.III.1.5], a Jordan algebra $\mathbb{A} = (\mathbb{V}, \circ)$ over $\mathbb{R}$ with a unit element $e \in \mathbb{V}$ is Euclidean if and only if the symmetric bilinear form $\operatorname{tr}(x \circ y)$ is positive definite, we may define another inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{V}$ by

$$\langle x, y \rangle \ := \ \operatorname{tr}(x \circ y), \quad \forall\, x, y \in \mathbb{V}. \tag{6}$$

By the associativity of $\operatorname{tr}(\cdot)$ (see [11, Prop.II.4.3]), the inner product $\langle \cdot, \cdot \rangle$ is associative, i.e., for all $x, y, z \in \mathbb{V}$, it holds that $\langle x \circ y, z \rangle = \langle y, x \circ z \rangle$.

Unless otherwise stated, in the rest of this paper, we always assume that $\mathbb{A} = (\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ is a Euclidean Jordan algebra of rank r and $\dim(\mathbb{V}) = n$ with $\langle \cdot, \cdot \rangle$ defined as in (6).

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a scalar-valued function. By Theorem 2.1, it is natural to define a vector-valued function associated with the Euclidean Jordan algebra $\mathbb{A} = (\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ by

$$\varphi_{\mathbb{V}}(x) := \varphi(\lambda_1(x))c_1 + \varphi(\lambda_2(x))c_2 + \cdots + \varphi(\lambda_r(x))c_r, \tag{7}$$

where $x \in \mathbb{V}$ has the spectral decomposition $x = \sum_{j=1}^{r} \lambda_j(x)c_j$. The function $\varphi_{\mathbb{V}}$ is also called *Löwner operator* in [23] and shown to inherit many properties from φ. Especially, when $\varphi(t)$ is chosen as $\max\{0, t\}$ and $\min\{0, t\}$ for $t \in \mathbb{R}$, the Löwner operator $\varphi_{\mathbb{V}}(\cdot)$ respectively becomes the metric projection operator onto $\mathcal{K}$ and $-\mathcal{K}$:

$$(x)_+ := \sum_{j=1}^{r} \max\big\{0, \lambda_j(x)\big\}c_j \quad \text{and} \quad (x)_- := \sum_{j=1}^{r} \min\big\{0, \lambda_j(x)\big\}c_j. \tag{8}$$

A Euclidean Jordan algebra is called *simple* if it is not the direct sum of two Euclidean Jordan algebras. By [11, Prop.III.4.4-4.5 & Theorem V.3.7], each Euclidean Jordan algebra is, in a unique way, a direct sum of simple Euclidean Jordan algebras. Also, the symmetric cone in a given Euclidean Jordan algebra is, in a unique way, a direct sum of symmetric cones in the constituent simple Euclidean Jordan algebras. In the sequel, we assume that $\mathbb{V} \equiv \mathbb{V}_1 \times \cdots \times \mathbb{V}_m$ and $\mathcal{K} \equiv \mathcal{K}^1 \times \cdots \times \mathcal{K}^m$, where each $\mathbb{A}_i = (\mathbb{V}_i, \circ, \langle \cdot, \cdot \rangle)$ is a simple Euclidean Jordan algebra and $\mathcal{K}^i$ is a symmetric cone in $\mathbb{V}_i$. Corresponding to the Cartesian structure of $\mathbb{V}$ and $\mathcal{K}$, let $\zeta = (\zeta_1, \ldots, \zeta_m)$ with $\zeta_i \in \mathbb{V}_i$ and $F(\zeta) = (F_1(\zeta), \ldots, F_m(\zeta))$ with $F_i : \mathbb{V} \to \mathbb{V}_i$.

To close this section, we recall the definitions of uniform Cartesian P-property [5, 14] and uniform Jordan P-property [24].

Definition 2.1 *The mapping $F = (F_1, F_2, \ldots, F_m)$ with $F_i : \mathbb{V} \to \mathbb{V}_i$ is said to have*

(a) *the uniform Cartesian P-property if there exists a positive scalar ρ such that for any $\zeta, \xi \in \mathbb{V}$, there is an index $i \in \{1, 2, \ldots, m\}$ such that*

$$\langle \zeta_i - \xi_i,\ F_i(\zeta) - F_i(\xi) \rangle \geq \rho \|\zeta - \xi\|^2.$$

(b) *the uniform Jordan P-property if there is a scalar $\rho > 0$ such that for any $\zeta, \xi \in \mathbb{V}$,*

$$\lambda_{\max}\left[(\zeta - \xi) \circ (F(\zeta) - F(\xi)) \right] \geq \rho \|\zeta - \xi\|^2$$

where $\lambda_{\max}(x)$ denotes the largest eigenvalue of a vector $x \in \mathbb{V}$.

3 Properties of the function Ψ_α

This section is devoted to the favorable properties of the implicit Lagrangian function Ψ_α. Most of the properties have been given by Kong, Tuncel and Xiu [14], and we supplement some ones that play an important role in the convergence analysis of the algorithms. For this purpose, let $r_\alpha : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ and $R_\alpha : \mathbb{V} \to \mathbb{V}$ be respectively defined by

$$r_\alpha(x, y) \ := \ x - (x - \alpha y)_+ \quad \text{for} \ \ \alpha > 0 \tag{9}$$

and

$$R_\alpha(\zeta) \ := \ r_\alpha(\zeta, F(\zeta)) \quad \text{for} \ \ \alpha > 0. \tag{10}$$

Then, by using the same arguments as those of [4, Lemma 1] and [21, Theorem 4.2], it is easy to obtain the following properties of r_α, and we omit the proof for simplicity.

Lemma 3.1 *Let r_α be defined as in (9). Then, there hold that*

(a) $\min\{1, \alpha\}\|r_1(x, y)\| \leq \|r_\alpha(x, y)\| \leq \max\{1, \alpha\}\|r_1(x, y)\|$ *for all $x, y \in \mathbb{V}$ and $\alpha > 0$;*

(b) $\min\{1, \alpha\}\|r_1(x, y)\| \leq \|r_\alpha(y, x)\| \leq \max\{1, \alpha\}\|r_1(x, y)\|$ *for all $x, y \in \mathbb{V}$ and $\alpha > 0$;*

(c) $\alpha^{-1}(\alpha - 1)\|r_1(x, y)\|^2 \leq \psi_\alpha(x, y) \leq (\alpha - 1)\|r_1(x, y)\|^2$ *for all $x, y \in \mathbb{V}$ and $\alpha > 1$.*

Now we give a proposition to summarize the favorable properties of the function Ψ_α.

Proposition 3.1 *Let Ψ_α be defined as in (5). Then the following statements hold:*

(a) $\Psi_\alpha(\zeta) \geq 0$ *for all* $\zeta \in \mathbb{V}$*, and* $\Psi_\alpha(\zeta) = 0$ *if and only if* $\zeta \in \mathbb{V}$ *solves the SCCP (1).*

(b) Ψ_α *is continuously differentiable everywhere on* $\mathbb{V}$ *with the gradient given by*

$$\nabla\Psi_\alpha(\zeta) = \nabla_x\psi_\alpha(\zeta, F(\zeta)) + \nabla F(\zeta)\nabla_y\psi_\alpha(\zeta, F(\zeta)).$$

(c) $\nabla_x\psi_\alpha(\zeta, F(\zeta)) + \nabla_y\psi_\alpha(\zeta, F(\zeta)) = 0$ *if and only if* $\zeta \in \mathbb{V}$ *solves the SCCP (1).*

(d) $\big\langle \nabla_x\psi_\alpha(\zeta, F(\zeta)), \nabla_y\psi_\alpha(\zeta, F(\zeta))\big\rangle \geq 0$ *for any* $\zeta \in \mathbb{V}$.

(e) $\big\|\nabla_x\psi_\alpha(\zeta, F(\zeta)) + \nabla_y\psi_\alpha(\zeta, F(\zeta))\big\|^2 \geq \dfrac{(\alpha^2-1)^2}{\alpha^2(\alpha^2+1)}\|R_\alpha(\zeta)\|^2$ *for any* $\zeta \in \mathbb{V}$.

(f) $\big\|\nabla_x\psi_\alpha(\zeta, F(\zeta)) + \nabla_y\psi_\alpha(\zeta, F(\zeta))\big\|^2 \leq 2\alpha(\alpha-1)\Psi_\alpha(\zeta)$ *for any* $\zeta \in \mathbb{V}$.

Proof. The proof of parts (a)–(d) can be found in the literature [14]. To prove parts (e) and (f), we only need to show that for all $x, y \in \mathbb{V}$,

$$\begin{aligned} \big\|\nabla_x\psi_\alpha(x,y) + \nabla_y\psi_\alpha(x,y)\big\|^2 &\geq \frac{(\alpha^2-1)^2}{\alpha^2(\alpha^2+1)}\left[\|r_\alpha(x,y)\|^2 + \|r_\alpha(y,x)\|^2\right], &(11)\\ \big\|\nabla_x\psi_\alpha(x,y) + \nabla_y\psi_\alpha(x,y)\big\|^2 &\leq 2\alpha(\alpha-1)\psi_\alpha(x,y). &(12)\end{aligned}$$

From [14], it follows that for any $x, y \in \mathbb{V}$,

$$\begin{aligned} \nabla_x\psi_\alpha(x,y) &= y + \alpha^{-1}\left[(x-\alpha y)_+ - x - \alpha(y-\alpha x)_+\right], \\ \nabla_y\psi_\alpha(x,y) &= x + \alpha^{-1}\left[(y-\alpha x)_+ - y - \alpha(x-\alpha y)_+\right]. \qquad (13)\end{aligned}$$

Therefore, we have

$$\begin{aligned} \big\|\nabla_x\psi_\alpha(x,y) + \nabla_y\psi_\alpha(x,y)\big\|^2 &= \frac{(\alpha-1)^2}{\alpha^2}\big\|\left[x-(x-\alpha y)_+\right] + \left[y-(y-\alpha x)_+\right]\big\|^2 \\ &= \frac{(\alpha-1)^2}{\alpha^2}\left[\big\|x-(x-\alpha y)_+\big\|^2 + \big\|y-(y-\alpha x)_+\big\|^2\right] \\ &\quad + \frac{2(\alpha-1)^2}{\alpha^2}\big\langle x-(x-\alpha y)_+,\ y-(y-\alpha x)_+\big\rangle. \qquad (14)\end{aligned}$$

From part (d), we see that $\langle \nabla_x\psi_\alpha(x,y), \nabla_y\psi_\alpha(x,y)\rangle \geq 0$ for any $x, y \in \mathbb{V}$, that is,

$$\begin{aligned} 0 &\leq \left\langle y-(y-\alpha x)_+ + \frac{1}{\alpha}[(x-\alpha y)_+ - x],\ x-(x-\alpha y)_+ + \frac{1}{\alpha}[(y-\alpha x)_+ - y]\right\rangle \\ &= -\frac{1}{\alpha}\big\|x-(x-\alpha y)_+\big\|^2 - \frac{1}{\alpha}\big\|y-(y-\alpha x)_+\big\|^2 \\ &\quad + \left(1+\frac{1}{\alpha^2}\right)\big\langle x-(x-\alpha y)_+,\ y-(y-\alpha x)_+\big\rangle.\end{aligned}$$

This in turn implies that

$$\begin{aligned}&\left\langle x-(x-\alpha y)_+, y-(y-\alpha x)_+\right\rangle \\ \geq\ & \frac{\alpha}{\alpha^2+1}\left[\|x-(x-\alpha y)_+\|^2+\|y-(y-\alpha x)_+\|^2\right] \quad \forall\, x,y\in\mathbb{V}. \quad (15)\end{aligned}$$

Combining (14) with (15) and noting that $\alpha>0$, we immediately obtain

$$\begin{aligned}&\left\|\nabla_x\psi_\alpha(x,y)+\nabla_y\psi_\alpha(x,y)\right\|^2 \\ \geq\ & \frac{(\alpha-1)^2(\alpha+1)^2}{\alpha^2(\alpha^2+1)}\left[\|x-(x-\alpha y)_+\|^2+\|y-(y-\alpha x)_+\|^2\right] \\ =\ & \frac{(\alpha^2-1)^2}{\alpha^2(\alpha^2+1)}\left[\|r_\alpha(x,y)\|^2+\|r_\alpha(y,x)\|^2\right].\end{aligned}$$

This completes the proof of (11). To see inequality (12), we verify the following:

$$\begin{aligned}\left\|\nabla_x\psi_\alpha(x,y)+\nabla_y\psi_\alpha(x,y)\right\|^2 &= \frac{(\alpha-1)^2}{\alpha^2}\left\|r_\alpha(x,y)+r_\alpha(y,x)\right\|^2 \\ &\leq 2\frac{(\alpha-1)^2}{\alpha^2}\left(\|r_\alpha(x,y)\|^2+\|r_\alpha(y,x)\|^2\right) \\ &\leq \frac{(\alpha-1)^2}{\alpha^2}\cdot 2\alpha^2\|r_1(x,y)\|^2 \\ &\leq 2(\alpha-1)^2\cdot\frac{\alpha}{\alpha-1}\psi_\alpha(x,y) \\ &= 2\alpha(\alpha-1)\psi_\alpha(x,y),\end{aligned}$$

where the first equality is due to the first equation of (14) and the definition of r_α, and the second inequality holds by Lemma 3.1(a)-(b). Thus, we complete the proof. □

The assertions of Prop.3.1(e)-(f) are new, and they play a key role in establishing the rate of convergence result of the nonmonotone descent algorithm of this paper. When $\mathbb{V}$ reduces to the Euclidean space $\mathbb{R}^n$ with the standard inner product and Jordan product defined as the componentwise product of the vectors, Prop.3.1(e) implies the second result of [18, Lemma 1] by observing $\alpha>1$ and the following inequalities

$$\begin{aligned}\left\|\nabla_x\psi_\alpha(\zeta,F(\zeta))+\nabla_y\psi_\alpha(\zeta,F(\zeta))\right\| &\geq \frac{\alpha^2-1}{\alpha\sqrt{\alpha^2+1}}\|R_\alpha(\zeta)\| \\ &\geq \frac{(\alpha-1)(\alpha+1)}{\alpha\sqrt{\alpha^2+1}}\|R_1(\zeta)\| \\ &\geq \frac{\alpha-1}{\alpha}\|R_1(\zeta)\|\end{aligned}$$

where the first inequality is by Lemma 3.1(b) and the second one is due to $\alpha+1>\sqrt{\alpha^2+1}$.

The following results for Ψ_α can be found in [14, Corollary 6.4] and [14, Theorem 6.3].

Proposition 3.2 *Assume that F has the uniform Cartesian P-property. Then,*

(a) *Each stationary point of Ψ_α is a solution of the SCCP (1).*

(b) *If, in addition, F is Lipschitz continuous with constant $L > 0$, then for any $\zeta \in \mathbb{V}$,*

$$\frac{1}{(\alpha-1)(2+L)^2}\Psi_\alpha(\zeta) \leq \|\zeta - \zeta^*\|^2 \leq \frac{\alpha(1+L)^2}{(\alpha-1)\rho^2}\Psi_\alpha(\zeta),$$

where ζ^* be the unique solution of (1), and the constant ρ is same as in Def.2.1.

It is well known that the coerciveness of the merit function plays an important role in the convergence analysis of the unconstrained reformulation methods for the complementarity problems. The next proposition presents a mild condition to guarantee the coerciveness of Ψ_α, whose proof can be found in [19, Theorem 4.1].

Proposition 3.3 *The function Ψ_α is coercive under the following condition that*

(C.1) *F has the uniform Jordan P-property and the linear growth, i.e., there exists a constant $C > 0$ such that for any $\zeta \in \mathbb{V}$, $\|F(\zeta)\| \leq \|F(0)\| + C\|\zeta\|$.*

Particularly, if F is given as in (2) with $\mathcal{L}$ having the P-property, then Ψ_α is coercive.

To close this section, we present the direction $d(\zeta)$ that will be employed to design our derivative-free algorithm. Specifically, let the mapping $d : \mathbb{V} \to \mathbb{V}$ be given by

$$d(\zeta) := -\theta\nabla_x\psi_\alpha(\zeta, F(\zeta)) - (1-\theta)\nabla_y\psi_\alpha(\zeta, F(\zeta)) \quad \forall\theta \in [0,1]. \qquad (16)$$

Such vector d enjoys the properties stated as in the following proposition.

Proposition 3.4 *Suppose that ∇F is positive definite. Then, for sufficiently small $\theta > 0$,*

$$d(\zeta)^T\nabla\Psi_\alpha(\zeta) < 0 \quad \text{when} \quad d(\zeta) \neq 0.$$

If F is strongly monotone with modulus $\mu > 0$ and $S \subseteq \mathbb{V}$ is any bounded set, then there exists $\bar{\theta} \in (0,1)$ such that for all $\theta \leq \bar{\theta}$,

$$d(\zeta)^T\nabla\Psi_\alpha(\zeta) \leq -\frac{1}{2}\theta\big\|\nabla_x\psi_\alpha(\zeta, F(\zeta)) + \nabla_y\psi_\alpha(\zeta, F(\zeta))\big\|^2 \quad \forall\zeta \in S.$$

Proof. By the formula of $\nabla\Psi_\alpha$ and Prop.3.1(d), for any $\theta \in [0,1]$, we have

$$\begin{aligned} d(\zeta)^T\nabla\Psi_\alpha(\zeta) &= -\theta\|\nabla_x\psi_\alpha(\zeta,F(\zeta))\|^2 - (1-\theta)\langle\nabla_x\psi_\alpha(\zeta,F(\zeta)),\nabla_y\psi_\alpha(\zeta,F(\zeta))\rangle \\ &\quad -\theta\,\langle\nabla_x\psi_\alpha(\zeta,F(\zeta)),\ \nabla F(\zeta)\nabla_y\psi_\alpha(\zeta,F(\zeta))\rangle \\ &\quad -(1-\theta)\,\langle\nabla_y\psi_\alpha(\zeta,F(\zeta)),\ \nabla F(\zeta)\nabla_y\psi_\alpha(\zeta,F(\zeta))\rangle \\ &\le -\theta\|\nabla_x\psi_\alpha(\zeta,F(\zeta))\|^2 - \theta\,\langle\nabla_x\psi_\alpha(\zeta,F(\zeta)),\nabla F(\zeta)\nabla_y\psi_\alpha(\zeta,F(\zeta))\rangle \\ &\quad -(1-\theta)\,\langle\nabla_y\psi_\alpha(\zeta,F(\zeta)),\nabla F(\zeta)\nabla_y\psi_\alpha(\zeta,F(\zeta))\rangle\,. \end{aligned} \tag{17}$$

Notice that for sufficiently small $\theta > 0$ and any given $\zeta \in \mathbb{V}$, the vector $d(\zeta) \neq 0$ must imply that $\nabla_y\psi_\alpha(\zeta,F(\zeta)) \neq 0$. Thus, the last term of the right hand side is always strictly negative by the positive definiteness of ∇F, whereas the first two terms are sufficiently small. Therefore, we obtain that $d(\zeta)^T\nabla\Psi_\alpha(\zeta) < 0$ whenever $d(\zeta) \neq 0$.

Since ∇F is continuous and S is bounded, there exists a constant $\nu > 0$ such that

$$\|\nabla F(\zeta)\| \le \nu \quad \forall\, \zeta \in S. \tag{18}$$

On the other hand, using the strong monotonicity of F, we have

$$\langle\nabla F(\zeta)u,u\rangle \ge \mu\|u\|^2 \quad \forall\, \zeta,u \in \mathbb{V}. \tag{19}$$

Now, from equations (17)–(19), it follows that for any $\theta \in [0,1]$ and $\zeta \in S$,

$$\begin{aligned} d(\zeta)^T\nabla\Psi_\alpha(\zeta) &\le -\theta\|\nabla_x\psi_\alpha(\zeta,F(\zeta))\|^2 - (1-\theta)\mu\|\nabla_y\psi_\alpha(\zeta,F(\zeta))\|^2 \\ &\quad +\theta\nu\|\nabla_x\psi_\alpha(\zeta,F(\zeta))\|\cdot\|\nabla_y\psi_\alpha(\zeta,F(\zeta))\| \\ &= -\frac{1}{2}\theta\Big(\|\nabla_x\psi_\alpha(\zeta,F(\zeta))\| + \|\nabla_y\psi_\alpha(\zeta,F(\zeta))\|\Big)^2 \\ &\quad -\frac{1}{2}\theta\big\|\nabla_x\psi_\alpha(\zeta,F(\zeta))\big\|^2 - \frac{2(1-\theta)\mu-\theta}{2}\big\|\nabla_y\psi_\alpha(\zeta,F(\zeta))\big\|^2 \\ &\quad +\theta(\nu+1)\|\nabla_x\psi_\alpha(\zeta,F(\zeta))\|\cdot\|\nabla_y\psi_\alpha(\zeta,F(\zeta))\|. \end{aligned} \tag{20}$$

If $\theta \le 2\mu/(2\mu+1)$, then the last inequality can be rewritten as

$$\begin{aligned} d(\zeta)^T\nabla\Psi_\alpha(\zeta) &\le -\frac{1}{2}\theta\Big(\|\nabla_x\psi_\alpha(\zeta,F(\zeta))\| + \|\nabla_y\psi_\alpha(\zeta,F(\zeta))\|\Big)^2 \\ &\quad -\left(\sqrt{\frac{\theta}{2}}\big\|\nabla_x\psi_\alpha(\zeta,F(\zeta))\big\| - \sqrt{\frac{2\mu-(2\mu+1)\theta}{2}}\big\|\nabla_y\psi_\alpha(\zeta,F(\zeta))\big\|\right)^2 \quad (21) \\ &\quad +\Big(\theta(\nu+1) - \sqrt{2\mu\theta-(2\mu+1)\theta^2}\Big)\,\|\nabla_x\psi_\alpha(\zeta,F(\zeta))\|\|\nabla_y\psi_\alpha(\zeta,F(\zeta))\|. \end{aligned}$$

If $\theta(\nu+1) \le \sqrt{2\mu\theta-(2\mu+1)\theta^2}$, that is, $\theta \le 2\mu/(2\mu+1+(\nu+1)^2)$, then using (21) and the Cauchy-Schwartz inequality yields

$$\begin{aligned} d(\zeta)^T\nabla\Psi_\alpha(\zeta) &\le -\frac{1}{2}\theta\Big(\|\nabla_x\psi_\alpha(\zeta,F(\zeta))\| + \|\nabla_y\psi_\alpha(\zeta,F(\zeta))\|\Big)^2 \\ &\le -\frac{1}{2}\theta\big\|\nabla_x\psi_\alpha(\zeta,F(\zeta)) + \nabla_y\psi_\alpha(\zeta,F(\zeta))\big\|^2. \end{aligned}$$

Thus, by setting

$$\bar{\theta} := \min\left\{\frac{2\mu}{2\mu+1}, \frac{2\mu}{2\mu+1+(\nu+1)^2}\right\} = \frac{2\mu}{2\mu+1+(\nu+1)^2}, \tag{22}$$

we obtain the desired result. The proof is complete. □

4 Nonmonotone derivative-free algorithm

In this section, we utilize the direction $d(\zeta)$ defined by (16) to design a derivative-free algorithm. By Prop.3.4, $d(\zeta)$ with $\theta \in [0,1]$ may not satisfy the descent condition. Moreover, the technique of nonmonotone line search is often more effective than the Armijo-type line search. So, we adopt a nonmonotone line search rule to seek a suitable stepsize.

Algorithm 4.1

(Step 0) Choose $\zeta^0 \in \mathbb{V}$, $\epsilon \geq 0$, $\theta \in [0,1]$ and $\gamma, \delta \in (0,1)$. Let $M > 0$ be an integer. Set $k := 0$.

(Step 1) If $\Psi_\alpha(\zeta^k) \leq \epsilon$, then stop. Otherwise, go to Step 2.

(Step 2) Let $m(0) = 0$, $0 \leq m(k) \leq \min\{m(k-1)+1, M-1\}$ for $k \geq 1$. Let l_k be the smallest nonnegative integer l satisfying

$$\Psi_\alpha(\zeta^k + \gamma^l d^k) \leq \max_{0\leq j\leq m(k)} \Psi_\alpha(\zeta^{k-j}) - \delta\gamma^{2l} h(\zeta^k), \tag{23}$$

where $d^k := d(\zeta^k)$ with $d(\zeta)$ defined as in (16), and

$$h(\zeta) := \|\nabla_x \psi_\alpha(\zeta, F(\zeta)) + \nabla_y \psi_\alpha(\zeta, F(\zeta))\|^2. \tag{24}$$

(Step 3) Set $\zeta^{k+1} := \zeta^k + \gamma^{l_k} d^k$ and $k := k+1$, and then go to Step 1.

Observe that no derivatives of F are needed to compute the search direction or the stepsize in Algorithm 4.1. Hence, Algorithm 4.1 requires little computation and storing work at each iteration. Since θ is any fixed constant in $[0,1]$, the direction d^k is different from the one used in [18] and at each iteration may not satisfy the descent condition $(d^k)^T \nabla \Psi_\alpha(\zeta^k) < 0$. Based on this, a nonmonotone line search rule is used in Step 2. The line search rule is different from the ones adopted in [2, 6] where the gradient of the merit (or objective) function is needed, and when $m(k) \equiv 0$, the nonmonotone line search reduces to the Armijo line search. Particularly, if θ is restricted to be less than $\bar{\theta}$ given by (22) and F is strongly monotone, then Prop.3.4 implies that Algorithm 4.1

will become a nonmonotone derivative-free descent algorithm.

In what follows, we study the convergence of Algorithm 4.1. To the end, assume that Algorithm 4.1 generates an infinite sequence $\{\zeta^k\}$, i.e., $\epsilon = 0$. We define the level set

$$\mathcal{L}(\Psi_\alpha, \zeta^0) := \big\{\zeta \in \mathbb{V} \mid \Psi_\alpha(\zeta) \le \Psi_\alpha(\zeta^0)\big\}.$$

Then $\mathcal{L}(\Psi_\alpha, \zeta^0)$ is bounded under one of the condition given in Prop.3.3. By the continuity of $F(\cdot)$, we know that $\mathcal{D}(\zeta^0) := \sup\big\{\|d(\zeta)\| \mid \zeta \in \mathcal{L}(\Psi_\alpha, \zeta^0)\big\}$ is finite. Consequently,

$$\mathcal{B}(\zeta^0) := \mathcal{L}(\Psi_\alpha, \zeta^0) + \big\{\zeta \in \mathbb{V} \mid \|\zeta\| \le \mathcal{D}(\zeta^0)\big\}$$

is also bounded under the condition stated in Prop.3.3.

Lemma 4.1 *Let $\{\zeta^k\}$ be the sequence generated by Algorithm 4.1. Then,*

(a) *the sequence $\{\zeta^k\}$ is contained in $\mathcal{L}(\Psi_\alpha, \zeta^0)$;*

(b) $\max\limits_{1\le i\le M} \Psi_\alpha(\zeta^{Mp+i}) \le \max\limits_{1\le i\le M} \Psi_\alpha(\zeta^{M(p-1)+i}) - \delta \min\limits_{0\le i\le M-1} \gamma^{2l_{(Mp+i)}} h(\zeta^{Mp+i})$ *for any $p \ge 1$.*

Proof. (a) For each $k \ge 0$, let $\sigma(k)$ be an integer from $[k - m(k), k]$ such that

$$\Psi_\alpha(\zeta^{\sigma(k)}) = \max_{0\le j\le m(k)} \Psi_\alpha(\zeta^{k-j}).$$

Then, the line search condition (23) can be rewritten as

$$\Psi_\alpha(\zeta^{k+1}) \le \Psi_\alpha(\zeta^{\sigma(k)}) - \delta\gamma^{2l_k} h(\zeta^k). \tag{25}$$

Noting that $m(k+1) \le m(k) + 1$ and $h(\zeta) \ge 0$ for any $\zeta \in \mathbb{V}$, we have from (23) that

$$\begin{aligned}\Psi_\alpha(\zeta^{\sigma(k+1)}) = \max_{0\le j\le m(k+1)} \Psi_\alpha(\zeta^{k+1-j}) &\le \max_{0\le j\le m(k)+1} \Psi_\alpha(\zeta^{k+1-j}) \\ &= \max\{\Psi_\alpha(\zeta^{\sigma(k)}), \Psi_\alpha(\zeta^{k+1})\} \\ &= \Psi_\alpha(\zeta^{\sigma(k)}),\end{aligned}$$

where the last equality is from (25) and the nonnegativity of $h(\zeta^k)$. This shows that the sequence $\{\Psi_\alpha(\zeta^{\sigma(k)})\}$ is nonincreasing. Noting that $\zeta^{\sigma(0)} = \zeta^0$, we then have $\Psi_\alpha(\zeta^k) \le \Psi_\alpha(\zeta^0)$ for all k, which in turn implies $\{\zeta^k\} \subseteq \mathcal{L}(\Psi_\alpha, \zeta^0)$.

(b) We only need to show that the following inequality holds for $j = 1, 2, \ldots, M$:

$$\Psi_\alpha(\zeta^{Mp+j}) \le \max_{1\le i\le M} \Psi_\alpha(\zeta^{M(p-1)+i}) - \delta\gamma^{2l_{(Mp+j-1)}} h(\zeta^{Mp+j-1}) \quad \forall p \ge 1. \tag{26}$$

Notice that the linear search condition (23) implies

$$\Psi_\alpha(\zeta^{Mp+1}) \le \max_{0\le i\le m(Mp)} \Psi_\alpha(\zeta^{Mp-i}) - \delta\gamma^{2l_{Mp}} h(\zeta^{Mp}),$$

which together with $m(Mp) \le M-1$ shows that inequality (26) holds for $j = 1$. Suppose that (26) holds for any $1 \le j \le M-1$. Then, from the nonnegativity of $h(\zeta)$, it follows that

$$\max_{1\le i\le j} \Psi_\alpha(\zeta^{Mp+i}) \le \max_{1\le i\le M} \Psi_\alpha(\zeta^{M(p-1)+i}).$$

Consequently, by using (23), the induction hypothesis and $m(Mp+j) \le M-1$, we get

$$\begin{aligned}
\Psi_\alpha(\zeta^{Mp+j+1}) &\le \max_{0\le i\le m(Mp+j)} \Psi_\alpha(\zeta^{Mp+j-i}) - \delta\gamma^{2l_{(Mp+j)}} h(\zeta^{Mp+j}) \\
&\le \max\left\{ \max_{1\le i\le M} \Psi_\alpha(\zeta^{M(p-1)+i}), \max_{1\le i\le j} \Psi_\alpha(\zeta^{Mp+j}) \right\} - \delta\gamma^{2l_{(Mp+j)}} h(\zeta^{Mp+j}) \\
&\le \max_{1\le i\le M} \Psi_\alpha(\zeta^{M(p-1)+i}) - \delta\gamma^{2l_{(Mp+j)}} h(\zeta^{Mp+j}).
\end{aligned}$$

This shows that (26) also holds for $j+1$. By induction, we prove that (26) is true for all $1 \le j \le M$. Consequently, the assertion of part (b) follows. □

Now we are in a position to state and prove our convergent result for Algorithm 4.1.

Theorem 4.1 *Let $\{\zeta^k\}$ be the sequence generated by Algorithm 4.1. Suppose that F is Lipschitz continuous and satisfies the condition in Prop.3.3, and $\nabla F(\cdot)$ is Lipschitz continuous on $\mathcal{B}(\zeta^0)$. Then, the following results hold.*

(a) *The sequence $\{\zeta^k\}$ is bounded.*

(b) *The sequence $\{\Psi_\alpha(\zeta^k)\}$ is convergent.*

(c) $\lim_{k\to\infty} \gamma^{2l_k} h(\zeta^k) = 0$, $\lim_{k\to\infty} \gamma^{l_k}\|d^k\| = 0$ *and* $\lim_{k\to\infty} \|\zeta^{k+1} - \zeta^k\| = 0$.

(d) *Each accumulation point of $\{\zeta^k\}$ either is a solution of the SCCP (1) or satisfies*

$$\frac{|\nabla\Psi_\alpha(\zeta)^T d(\zeta)|}{h(\zeta)} = 0. \tag{27}$$

Proof. (a) By Prop.3.3, $\mathcal{L}(\Psi_\alpha, \zeta^0)$ is bounded, and the result holds by Lemma 4.1(a).

(b) First, by the proof of Lemma 4.1(a), the sequence $\{\Psi_\alpha(\zeta^{\sigma(k)})\}$ is nonincreasing. This together with the nonnegativity of $\Psi_\alpha(\zeta)$ for any $\zeta \in \mathbb{V}$ implies that $\{\Psi_\alpha(\zeta^{\sigma(k)})\}$ admits a limit when $k \to \infty$. Let j be an integer such that $1 \le j \le M+1$. We first by induction on j show that

$$\lim_{k\to\infty} \|\zeta^{\sigma(k)-j+1} - \zeta^{\sigma(k)-j}\| = 0, \tag{28}$$

$$\lim_{k\to\infty} \Psi_\alpha(\zeta^{\sigma(k)}) = \lim_{k\to\infty} \Psi_\alpha(\zeta^{\sigma(k)-j}), \tag{29}$$

where $\sigma(k)$ is defined as in Lemma 4.1, and the sequences are considered for sufficiently large k such that $\sigma(k) \ge k - M > 1$. If $j = 1$, then using (25) with k replaced by $\sigma(k) - 1$, we obtain that

$$\Psi_\alpha(\zeta^{\sigma(k)}) \le \Psi_\alpha(\zeta^{\sigma(\sigma(k)-1)}) - \delta\gamma^{2l_{\sigma(k)-1}} h(\zeta^{\sigma(k)-1}). \tag{30}$$

Since $\{\Psi_\alpha(\zeta^{\sigma(k)})\}$ admits a limit, taking limits to the both sides of (30) yields

$$\lim_{k\to\infty} \gamma^{2l_{\sigma(k)-1}} h(\zeta^{\sigma(k)-1}) = 0.$$

From the definition of $d(\zeta)$ and $h(\zeta)$, it is easy to verify that

$$h(\zeta) \ge \|d(\zeta)\|^2 \quad \text{for any } \zeta \in \mathbb{V}.$$

Using the last two equations, it then follows that

$$0 \ge \lim_{k\to\infty} \|\gamma^{l_{\sigma(k)-1}} d^{\sigma(k)-1}\| = \lim_{k\to\infty} \|\zeta^{\sigma(k)} - \zeta^{\sigma(k)-1}\| \ge 0. \tag{31}$$

On the other hand, since Ψ_α is continuously differentiable everywhere and $\mathcal{L}(\Psi_\alpha, \zeta^0)$ is bounded, the function Ψ_α is Lipschitz continuous on $\mathcal{L}(\Psi_\alpha, \zeta^0)$. This means that there exists a constant $L_2 > 0$ such that

$$|\Psi_\alpha(\zeta) - \Psi_\alpha(\xi)| \le L_2\|\zeta - \xi\| \quad \forall \zeta, \xi \in \mathcal{L}(\Psi_\alpha, \zeta^0). \tag{32}$$

From equations (31)–(32), we immediately obtain

$$\lim_{k\to\infty} \Psi_\alpha(\zeta^{\sigma(k)}) = \lim_{k\to\infty} \Psi_\alpha(\zeta^{\sigma(k)-1}).$$

This shows that (28) and (29) hold at each k for $j = 1$. Now assume that (29) holds for a given j. Using (25) with k replaced by $\sigma(k) - j - 1$, we have

$$\Psi_\alpha(\zeta^{\sigma(k)-j}) \le \Psi_\alpha(\zeta^{\sigma(\sigma(k)-j-1)}) - \delta\gamma^{2l_{\sigma(k)-j-1}} h(\zeta^{\sigma(k)-j-1}).$$

Taking limits for $k \to \infty$ and recalling (29) give

$$\lim_{k\to\infty} \gamma^{2l_{\sigma(k)-j-1}} h(\zeta^{\sigma(k)-j-1}) = 0.$$

This together with $h(\zeta^{\sigma(k)-j-1}) \ge \|d^{\sigma(k)-j-1}\|^2$ implies

$$0 \ge \lim_{k\to\infty} \gamma^{l_{\sigma(k)-j-1}} \|d^{\sigma(k)-j-1}\| = \lim_{k\to\infty} \|\zeta^{\sigma(k)-j} - \zeta^{\sigma(k)-j-1}\| = 0.$$

Combining with (29) and (32), we then obtain

$$\lim_{k\to\infty} \Psi_\alpha(\zeta^{\sigma(k)}) = \lim_{k\to\infty} \Psi_\alpha(\zeta^{\sigma(k)-j-1}).$$

The last two equations show that (28) and (29) hold when replacing j with $j+1$, and hence (28) and (29) hold for any given $j \in \{1, \ldots, M\}$. Let $\hat{\sigma}(k) = \sigma(k+M+1)$. Then,

$$\begin{aligned} \zeta^{\hat{\sigma}(k)} &= \zeta^k + (\zeta^{k+1} - \zeta^k) + \cdots + (\zeta^{\hat{\sigma}(k)} - \zeta^{\hat{\sigma}(k)-1}) \\ &= \zeta^k + \sum_{j=1}^{\hat{\sigma}(k)-k} (\zeta^{\hat{\sigma}(k)-j+1} - \zeta^{\hat{\sigma}(k)-j}). \end{aligned} \tag{33}$$

Notice that $\sigma(k+M+1) \le k+M+1$ and $\hat{\sigma}(k) - k \le M+1$, and therefore, from (33) and (28), it follows

$$\lim_{k\to\infty} \|\zeta^k - \zeta^{\hat{\sigma}(k)}\| = 0. \tag{34}$$

Since $\{\Psi_\alpha(\zeta^{\sigma(k)})\}$ has a limit, using (32) and (34), we have

$$\lim_{k\to\infty} \Psi_\alpha(\zeta^k) = \lim_{k\to\infty} \Psi_\alpha(\zeta^{\hat{\sigma}(k)}) = \lim_{k\to\infty} \Psi_\alpha(\zeta^{\sigma(k+M+1)}) = \lim_{k\to\infty} \Psi_\alpha(\zeta^{\sigma(k)}).$$

Thus, we complete the proof of assertion (b).

(c) From the line search condition (23) and part (b), it readily follows

$$\lim_{k\to\infty} \gamma^{2l_k} h(\zeta^k) = 0.$$

This together with $h(\zeta^k) \ge \|d^k\|^2$ and $\|\gamma^{l_k} d^k\| = \|\zeta^{k+1} - \zeta^k\|$ yields

$$\lim_{k\to\infty} \gamma^{l_k} \|d^k\| = \lim_{k\to\infty} \|\zeta^{k+1} - \zeta^k\| = 0.$$

Consequently, the assertions of part (c) hold.

(d) If $l_k = 0$ fails for the line search condition (23), then we have

$$\begin{aligned} \Psi_\alpha(\zeta^k + \gamma^{l_k-1} d^k) &> \max_{0\le j\le m(k)} \Psi_\alpha(\zeta^{k-j}) - \delta\gamma^{2(l_k-1)} h(\zeta^k) \\ &\ge \Psi_\alpha(\zeta^k) - \delta\gamma^{2(l_k-1)} h(\zeta^k). \end{aligned} \tag{35}$$

Since $F(\cdot)$ and $\nabla F(\cdot)$ are Lipschitz continuous on $\mathcal{B}(\zeta^0)$, it is clear that $\nabla\Psi_\alpha(\cdot)$ is Lipschitz continuous on this bounded set, i.e., there exists a constant $L_3 > 0$ such that

$$\|\nabla\Psi_\alpha(\zeta) - \nabla\Psi_\alpha(\xi)\| \le L_3 \|\zeta - \xi\| \quad \forall \zeta, \xi \in \mathcal{B}(\zeta^0). \tag{36}$$

Notice that ζ^k and $\zeta^k + td^k$ for any $t \in [0,1]$ belong to the set $\mathcal{B}(\zeta^0)$. By the mean-value theorem and the Lipschitz continuity of $\nabla\Psi_\alpha$ on $\mathcal{B}(\zeta^0)$, it then follows that

$$\begin{aligned}
& \Psi_\alpha(\zeta^k + td^k) - \Psi_\alpha(\zeta^k) \\
= \; & t\nabla\Psi_\alpha(\zeta^k)^T d^k + \int_0^t [\nabla\Psi_\alpha(\zeta^k + sd^k) - \nabla\Psi_\alpha(\zeta^k)]^T d^k ds \\
\leq \; & t\nabla\Psi_\alpha(\zeta^k)^T d^k + \int_0^t L_3\|d^k\|^2 s ds \\
= \; & t\nabla\Psi_\alpha(\zeta^k)^T d^k + (1/2)L_3 t^2\|d^k\|^2 \\
\leq \; & t\nabla\Psi_\alpha(\zeta^k)^T d^k + (1/2)L_3 t^2 h(\zeta^k) \\
\leq \; & -\delta t^2 h(\zeta^k) \quad \text{for all } t \in \left[0, \frac{2|\nabla\Psi_\alpha(\zeta^k)^T d^k|}{h(\zeta^k)(2\delta + L_3)}\right].
\end{aligned} \tag{37}$$

Combining the inequality (37) with (35), we obtain that

$$\gamma^{l_k - 1} > \frac{2|\nabla\Psi_\alpha(\zeta^k)^T d^k|}{h(\zeta^k)(2\delta + L_3)}.$$

If $l_k = 0$ succeeds for the line search condition (23), then $\gamma^{l_k} = 1$. Thus, there exists some constant $C_1 = 2\gamma/(2\delta + L_3) > 0$ such that

$$\gamma^{l_k} > \min\left\{1, C_1 \frac{|\nabla\Psi_\alpha(\zeta^k)^T d^k|}{h(\zeta^k)}\right\} \quad \text{for all } k. \tag{38}$$

Now let ζ^* be an accumulation point of $\{\zeta^k\}$ and $\{\zeta^k\}_{k\in K}$ be the subsequence such that

$$\lim_{k\to\infty, k\in K} \zeta^k = \zeta^*.$$

By part (c), $\lim_{k\to\infty} \gamma^{2l_k} h(\zeta^k) = 0$. If $\lim_{k\to\infty, k\in K} h(\zeta^k) = h(\zeta^*) = 0$, then

$$\left\|\nabla_x\psi_\alpha(\zeta^*, F(\zeta^*)) + \nabla_x\psi_\alpha(\zeta^*, F(\zeta^*))\right\| = 0.$$

By Proposition 3.1 (c), ζ^* is a solution of the SCCP. If $\lim_{k\to\infty} h(\zeta^k) \neq 0$, then there holds $\lim_{k\to\infty} \gamma^{l_k} = 0$. This together with (38) implies

$$0 = \lim_{k\to\infty} \frac{|\nabla\Psi_\alpha(\zeta^k)^T d^k|}{h(\zeta^k)} = \frac{|\nabla\Psi_\alpha(\zeta^*)^T d(\zeta^*)|}{h(\zeta^*)}.$$

Thus, we complete the proof. □

Theorem 4.1 states that, when θ is any fixed real number in $[0, 1]$, the nonmonotone derivative-free algorithm converges in terms of the value of merit function Ψ_α and the sequence $\{\zeta^k\}$ is bounded for a large class of SCCPs which may even not be monotone. If θ is chosen to be less than $\bar{\theta}$ and F is strongly monotone, then by Prop.3.4,

$$|\nabla\Psi_\alpha(\zeta)^T d(\zeta)| \geq \frac{1}{2}\theta h(\zeta) \quad \forall \zeta \in \mathcal{B}(\zeta^0).$$

This implies that any accumulation point of $\{\zeta^k\}$ can not satisfy (27), and consequently, each accumulation of $\{\zeta^k\}$ is a solution of the SCCP (1). In fact, under this case, $\{\zeta^k\}$ converges to the solution of (1) at a R-linear rate. We next prove the assertion.

Theorem 4.2 *Let $\{\zeta^k\}$ be the sequence generated by Algorithm 4.1. Suppose that F is strongly monotone and Lipschitz continuous, and $\nabla F(\cdot)$ is Lipschitz continuous on $\mathcal{B}(\zeta^0)$. If $\theta \leq \bar{\theta}$ with $\bar{\theta}$ given by (22), then there exist constants $\nu_0 > 0$ and $\nu_6 \in (0,1)$ such that*

$$\Psi_\alpha(\zeta^k) \leq \nu_0 \nu_6^k \Psi_\alpha(\zeta^1).$$

Moreover, $\{\zeta^k\}$ converges to the unique solution ζ^* of the SCCP (1) with R-linear rate.

Proof. Since strong monotonicity implies the uniform Jordan P-property, which by Prop.3.3 implies that $\mathcal{B}(\zeta^0)$ is bounded and all results of Theorem 4.1 hold.

To prove the conclusion, we first show that there exist constants $\nu_1,\ \nu_2 > 0$ such that

$$\Psi_\alpha(\zeta^{k+1}) \leq \nu_1 \Psi_\alpha(\zeta^k) \quad \text{for all } k \geq 0, \tag{39}$$

and

$$h(\zeta^{k+1}) \leq \nu_2 h(\zeta^k) \quad \text{for all } k \geq 0. \tag{40}$$

Because $\theta \leq \bar{\theta}$ and F is strongly monotone, using (37) and Proposition 3.4 yields

$$\begin{aligned}\Psi_\alpha(\zeta^{k+1}) - \Psi_\alpha(\zeta^k) &\leq \gamma^{l_k}\left[\nabla\Psi_\alpha(\zeta^k)^T d^k + (1/2)L_3\gamma^{l_k}h(\zeta^k)\right] \\ &\leq -\frac{1}{2}\gamma^{l_k}(\theta - L_3\gamma^{l_k})h(\zeta^k). \end{aligned}\tag{41}$$

By Proposition 3.1 (e)–(f), Lemma 3.1 (a) and (c), it is easy to verify that

$$h(\zeta) \geq \frac{(\alpha-1)^2}{\alpha^2}\|R_\alpha(\zeta)\|^2 \geq \frac{(\alpha-1)^2}{\alpha^2}\|R_1(\zeta)\|^2 \geq \frac{\alpha-1}{\alpha^2}\Psi_\alpha(\zeta) \quad \forall \zeta \in \mathbb{V}, \tag{42}$$

and

$$h(\zeta) \leq 2\alpha(\alpha-1)\Psi_\alpha(\zeta) \quad \forall \zeta \in \mathbb{V}. \tag{43}$$

Therefore, if $\theta - L_3\gamma^{l_k} \geq 0$, equations (41) and (42) imply

$$\begin{aligned}\Psi_\alpha(\zeta^{k+1}) &\leq \Psi_\alpha(\zeta^k) - \frac{1}{2}\gamma^{l_k}(\theta - L_3\gamma^{l_k})\frac{\alpha-1}{\alpha^2}\Psi_\alpha(\zeta^k) \\ &= \left[1 - \frac{1}{2}\gamma^{l_k}(\theta - L_3\gamma^{l_k})\frac{\alpha-1}{\alpha^2}\right]\Psi_\alpha(\zeta^k) \leq \Psi_\alpha(\zeta^k);\end{aligned}$$

whereas if $\theta - L_3\gamma^{l_k} < 0$, equations (41) and (43) lead to

$$\begin{aligned}\Psi_\alpha(\zeta^{k+1}) &\leq \left[1 - \gamma^{l_k}(\theta - L_3\gamma^{l_k})\alpha(\alpha-1)\right]\Psi_\alpha(\zeta^k) \\ &\leq [1 + (L_3 - \theta)\alpha(\alpha-1)]\Psi_\alpha(\zeta^k).\end{aligned}$$

This shows that (39) holds with $\nu_1 := \max\{1, 1 + (L_3 - \theta)\alpha(\alpha - 1)\}$. Using (43), (39) and (42), we have

$$h(\zeta^{k+1}) \le 2\alpha(\alpha-1)\Psi_\alpha(\zeta^{k+1}) \le 2\alpha(\alpha-1)\nu_1\Psi_\alpha(\zeta^k) \le 2\nu_1\alpha^3 h(\zeta^k),$$

which implies that (40) holds with $\nu_2 := 2\nu_1\alpha^3 > 0$.

Now for any $p \ge 1$, let $\phi(p)$ be any index in $[Mp+1, M(p+1)]$ satisfying

$$\Psi_\alpha(\zeta^{\phi(p)}) := \max_{1\le i\le M} \Psi_\alpha(\zeta^{Mp+i}).$$

From Lemma 4.1 (b), it then follows

$$\Psi_\alpha(\zeta^{\phi(p)}) \le \Psi_\alpha(\zeta^{\phi(p-1)}) - \delta \min_{0\le i\le M-1} \gamma^{2l_{(Mp+i)}} h(\zeta^{Mp+i}).$$

Notice that $\gamma^{l_k} \ge \min\{1, C_1\theta/2\}$ for all k by using (38) and the second assertion of Proposition 3.4. Hence, there exists a constant $\nu_3 := \delta\min\{1, C_1\theta/2\} > 0$ such that

$$\Psi_\alpha(\zeta^{\phi(p)}) \le \Psi_\alpha(\zeta^{\phi(p-1)}) - \nu_3 \min_{0\le i\le M-1} h(\zeta^{Mp+i}). \tag{44}$$

Let $s(p)$ and $w(p)$ be any indices in $[Mp+1, M(p+2)]$ for which

$$h(\zeta^{s(p)}) := \min_{1\le i\le 2M} h(\zeta^{Mp+i}) \quad \text{and} \quad \Psi_\alpha(\zeta^{w(p)}) := \min_{1\le i\le 2M} \Psi_\alpha(\zeta^{Mp+i}), \tag{45}$$

and denote by ν_4 the constant given by

$$\nu_4 = \left[\nu_3 + \frac{\alpha^2}{\alpha-1}\nu_2^{4M}\right]^{-1}. \tag{46}$$

We now define an infinite subsequence $\{k_i : i \ge 0\} \subset \{1, 2, \ldots\}$ as follows. Let $k_0 = \phi(0)$. Suppose that $k_i = \phi(\bar{p})$ has been chosen for some $\bar{p}$. Define

$$k_{i+1} := \begin{cases} w(\bar{p}+1) & \text{if } h(\zeta^{s(\bar{p}+1)}) \le \nu_4\Psi_\alpha(\zeta^{\phi(\bar{p})}) \\ \phi(\bar{p}+3) & \text{otherwise.} \end{cases} \tag{47}$$

For the subsequence $\{k_i\}$ defined as above, it is obvious that

$$k_{i+1} - k_i \le 4M. \tag{48}$$

In addition, there necessarily exists a constant $\nu_5 \in (0,1)$ such that

$$\Psi_\alpha(\zeta^{k_{i+1}}) \le \nu_5\Psi_\alpha(\zeta^{k_i}), \quad \text{for all } i \ge 1. \tag{49}$$

In fact, if $h(\zeta^{s(\bar{p}+1)}) \le \nu_4\Psi_\alpha(\zeta^{\phi(\bar{p})})$, from (42), (40) and (48), it follows that

$$\Psi_\alpha(\zeta^{k_{i+1}}) \le \frac{\alpha^2}{\alpha-1}h(\zeta^{k_{i+1}}) \le \frac{\alpha^2}{\alpha-1}\nu_2^{4M}h(\zeta^{s(\bar{p}+1)}) \le \frac{\alpha^2}{\alpha-1}\nu_2^{4M}\nu_4\Psi_\alpha(\zeta^{k_i}).$$

If $h(\zeta^{s(\bar{p}+1)}) > \nu_4 \Psi_\alpha(\zeta^{\phi(\bar{p})})$, using (44) and (45) yields

$$\Psi_\alpha(\zeta^{k_{i+1}}) \leq (1-\nu_3\nu_4)\Psi_\alpha(\zeta^{k_i})$$

By the choice of ν_4, the last two equations imply that (49) holds with $\nu_5 = (1-\nu_3\nu_4)$.

For any $k \geq 1$, assume that $k \in [k_i, k_{i+1})$ for some i. Then from (48) we have that

$$k - k_i \leq 4M \quad \text{and} \quad k_i \leq 4Mi + k_0. \tag{50}$$

Using equation (50) and noting that $1 \leq k_0 \leq M$ give

$$i \geq \frac{k_i - k_0}{4M} \geq \frac{k - 4M - k_0}{4M} \geq \frac{k}{4M} - \frac{5}{4}. \tag{51}$$

Thus, by (39), (49), (50)–(51), we obtain

$$\begin{aligned} \Psi_\alpha(\zeta^k) &\leq \nu_1^{k-k_i}\Psi_\alpha(\zeta^{k_i}) \leq \nu_1^{4M}\nu_5^{i}\Psi_\alpha(\zeta^{k_0}) \\ &\leq \nu_1^{4M}\nu_5^{(k/(4M)-5/4)}\Psi_\alpha(\zeta^{k_0}) \\ &\leq \nu_1^{5M}\nu_5^{(k/(4M)-5/4)}\Psi_\alpha(\zeta^{1}). \end{aligned}$$

Letting $\nu_0 = \nu_1^{5M}\nu_5^{-5/4}$ and $\nu_6 = \nu_5^{1/(4M)}$ and noting that $\nu_5 = (1-\nu_3\nu_4) < 1$, we prove the first part of the conclusion. The second part is direct since $\{\Psi_\alpha(\zeta^k)\}$ converges Q-linearly to zero and $\|\zeta^k - \zeta^*\| \leq \frac{L+1}{\rho}\sqrt{\frac{\alpha}{\alpha-1}}\sqrt{\Psi_\alpha(\zeta^k)}$ by Prop.3.2(b). □

Theorem 4.2 is the first rate of convergence result for the class of derivative-free descent methods with a nonmonotone line search rule for the non-polyhedral SCCPs. In the next section, we compare the numerical performance of Algorithm 4.1 with that of Algorithm 4.2 descried as below, which is a monotone descent derivative-free method similar to the one in [26] for the NCPs. The stepsize and the search direction of Algorithm 4.2 are adjusted during the backtracking search of Armijo-type.

Algorithm 4.2

(Step 0) Choose $\zeta^0 \in \mathbb{V}$, $\epsilon \geq 0$, $\delta \in (0,1)$, $\gamma \in (0,1)$, and a sufficiently small $\beta \in (0,1)$. Set $k := 0$.

(Step 1) If $\Psi_\alpha(\zeta^k) \leq \varepsilon$, then stop. Otherwise, go to Step 2.

(Step 2) Let l_k be the smallest nonnegative integer l satisfying

$$\Psi_\alpha(\zeta^k + \gamma^l d^k(\beta^l)) \leq \Psi_\alpha(\zeta^k) - \delta\gamma^{2l}h(\zeta^k), \tag{52}$$

where $h(\zeta)$ is defined as in (24) and

$$d^k(\beta^l) := -\beta^l\nabla_x\psi_\alpha(\zeta^k, F(\zeta^k)) - (1-\beta^l)\nabla_y\psi_\alpha(\zeta^k, F(\zeta^k)). \tag{53}$$

(Step 3) Set $\zeta^{k+1} := \zeta^k + \gamma^{l_k} d^k(\beta^{l_k})$, $k := k+1$, and go to Step 1.

5 Numerical experiments

In this section, we test the performance of Algorithms 4.1 and 4.2 for the affine SOCCP

$$\zeta \in K_+^n, \quad F(\zeta) = M\zeta + b \in K_+^n, \quad \langle \zeta, F(\zeta) \rangle = 0, \tag{54}$$

where $K_+^n = K_+^{n_1} \times \cdots \times K_+^{n_m}$ with $n_1 + \cdots + n_m = n$, $M \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$.

During the testing, we set $M \equiv \mathrm{diag}(M_1, \cdots, M_m)$ with $M_i = N_i N_i^T + \tau I_i$ for all i, where $\tau \geq 0$ is a given parameter, I_i is an $n_i \times n_i$ identity matrix, and each $N_i \in \mathbb{R}^{n_i\times n_i}$ was generated randomly such that it has 1% nonzero density with the nonzero entries from a normal distribution of mean -1 and variance 4. It is not hard to see that the matrix M generated by such a way is positive semidefinite (respectively, positive definite) if $\tau = 0$ (respectively, $\tau > 0$), which means that the corresponding F is strongly monotone (or monotone). The vector b was obtained by setting $b = -Mw$ with $w = (w_1, \ldots, w_m) \in K_+^n$, where $w_i \in K_+^{n_i}$ was generated as follows: let the elements of w_i be chosen randomly from a normal distribution with mean -1 and variance 4, and then set the first element w_{i1} of w_i to be $\|w_{i2}\|$, where w_{i2} is a vector composed of the rest $n_i - 1$ components of w_i. In this way, the affine SOCCP is guaranteed to have a solution $\zeta^* = w$.

All experiments were done with a PC of Pentium 4 with 2.8GHz CPU and 512MB memory. The computer codes were written in Matlab 6.5. During the tests, we chose n_i and m such that $n_1 = \cdots = n_m = 10$ and $m = 100$. We set $m(k)$ in Algorithm 4.1 as

$$m(k) := \begin{cases} 0 & k < 5 \\ \min\{m(k-1)+1, M-1\} & \text{otherwise} \end{cases} \quad \text{with} \quad M = 6.$$

We started Algorithms 4.1 and 4.2 from the initial point $\zeta^0 = (\bar{\zeta}^{n_i}, \ldots, \bar{\zeta}^{n_m})$ with $\bar{\zeta}^{n_i} = (10, \omega_i/\|\omega_i\|)$, where $\omega_i \in \mathbb{R}^{n_i-1}$ for all i were generated randomly by Matlab's **rand.m**. The parameters γ and δ in the two algorithms, and β in Algorithm 4.2 were chosen as

$$\gamma = 0.2, \quad \delta = 10^{-10}, \quad \text{and} \quad \beta = 0.1.$$

The algorithms were terminated once one of the following conditions is satisfied:

(a) $\min\left\{\Psi_\alpha(\zeta^k), |\langle \zeta^k, F(\zeta^k)\rangle|\right\} \leq 10^{-5}$;

(b) The stepsize is less than 10^{-8};

(c) The maximum iteration number is over 5×10^5.

If the algorithms are stopped under condition (a), we say that they solve the test problem successfully, and otherwise say that they fail to the test problem.

We first tested the influence of α for the iterations and the function evaluations needed by Algorithms 4.1 and 4.2 for solving (54) with τ in each M_i chosen as 0.1. For every $\alpha = 2, 5, 10, 20, 40, 50, 60, 80, 100, 150, 200$, we applied Algorithm 4.1 with $\theta = 0.95$ and Algorithm 4.2, respectively, for solving the same 50 test problems generated as above. The the average iteration and average function evaluation were respectively taken as the average of iterations and function evaluations of the test problems solved successfully. The testing results show that Algorithm 4.1 with $\alpha = 2$ failed for 4 test problems due to too small stepsize, and successfully solved all test problems with the other α; whereas Algorithm 4.2 with $\alpha = 2$ and $\alpha = 5$ failed for 11 and 1 test problems, respectively, due to too small stepsize, and successfully solved all test problems with the rest α.

Figures 1 and 2 depict the curves of the average function evaluation and the average iteration, respectively, of Algorithms 4.1 and 4.2 with respect to α. From these figures, we see that the number of function evaluations and the iteration times needed by Algorithm 4.1 and Algorithm 4.2 increase with α. Taking into account that the global convergence of the two algorithms is not stable when α is close to 1 (for example they fail to some test problems when $\alpha = 2$), a desirable choice for α should be in the interval $[10, 50]$. Also, the average function evaluation and the average iteration of Algorithm 4.2 are more than those of Algorithm 4.1, especially when $\alpha > 40$. This implies that the non-monotone derivative-free method has better performance than the monotone descent one.

Then, we tested the influence of θ for the rate of convergence of Algorithm 4.1, by using this algorithm with $\alpha = 15$ and four different θ to solve a test example generated as above with $\tau = 0.01$. Figure 3 below depicts the convergence curve of Algorithm 4.1. From this figure, we see that the curve corresponding to $\theta = 0.5$ has the largest slope rate, the curve corresponding to $\theta = 10^{-4}$ has the smallest slope rate, and the curve corresponding to a smaller θ has a smaller slope rate when $\theta \leq 0.1$. This shows that Algorithm 4.1 with a smaller θ has a better rate of convergence, and it has the worst rate of convergence when $\theta = 0.5$. This coincides with the theoretical results of Theorem 4.2.

We also tested the influence of θ for the performance of Algorithm 4.1. Specifically, for every $\theta = 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95$, we em-

Figure 1: Influence of α on the average function evaluation of Algorithms 4.1 and 4.2

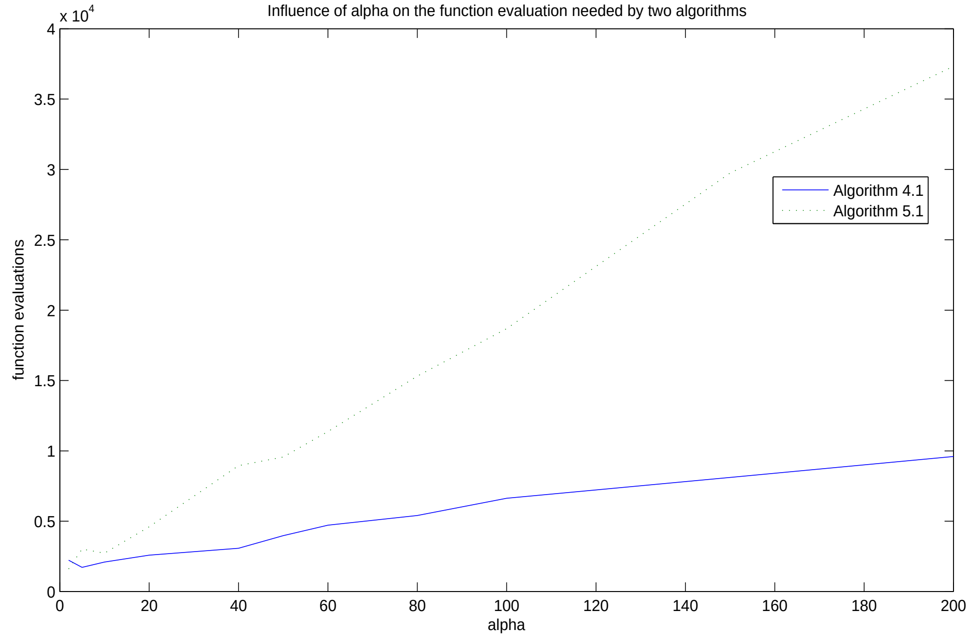
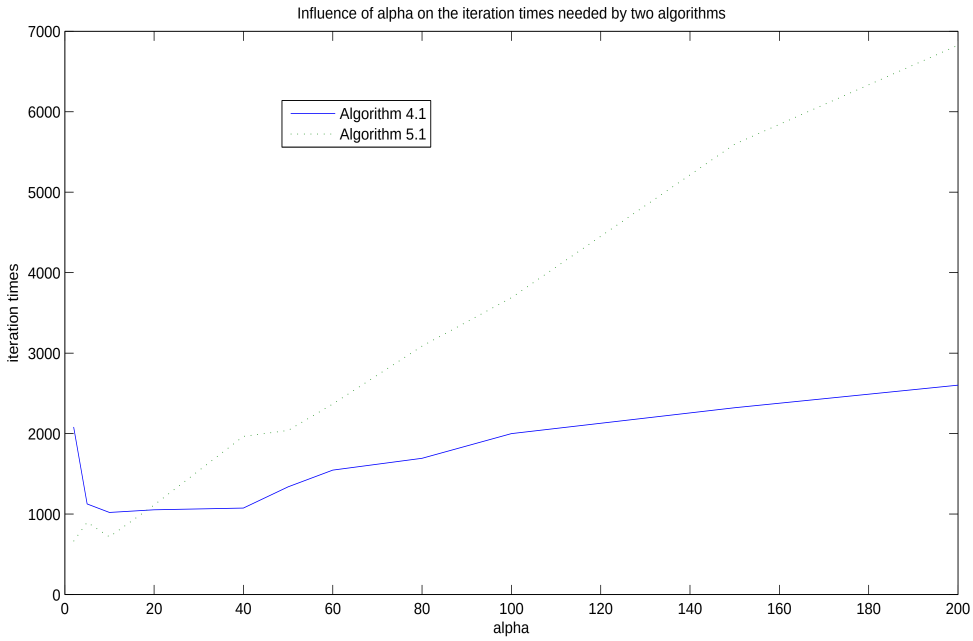


Figure 2: Influence of α on the average iteration of Algorithms 4.1 and 4.2



ployed Algorithm 4.1 with $\alpha = 15$ to solve the same 50 test problems generated as above with $\tau = 0$. Note that this class of problems is more difficult than the one used above since the mapping F is now only monotone, instead of strongly monotone. The testing results show that Algorithm 4.1 successfully solved all test problems with all these θ. This shows that Algorithm 4.1 is also suitable for the solution of monotone SCCPs although the global convergence of the sequence generated is not established for this class of problems. Figure 4 below depicts the curves of the function evaluation and the iteration times of Algorithm 4.1 with respect to θ. From this figure, we see that Algorithm 4.1 has the worst performance when $\theta = 0.5$, and a desirable θ should be from the interval $[0.2, 0.4]$ or $[0.9, 1)$.

6 Conclusion

We have extended the derivative-free method [18] for the NCP to the general SCCPs by using a different search direction. It was shown that the algorithm is convergent in terms of the value of Ψ_α for a large class of SCCPs which may not even be monotone, whereas if $\theta \le \bar{\theta}$ with $\bar{\theta}$ given by (22) and F is strongly monotone, the sequence generated by the algorithm converge globally to the solution of the problem at a R-linear rate. It is interesting to note that the linear convergence rate of the nonmonotone descent algorithm is obtained without requiring any convexity of Ψ_α, and the relation among $R_1(\zeta), h(\zeta)$ and $\Psi_\alpha(\zeta)$ plays a key role. In the future research, it is worthwhile to study the convergence rate of nomonotone derivative-free methods based on other merit functions, and explore other derivative-free methods for the SCCPs, for example, the pattern search algorithms.

References

[1] J.-S. Chen, H.-T. Gao and S.-H. Pan, *A derivative-free R-linearly convergent algorithm based on the generalized Fischer-Burmeister merit function*, Journal of Computational and Applied Mathematics, vol. 232, pp. 455-471, 2009.

[2] Y.-H. Dai, *On the nonmonotone line search*, Journal of Optimization Theory and Applications, vol. 112, pp. 315-330, 2002.

[3] L. Faybusovich, *Euclidean Jordan algebras and interior-point algorithms*, J. Positivity, vol. 1, pp. 331-357, 1997.

Figure 3: Convergence process of Algorithm 4.1 with different θ

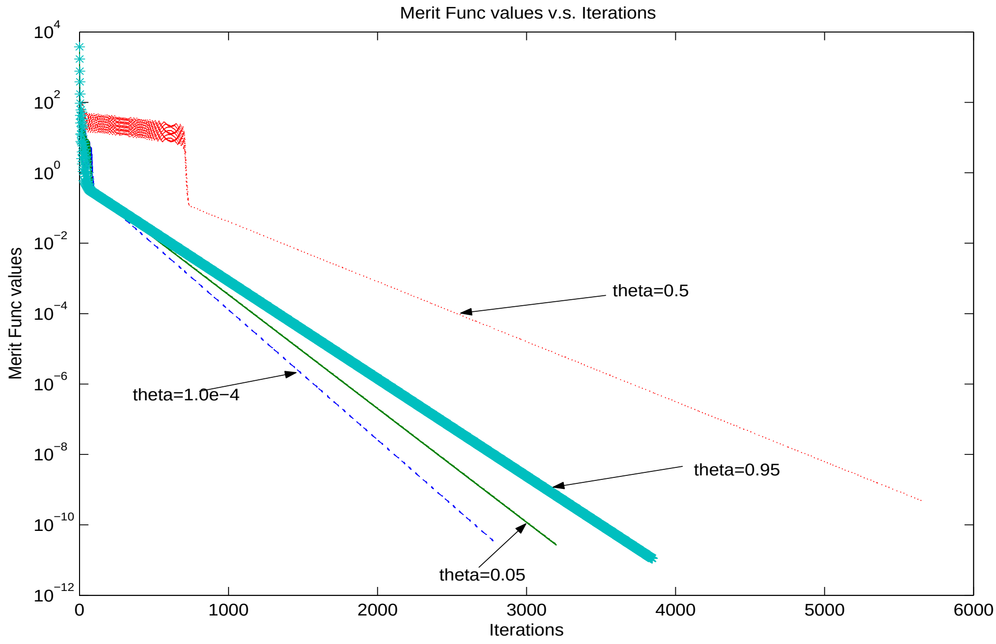
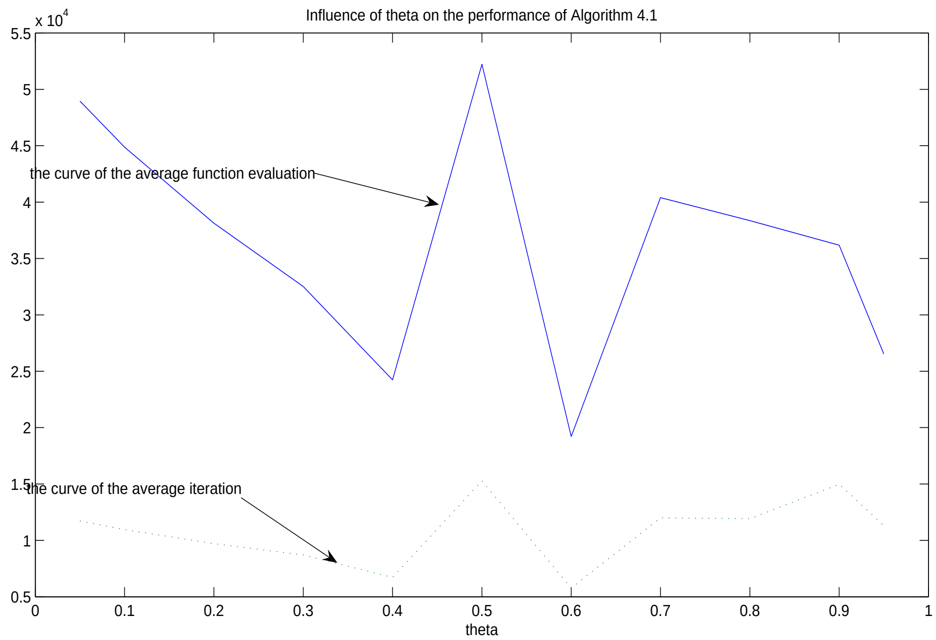


Figure 4: Influence of θ on the performance of Algorithm 4.1



[4] E. M. Gafni and D. P. Bertsekas, *Two-meric projection methods for constrained optimization*, SIAM Journal on Control and Optimization, vol. 22, pp. 936-964, 1984.

[5] X. Chen and H. Qi, *Cartesian P-property and its applications to the semidefinite linear complementarity problem*, Mathematical Programming, vol. 106, pp. 177-201, 2006.

[6] L. Grippo, F. Lampariello and S. Lucidi, *A nonmonotone line search technique for Newton's method*, SIAM Journal on Numerical Analysis, vol. 23, pp. 707-716, 1986.

[7] C. Geiger and C. Kanzow, *On the resolution of monontone complementarity problems*, Computational Optimization and Applications, vol. 5, pp. 155-173, 1996.

[8] Z.-H. Huang and T. Ni, *Smoothing algorithms for complementarity problems over symmetric cones*, Computational Optimization and Applications, vol. 45, pp. 557-579, 2010.

[9] H. Jiang, *Unconstrained minimization approaches to nonlinear complementarity problems*, Journal of Global Optimization, vol. 9, pp. 169-181, 1996.

[10] C. Kanzow, *Nonlinear complementarity as unconstrained optimization*, Journal of Optimization Theory and Applications, vol. 88, pp. 139-155, 1996.

[11] J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs, Oxford University Press, New York, 1994.

[12] C. Kanzow, Y. Yamashita, and M. Fukushima, *New NCP functions and their properties*, Journal of Optimization Theory and Applications, vol. 97, pp. 115-135, 1997.

[13] L. C. Kong, J. Sun and N. H. Xiu, *A regularized smoothing Newton method for symmetric cone complementarity problems*, SIAM Journal on Optimization, vol. 19, pp. 1028-1047, 2008.

[14] L. C. Kong, L. Tuncel and N. H. Xiu, *Vector-valued implicit Lagrangian for symmetric cone complementarity problems*, Asia-Pacific Journal of Operational Research, vol. 26, pp. 199-233, 2009.

[15] Z.-Q. Luo and O. L. Mangasarian, *New error bounds for the linear complementarity problem*, Mathematics of Operations Research, vol. 19, pp. 880-892, 1994.

[16] Y. Liu, L. Zhang and Y. Wang, *Some propeties of a class of merit functions for symmetric cone complementarity problems*, Asia-Pacific Journal of Operational Research, vol. 23, pp. 473-496, 2006.

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

[18] O. L. Mangasarian and M. V. Solodov, *A linearly convergent derivative-free descent method for strongly monotone complementarity problems*, Computational Optimization and Applications, vol. 14, pp. 5-16, 1999.

[19] S.-H. Pan and J.-S. Chen, *Growth behavior of two classes of merit functions for symmetric cone complementarity problems*, Journal of Optimization Theory and Applications, vol. 141, pp. 167-191, 2009.

[20] S.-H. Pan and J.-S. Chen, *A linearly convergent derivative-free descent method for the second-order cone complementarity problem*, Optimization, vol. 59, pp. 1173-1197, 2010.

[21] J. M. Peng, *Equivalence of varational inequality problems to unconstrained minimization*, Mathematical Programming, vol. 78, pp. 347-355, 1997.

[22] S. Schmieta and F. Alizadeh, *Extensions of primal-dual interior-point algorithms to symmetric cones*, Mathematical Programming, vol. 96, pp. 409-438, 2003.

[23] D. Sun and J. Sun, *Löwner's operator and spectral functions on Euclidean Jordan algebras*, Mathematics of Operations Research, vol. 33, pp. 421-425, 2008.

[24] J. Tao and M. S. Gowda, *Some P-properties for nonlinear transformations on Euclidean Jordan algebras*, Mathematics of Operations Research, vol. 30, pp. 985-1004, 2005.

[25] P. Tseng, N. Yamashita, and M. Fukushima, *Equivalence of complementarity problems to differentiable minimization: A unified approach*, SIAM Journal on Optimization, vol. 6, pp. 446-460, 1996.

[26] K. Yamada, N. Yamashita, and M. Fukushima, *A new derivative-free descent method for the nonlinear complementarity problems*, in Nonlinear Optimization and Related Topics edited by G.D. Pillo and F. Giannessi, Kluwer Academic Publishers, Netherlands, pp. 463-487, 2000.

[27] N. Yamashita and M. Fukushima, *On stationary points of the implicit Lagrangian for nonlinear complementarity problems*, Journal of Optimization Theory and Applications, vol. 84, pp. 653-663, 1995.

[28] N. Yamashita and M. Fukushima, *A new merit function and a descent method for semidefinite complementarity problems*, Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, edited by

M. Fukushima, M. Fashumio and L. Qi, Kluwer Academic Publishers, pp. 405-420, 1998.

[29] A. Yoshise, *Interior point trajectiries and homogeneous model for non-linear complementarity problems over symmetric cones*, SIAM Journal on Optimization, vol. 17, pp. 1129-1153, 2006.