Comput Optim Appl (2012) 51:623–648
DOI 10.1007/s10589-010-9359-x

# Neural networks for solving second-order cone constrained variational inequality problem

**Juhe Sun · Jein-Shan Chen · Chun-Hsu Ko**

Received: 25 April 2010 / Published online: 6 October 2010


**Abstract** In this paper, we consider using the neural networks to efficiently solve the second-order cone constrained variational inequality (SOCCVI) problem. More specifically, two kinds of neural networks are proposed to deal with the Karush-Kuhn-Tucker (KKT) conditions of the SOCCVI problem. The first neural network uses the Fischer-Burmeister (FB) function to achieve an unconstrained minimization which is a merit function of the Karush-Kuhn-Tucker equation. We show that the merit function is a Lyapunov function and this neural network is asymptotically stable. The second neural network is introduced for solving a projection formulation whose solutions coincide with the KKT triples of SOCCVI problem. Its Lyapunov stability and global convergence are proved under some conditions. Simulations are provided to show effectiveness of the proposed neural networks.

**Keywords** Second-order cone · Variational inequality · Fischer-Burmeister function · Neural network · Lyapunov stable · Projection function

---

J. Sun is also affiliated with Department of Mathematics, National Taiwan Normal University.
J.-S. Chen is member of Mathematics Division, National Center for Theoretical Sciences, Taipei Office. The author's work is partially supported by National Science Council of Taiwan.

---

J. Sun
School of Science, Shenyang Aerospace University, Shenyang 110136, China
e-mail: juhesun@abel.math.ntnu.edu.tw

J.-S. Chen (✉)
Department of Mathematics, National Taiwan Normal University, Taipei 11677, Taiwan
e-mail: jschen@math.ntnu.edu.tw

C.-H. Ko
Department of Electrical Engineering, I-Shou University, Kaohsiung 840, Taiwan
e-mail: chko@isu.edu.tw

## 1 Introduction

Variational inequality (VI) problem, which was introduced by Stampacchia and his collaborators [19, 29, 30, 35, 36], has attracted much attention from researchers of engineering, mathematics, optimization, transportation science, and economics communities, see [1, 25, 26]. It is well known that VIs subsume many other mathematical problems, including the solution of systems of equations, complementarity problems, and a class of fixed point problems. For a complete discussion and history of the finite VI problem and its associated solution methods, we refer the interested readers to the excellent survey text by Facchinei and Pang [13], the monograph by Patriksson [34], the survey article by Harker and Pang [18], the Ph.D. thesis of Hammond [16] and the references therein.

In this paper, we are interested in solving the second-order cone constrained variational inequality (SOCCVI) problem whose constraints involve the Cartesian product of second-order cones (SOCs). The problem is to find $x \in C$ satisfying

$$\langle F(x), y - x\rangle \geq 0, \quad \forall y \in C, \tag{1}$$

where the set $C$ is finitely representable as

$$C = \{x \in \mathbb{R}^n : h(x) = 0,\ -g(x) \in \mathcal{K}\}. \tag{2}$$

Here $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product, $F : \mathbb{R}^n \to \mathbb{R}^n$, $h : \mathbb{R}^n \to \mathbb{R}^l$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable functions and $\mathcal{K}$ is a Cartesian product of second-order cones (or Lorentz cones), expressed as

$$\mathcal{K} = \mathcal{K}^{m_1} \times \mathcal{K}^{m_2} \times \cdots \times \mathcal{K}^{m_p}, \tag{3}$$

where $l \geq 0$, $m_1, m_2, \ldots, m_p \geq 1$, $m_1 + m_2 + \cdots + m_p = m$, and

$$\mathcal{K}^{m_i} := \{(x_{i1}, x_{i2}, \ldots, x_{im_i})^T \in \mathbb{R}^{m_i} \mid \|(x_{i2}, \ldots, x_{im_i})\| \leq x_{i1}\}$$

with $\|\cdot\|$ denoting the Euclidean norm and $\mathcal{K}^1$ the set of nonnegative reals $\mathbb{R}_+$. A special case of (3) is $\mathcal{K} = \mathbb{R}^n_+$, namely the nonnegative orthant in $\mathbb{R}^n$, which corresponds to $p = n$ and $m_1 = \cdots = m_p = 1$. When $h$ is affine, an important special case of the SOCCVI problem corresponds to the KKT conditions of the convex second-order cone program (CSOCP):

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & Ax = b,\ -g(x) \in \mathcal{K}, \end{array} \tag{4}$$

where $A \in \mathbb{R}^{l\times n}$ has full row rank, $b \in \mathbb{R}^l$, $g : \mathbb{R}^n \to \mathbb{R}^m$ and $f : \mathbb{R}^n \to \mathbb{R}$. Furthermore, when $f$ is a convex twice continuously differentiable function, problem (4) is equivalent to the following SOCCVI problem: Find $x \in C$ satisfying

$$\langle \nabla f(x), y - x\rangle \geq 0, \quad \forall y \in C,$$

where

$$C = \{x \in \mathbb{R}^n : Ax - b = 0,\ -g(x) \in \mathcal{K}\}.$$

For solving the constrained variational inequalities and complementary problems (CP), many computational methods have been proposed, see [3, 4, 6, 8, 13, 39] and references therein. These methods include the method based on merit function, interior method, Newton method, nonlinear equation method, projection method and its variant versions. Another class of techniques for solving the VI problem exploits the fact that the KKT conditions of a VI problem comprise a mixed complementarity problem (MiCP), involving both equations and nonnegativity constraints. In other words, the SOCCVI problem can be solved by analyzing its KKT conditions:

$$\begin{cases} L(x,\mu,\lambda) = 0, \\ \langle g(x),\lambda\rangle = 0, \qquad -g(x)\in\mathcal{K},\ \lambda\in\mathcal{K}, \\ h(x) = 0, \end{cases} \tag{5}$$

where $L(x,\mu,\lambda) = F(x) + \nabla h(x)\mu + \nabla g(x)\lambda$ is the variational inequality Lagrangian function, $\mu\in\mathbb{R}^l$ and $\lambda\in\mathbb{R}^m$. However, in many scientific and engineering applications, it is desirable to have a real-time solution for the VI and CP problems. Thus, at present, for solving the VI and CP problems, many researchers employ the neural network method which is a promising way to overcome this problem.

Neural networks for optimization were first introduced in the 1980s by Hopfield and Tank [20, 38]. Since then, neural networks have been applied to various optimization problems, including linear programming, nonlinear programming, variational inequalities, and linear and nonlinear complementarity problems; see [5, 9–11, 17, 21, 22, 24, 28, 40–44]. The main idea of the neural network approach for optimization is to construct a nonnegative energy function and establish a dynamic system that represents an artificial neural network. The dynamic system is usually in the form of first order ordinary differential equations. Furthermore, it is expected that the dynamic system will approach its static state (or an equilibrium point), which corresponds to the solution for the underlying optimization problem, starting from an initial point. In addition, neural networks for solving optimization problems are hardware-implementable; that is, the neural networks can be implemented by using integrated circuits. In this paper, we focus on neural network approach to the SOCCVI problem. Our neural networks will be aimed to solve the system (5) whose solutions are candidates of SOCCVI problem (1).

The rest of this paper is organized as follows. Section 2 introduces some preliminaries. In Sect. 3, the first neural network based on the Fischer-Burmeister function is proposed and studied. In Sect. 4, we show that the KKT system (5) is equivalent to a nonlinear projection formulation. Then, the model of neural network for solving the projection formulation is introduced and its stability is analyzed. In Sect. 5, illustrative examples are discussed. Section 6 gives the conclusion of this paper.

## 2 Preliminaries

In this section, we recall some preliminary results that will be used later and background materials of ordinary differential equations that will play an important role in the subsequent analysis. We begin with some concepts for a nonlinear mapping.

**Definition 2.1** Let $F = (F_1, \ldots, F_n)^T : \mathbb{R}^n \to \mathbb{R}^n$. Then, the mapping $F$ is said to be

(a) monotone if

$$\langle F(x) - F(y), x - y\rangle \geq 0, \quad \forall x, y \in \mathbb{R}^n.$$

(b) strictly monotone if

$$\langle F(x) - F(y), x - y\rangle > 0, \quad \forall x, y \in \mathbb{R}^n.$$

(c) strongly monotone with constant $\eta > 0$ if

$$\langle F(x) - F(y), x - y\rangle \geq \eta\|x - y\|^2, \quad \forall x, y \in \mathbb{R}^n.$$

(d) $F$ is said to be Lipschitz continuous with constant $\gamma$ if

$$\|F(x) - F(y)\| \leq \gamma\|x - y\|, \quad \forall x, y \in \mathbb{R}^n.$$

**Definition 2.2** Let $X$ be a closed convex set in $\mathbb{R}^n$. Then, for each $x \in \mathbb{R}^n$, there is a unique point $y \in X$ such that $\|x - y\| \leq \|x - z\|, \forall z \in X$. Here $y$ is known as the projection of $x$ onto the set $X$ with respect to Euclidean norm, that is,

$$y = P_X(x) = \arg\min_{z \in X} \|x - z\|.$$

The projection function $P_X(x)$ has the following property, called projection theorem [2], which is useful in our subsequent analysis.

**Property 2.1** *Let $X$ be a nonempty closed convex subset of $\mathbb{R}^n$. Then, for each $z \in \mathbb{R}^n$, $P_X(z)$ is the unique vector $\bar{z} \in X$ such that $(y - \bar{z})^T(z - \bar{z}) \leq 0, \ \forall y \in X$.*

Next, we recall some materials about first order differential equations (ODE):

$$\dot{w}(t) = H(w(t)), \qquad w(t_0) = w_0 \in \mathbb{R}^n, \tag{6}$$

where $H : \mathbb{R}^n \to \mathbb{R}^n$ is a mapping. We also introduce three kinds of stability that will be discussed later. These materials can be found in usual ODE textbooks, e.g. [31].

**Definition 2.3** A point $w^* = w(t^*)$ is called an equilibrium point or a steady state of the dynamic system (6) if $H(w^*) = 0$. If there is a neighborhood $\Omega^* \subseteq \mathbb{R}^n$ of $w^*$ such that $H(w^*) = 0$ and $H(w) \neq 0$ $\forall w \in \Omega^* \setminus \{w^*\}$, then $w^*$ is called an isolated equilibrium point.

**Lemma 2.1** *Assume that $H : \mathbb{R}^n \to \mathbb{R}^n$ is a continuous mapping. Then, for any $t_0 > 0$ and $w_0 \in \mathbb{R}^n$, there exists a local solution $w(t)$ for (6) with $t \in [t_0, \tau)$ for some $\tau > t_0$. If, in addition, $H$ is locally Lipschitz continuous at $w_0$, then the solution is unique; if $H$ is Lipschitz continuous in $\mathbb{R}^n$, then $\tau$ can be extended to $\infty$.*

If a local solution defined on $[t_0, \tau)$ cannot be extended to a local solution on a larger interval $[t_0, \tau_1)$, $\tau_1 > \tau$, then it is called a maximal solution, and the interval $[t_0, \tau)$ is the maximal interval of existence. Clearly, any local solution has an extension to a maximal one. We denote $[t_0, \tau(w_0))$ by the maximal interval of existence associated with $w_0$.

**Lemma 2.2** *Assume that $H : \mathbb{R}^n \to \mathbb{R}^n$ is continuous. If $w(t)$ with $t \in [t_0, \tau(w_0))$ is a maximal solution and $\tau(w_0) < \infty$, then* $\lim_{t\uparrow\tau(w_0)} \|w(t)\| = \infty$.

**Definition 2.4** (Lyapunov Stability) Let $w(t)$ be a solution for (6). An isolated equilibrium point $w^*$ is Lyapunov stable if for any $w_0 = w(t_0)$ and any $\varepsilon > 0$, there exists a $\delta > 0$ such that if $\|w(t_0) - w^*\| < \delta$, then $\|w(t) - w^*\| < \varepsilon$ for all $t \geq t_0$.

**Definition 2.5** (Asymptotic Stability) An isolated equilibrium point $w^*$ is said to be asymptotic stable if in addition to being Lyapunov stable, it has the property that if $\|w(t_0) - w^*\| < \delta$, then $w(t) \to w^*$ as $t \to \infty$.

**Definition 2.6** (Lyapunov function) Let $\Omega \subseteq \mathbb{R}^n$ be an open neighborhood of $\bar{w}$. A continuously differentiable function $V : \mathbb{R}^n \to \mathbb{R}$ is said to be a Lyapunov function at the state $\bar{w}$ over the set $\Omega$ for (6) if

$$\begin{cases} V(\bar{w}) = 0, \quad V(w) > 0, \quad \forall w \in \Omega \setminus \{\bar{w}\}, \\ \dot{V}(w) \leq 0, \quad \forall w \in \Omega \setminus \{\bar{w}\}. \end{cases} \tag{7}$$

**Lemma 2.3**

(a) *An isolated equilibrium point $w^*$ is Lyapunov stable if there exists a Lyapunov function over some neighborhood $\Omega^*$ of $w^*$.*
(b) *An isolated equilibrium point $w^*$ is asymptotically stable if there exists a Lyapunov function over some neighborhood $\Omega^*$ of $w^*$ such that $\dot{V}(w) < 0$, $\forall w \in \Omega^* \setminus \{w^*\}$.*

**Definition 2.7** (Exponential Stability) An isolated equilibrium point $w^*$ is exponentially stable if there exists a $\delta > 0$ such that arbitrary point $w(t)$ of (6) with the initial condition $w(t_0) = w_0$ and $\|w(t_0) - w^*\| < \delta$ is well defined on $[t_0, +\infty)$ and satisfies

$$\|w(t) - w^*\| \leq ce^{-\omega t}\|w(t_0) - w^*\| \quad \forall t \geq t_0,$$

where $c > 0$ and $\omega > 0$ are constants independent of the initial point.

## 3 Neural network model based on smoothed Fischer-Burmeister function

The smoothed Fischer-Burmeister function over the second-order cone defined below is used to construct a merit function by which the KKT system of SOCCVI is reformulated as an unconstrained smooth minimization problem. Furthermore, based on the minimization problem, we propose a neural network and study its stability in this section.

For any $a=(a_1;a_2),\ b=(b_1;b_2)\in\mathbb{R}\times\mathbb{R}^{n-1}$, we define their Jordan product as

$$a\cdot b=(a^Tb;\, b_1a_2+a_1b_2).$$

We denote $a^2=a\cdot a$ and $|a|=\sqrt{a^2}$, where for any $b\in\mathcal{K}^n$, $\sqrt{b}$ is the unique vector in $\mathcal{K}^n$ such that $b=\sqrt{b}\cdot\sqrt{b}$.

**Definition 3.1** A function $\phi:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n$ is called an SOC-complementarity function if it satisfies

$$\phi(a,b)=0 \quad\Longleftrightarrow\quad a\cdot b=0,\quad a\in\mathcal{K}^n,\ b\in\mathcal{K}^n.$$

A popular SOC-complementarity function is the Fischer-Burmeister function, which is semismooth [33] and defined as

$$\phi_{\mathrm{FB}}(a,b)=\left(a^2+b^2\right)^{1/2}-(a+b).$$

Then the smoothed Fischer-Burmeister function is given by

$$\phi_{\mathrm{FB}}^{\varepsilon}(a,b)=\left(a^2+b^2+\varepsilon^2e\right)^{1/2}-(a+b) \tag{8}$$

with $\varepsilon\in\mathbb{R}_+$ and $e=(1,0,\dots,0)^T\in\mathbb{R}^n$.

The following lemma gives the gradient of $\phi_{\mathrm{FB}}^{\varepsilon}$. Since the proofs can be found in [14, 33, 37], we here omit them.

**Lemma 3.1** *Let $\phi_{\mathrm{FB}}^{\varepsilon}$ be defined as in (8) and $\varepsilon\neq 0$. Then, $\phi_{\mathrm{FB}}^{\varepsilon}$ is continuously differentiable everywhere and*

$$\nabla_{\varepsilon}\phi_{\mathrm{FB}}^{\varepsilon}(a,b)=e^TL_z^{-1}L_{\varepsilon e},\qquad \nabla_a\phi_{\mathrm{FB}}^{\varepsilon}(a,b)=L_z^{-1}L_a-I,$$
$$\nabla_b\phi_{\mathrm{FB}}^{\varepsilon}(a,b)=L_z^{-1}L_b-I,$$

*where $z=\left(a^2+b^2+\varepsilon^2e\right)^{1/2}$, $I$ is identity mapping and $L_a=\left[\begin{smallmatrix}a_1 & a_2^T\\ a_2 & a_1I_{n-1}\end{smallmatrix}\right]$ for $a=(a_1;a_2)\in\mathbb{R}\times\mathbb{R}^{n-1}$.*

Using Definition 3.1 and KKT condition described in [37], we can see that the KKT system (5) is equivalent to the following unconstrained smooth minimization problem:

$$\min\Psi(w):=\frac{1}{2}\|S(w)\|^2. \tag{9}$$

Here $\Psi(w)$, $w = (\varepsilon, x, \mu, \lambda) \in \mathbb{R}^{1+n+l+m}$, is a merit function, and $S(w)$ is defined by

$$S(w) = \begin{pmatrix} \varepsilon \\ L(x, \mu, \lambda) \\ -h(x) \\ \phi_{\mathrm{FB}}^{\varepsilon}(-g_{m_1}(x), \lambda_{m_1}) \\ \vdots \\ \phi_{\mathrm{FB}}^{\varepsilon}(-g_{m_p}(x), \lambda_{m_p}) \end{pmatrix},$$

with $g_{m_i}(x), \lambda_{m_i} \in \mathbb{R}^{m_i}$. In other words, $\Psi(w)$ given in (9) is a smooth merit function for the KKT system (5).

Based on the above smooth minimization problem (9), it is natural to propose the first neural network for solving the KKT system (5) of SOCCVI problem:

$$\frac{dw(t)}{dt} = -\rho \nabla \Psi(w(t)), \qquad w(t_0) = w_0, \tag{10}$$

where $\rho > 0$ is a scaling factor.

*Remark 3.1* In fact, we can also adopt another merit function which is based on the FB function without the element $\varepsilon$. That is, we can define

$$S(x, \mu, \lambda) = \begin{pmatrix} L(x, \mu, \lambda) \\ -h(x) \\ \phi_{\mathrm{FB}}(-g_{m_1}(x), \lambda_{m_1}) \\ \vdots \\ \phi_{\mathrm{FB}}(-g_{m_p}(x), \lambda_{m_p}) \end{pmatrix}. \tag{11}$$

Then, the neural network model (10) could be obtained as well because $\|\phi_{\mathrm{FB}}\|^2$ is smooth [7]. However, it is observed that the gradient mapping $\nabla\Psi$ has more complicated formulas because $(-g_{m_i}(x))^2 + \lambda_{m_i}^2$ may lie on the boundary of SOC, or interior of SOC, see [7, 33], which will cost more expensive numerical computations. Thus, the one dimensional parameter $\varepsilon$ in use not only has no influence on the main result, but also will simplify the computational work.

To discuss properties of the neural network model (10), we make the following assumption which is used to avoid the singularity of $\nabla S(w)$, see [37].

**Assumption 3.1**

(a) *the gradients* $\{\nabla h_j(x) | j = 1, \dots, l\} \cup \{\nabla g_i(x) | i = 1, \dots, m\}$ *are linear independent.*

(b) $\nabla_x L(x, \mu, \lambda)$ *is positive definite on the null space of the gradients* $\{\nabla h_j(x) | j = 1, \dots, l\}$.

When SOCCVI problem corresponds to the KKT conditions of a convex second-order cone program (CSOCP) problem as (4) where both $h$ and $g$ are linear, the above

Assumption 3.1(b) is indeed equivalent to the well-used condition $\nabla^2 f(x)$ is positive definite, e.g. [42, Corollary 1].

**Proposition 3.1** *Let* $\Psi : \mathbb{R}^{1+n+l+m} \to \mathbb{R}_+$ *be defined as in* (9). *Then,* $\Psi(w) \geq 0$ *for* $w = (\varepsilon, x, \mu, \lambda) \in \mathbb{R}^{1+n+l+m}$ *and* $\Psi(w) = 0$ *if and only if* $(x, \mu, \lambda)$ *solves the KKT system* (5).

*Proof* The proof is straightforward. □

**Proposition 3.2** *Let* $\Psi : \mathbb{R}^{1+n+l+m} \to \mathbb{R}_+$ *be defined as in* (9). *Then, the following results hold.*

(a) *The function* $\Psi$ *is continuously differentiable everywhere with*

$$\nabla\Psi(w) = \nabla S(w) S(w),$$

*where*

$$\nabla S(w) = \begin{bmatrix} 1 & 0 & 0 & \operatorname{diag}\{\nabla_\varepsilon \phi^\varepsilon_{\mathrm{FB}}(-g_{m_i}(x), \lambda_{m_i})\}_{i=1}^p \\ 0 & \nabla_x L(x,\mu,\lambda)^T & -\nabla h(x) & -\nabla g(x)\operatorname{diag}\{\nabla_{g_{m_i}} \phi^\varepsilon_{\mathrm{FB}}(-g_{m_i}(x), \lambda_{m_i})\}_{i=1}^p \\ 0 & \nabla h(x)^T & 0 & 0 \\ 0 & \nabla g(x)^T & 0 & \operatorname{diag}\{\nabla_{\lambda_{m_i}} \phi^\varepsilon_{\mathrm{FB}}(-g_{m_i}(x), \lambda_{m_i})\}_{i=1}^p \end{bmatrix}.$$

(b) *Suppose that Assumption* 3.1 *holds. Then,* $\nabla S(w)$ *is nonsingular for any* $w \in \mathbb{R}^{1+n+l+m}$. *Moreover, if* $(0, x, \mu, \lambda) \in \mathbb{R}^{1+n+l+m}$ *is a stationary point of* $\Psi$, *then* $(x, \mu, \lambda) \in \mathbb{R}^{n+l+m}$ *is a KKT triple of the SOCCVI problem.*
(c) $\Psi(w(t))$ *is nonincreasing with respect to* $t$.

*Proof* Part(a) follows from the chain rule. For part(b), we know that $\nabla S(w)$ is nonsingular if and only if the following matrix

$$\begin{bmatrix} \nabla_x L(x,\mu,\lambda)^T & -\nabla h(x) & -\nabla g(x)\operatorname{diag}\{\nabla_{g_{m_i}} \phi^\varepsilon_{\mathrm{FB}}(-g_{m_i}(x), \lambda_{m_i})\}_{i=1}^p \\ \nabla h(x)^T & 0 & 0 \\ \nabla g(x)^T & 0 & \operatorname{diag}\{\nabla_{\lambda_{m_i}} \phi^\varepsilon_{\mathrm{FB}}(-g_{m_i}(x), \lambda_{m_i})\}_{i=1}^p \end{bmatrix}$$

is nonsingular. In fact, from [37, Theorem 3.1] and [37, Proposition 4.1], the above matrix is nonsingular and $(x, \mu, \lambda) \in \mathbb{R}^{n+l+m}$ is a KKT triple of the SOCCVI problem if $(0, x, \mu, \lambda) \in \mathbb{R}^{1+n+l+m}$ is a stationary point of $\Psi$. It remains to show part(c). By the definition of $\Psi(w)$ and (10), it is not difficult to compute

$$\frac{d\Psi(w(t))}{dt} = \nabla\Psi(w(t))^T \frac{dw(t)}{dt} = -\rho\|\nabla\Psi(w(t))\|^2 \leq 0. \tag{12}$$

Therefore, $\Psi(w(t))$ is a monotonically decreasing function with respect to $t$. □

Now, we are ready to analyze the behavior of the solution trajectory of (10) and establish properties of three kinds of stability for an isolated equilibrium point.

**Proposition 3.3**

(a) *If* $(x, \mu, \lambda) \in \mathbb{R}^{n+l+m}$ *is a KKT triple of SOCCVI problem, then* $(0, x, \mu, \lambda) \in \mathbb{R}^{1+n+l+m}$ *is an equilibrium point of* (10).
(b) *If Assumption* 3.1 *holds and* $(0, x, \mu, \lambda) \in \mathbb{R}^{1+n+l+m}$ *is an equilibrium point of* (10), *then* $(x, \mu, \lambda) \in \mathbb{R}^{n+l+m}$ *is a KKT triple of SOCCVI problem.*

*Proof* (a) From Proposition 3.1 and $(x, \mu, \lambda) \in \mathbb{R}^{n+l+m}$ being a KKT triple of SOCCVI problem, it is clear that $S(0, x, \mu, \lambda) = 0$. Hence, $\nabla\Psi(0, x, \mu, \lambda) = 0$. Besides, by Proposition 3.2, we know that if $\varepsilon \neq 0$, then $\nabla\Psi(\varepsilon, x, \mu, \lambda) \neq 0$. This shows that $(0, x, \mu, \lambda)$ is an equilibrium point of (10).

(b) It follows from $(0, x, \mu, \lambda) \in \mathbb{R}^{1+n+l+m}$ being an equilibrium point of (10) that $\nabla\Psi(0, x, \mu, \lambda) = 0$. In other words, $(0, x, \mu, \lambda)$ is the stationary point of $\Psi$. Then, the result is a direct consequence of Proposition 3.2(b). □

**Proposition 3.4**

(a) *For any initial state* $w_0 = w(t_0)$, *there exists exactly one maximal solution* $w(t)$ *with* $t \in [t_0, \tau(w_0))$ *for the neural network* (10).
(b) *If the level set* $\mathcal{L}(w_0) = \{w \in \mathbb{R}^{1+n+l+m} | \Psi(w) \leq \Psi(w_0)\}$ *is bounded, then* $\tau(w_0) = +\infty$.

*Proof* (a) Since $S$ is continuous differentiable, $\nabla S$ is continuous, and therefore, $\nabla S$ is bounded on a local compact neighborhood of $w$. That means $\nabla\Psi(w)$ is locally Lipschitz continuous. Thus, applying Lemma 2.1 leads to the desired result.

(b) This proof is similar to the proof of Case(i) in [5, Proposition 4.2]. □

*Remark 3.2* We wish to obtain the result that the level sets

$$\mathcal{L}(\Psi, \gamma) := \{w \in \mathbb{R}^{1+n+l+m} \mid \Psi(w) \leq \gamma\}$$

are bounded for all $\gamma \in \mathbb{R}$. However, we are not able to complete the argument. We suspect that there needs more subtle properties of $F$, $h$ and $g$ to finish it.

Next, we investigate the convergence of the solution trajectory of neural network (10).

**Theorem 3.1**

(a) *Let* $w(t)$ *with* $t \in [t_0, \tau(w_0))$ *be the unique maximal solution to* (10). *If* $\tau(w_0) = +\infty$ *and* $\{w(t)\}$ *is bounded, then* $\lim_{t\to\infty} \nabla\Psi(w(t)) = 0$.
(b) *If Assumption* 3.1 *holds and* $(\varepsilon, x, \mu, \lambda) \in \mathbb{R}^{1+n+l+m}$ *is the accumulation point of the trajectory* $w(t)$, *then* $(x, \mu, \lambda) \in \mathbb{R}^{n+l+m}$ *is a KKT triple of SOCCVI problem.*

*Proof* With Proposition 3.2(b) and (c) and Proposition 3.3, the arguments are exactly the same as those for [28, Corollary 4.3]. Thus, we omit them. □

**Theorem 3.2** *Let* $w^*$ *be an isolated equilibrium point of the neural network* (10). *Then the following results hold.*

(a) $w^*$ *is asymptotically stable.*
(b) *If Assumption* 3.1 *holds, then it is exponentially stable.*

*Proof* Since $w^*$ is an isolated equilibrium point of (10), there exists a neighborhood $\Omega^* \subseteq \mathbb{R}^{1+n+l+m}$ of $w^*$ such that

$$\nabla\Psi(w^*) = 0 \quad \text{and} \quad \nabla\Psi(w) \neq 0 \quad \forall w \in \Omega^* \setminus \{w^*\}.$$

Next, we argue that $\Psi(w)$ is indeed a Lyapunov function at $x^*$ over the set $\Omega^*$ for (10) by showing that the conditions in (7) are satisfied. First, notice that $\Psi(w) \geq 0$. Suppose that there is an $\bar{w} \in \Omega^* \setminus \{w^*\}$ such that $\Psi(\bar{w}) = 0$. Then, we can easily obtain that $\nabla\Psi(\bar{w}) = 0$, i.e., $\bar{w}$ is also an equilibrium point of (10), which clearly contradicts the assumption that $w^*$ is an isolated equilibrium point in $\Omega^*$. Thus, we prove that $\Psi(w) > 0$ for any $w \in \Omega^* \setminus \{w^*\}$. This together with (12) shows that the condition in (7) are satisfied. Because $w^*$ is isolated, from (12), we have

$$\frac{d\Psi(w(t))}{dt} < 0, \quad \forall w(t) \in \Omega^* \setminus \{w^*\}.$$

This implies that $w^*$ is asymptotically stable. Furthermore, if Assumption 3.1 holds, we can obtain that $\nabla S$ is nonsingular. In addition, we have

$$S(w) = S(w^*) + \nabla S(w^*)(w - w^*) + o(\|w - w^*\|), \quad \forall w \in \Omega^* \setminus \{w^*\}. \tag{13}$$

From $\|S(w(t))\|$ being a monotonically decreasing function with respect to $t$ and (13), we can deduce that

$$\begin{aligned}
\|w(t) - w^*\| &\leq \|(\nabla S(w^*))^{-1}\| \|S(w(t)) - S(w^*)\| + o(\|w(t) - w^*\|) \\
&\leq \|(\nabla S(w^*))^{-1}\| \|S(w(t_0)) - S(w^*)\| + o(\|w(t) - w^*\|) \\
&\leq \|(\nabla S(w^*))^{-1}\| \left[\|(\nabla S(w^*))\| \|w(t_0) - w^*\| + o(\|w(t_0) - w^*\|)\right] \\
&\quad + o(\|w(t) - w^*\|).
\end{aligned}$$

That is,

$$\begin{aligned}
\|w(t) - w^*\| - o(\|w(t) - w^*\|) \leq \|(\nabla S(w^*))^{-1}\| \big[&\|(\nabla S(w^*))\| \|w(t_0) - w^*\| \\
&+ o(\|w(t_0) - w^*\|)\big].
\end{aligned}$$

The above inequality implies that the neural network (10) is also exponentially stable. □

## 4 Neural network model based on projection function

In this section, we present that the KKT triple of SOCCVI problem is equivalent to the solution of a projection formulation. Based on this, we introduce another neural network model for solving the projection formulation and analyze the stability conditions and convergence.

Define the function $U : \mathbb{R}^{n+l+m} \to \mathbb{R}^{n+l+m}$ and vector $w$ in the following form:

$$U(w) = \begin{pmatrix} L(x, \mu, \lambda) \\ -h(x) \\ -g(x) \end{pmatrix}, \quad w = \begin{pmatrix} x \\ \mu \\ \lambda \end{pmatrix}, \tag{14}$$

where $L(x, \mu, \lambda) = F(x) + \nabla h(x)\mu + \nabla g(x)\lambda$ is the Lagrange function. To avoid confusion, we emphasize that, for any $w \in \mathbb{R}^{n+l+m}$, we have

$$\begin{aligned} &w_i \in \mathbb{R}, \quad \text{if } 1 \le i \le n + l, \\ &w_i \in \mathbb{R}^{m_{i-(n+l)}}, \quad \text{if } n + l + 1 \le i \le n + l + p. \end{aligned}$$

Then, we may write (14) as

$$\begin{aligned} &U_i = (U(w))_i = (L(x, \mu, \lambda))_i, \quad w_i = x_i,\ i = 1, \dots, n, \\ &U_{n+j} = (U(w))_{n+j} = -h_j(x), \quad w_{n+j} = \mu_j,\ j = 1, \dots, l, \\ &U_{n+l+k} = (U(w))_{n+l+k} = -g_k(x) \in \mathbb{R}^{m_k}, \quad w_{n+l+k} = \lambda_k \in \mathbb{R}^{m_k},\ k = 1, \dots, p, \\ &\sum_{k=1}^{p} m_k = m. \end{aligned}$$

With this, the KKT conditions (5) can be recast as

$$\begin{aligned} &U_i = 0, \quad i = 1, 2, \dots, n, n + 1, \dots, n + l, \\ &\langle U_J, w_J \rangle = 0, \quad U_J = (U_{n+l+1}, U_{n+l+2}, \dots, U_{n+l+p})^T \in \mathcal{K}, \\ &w_J = (w_{n+l+1}, w_{n+l+2}, \dots, w_{n+l+p})^T \in \mathcal{K}. \end{aligned} \tag{15}$$

Thus, $(x^*, \mu^*, \lambda^*)$ is a KKT triple for (1) if and only if $(x^*, \mu^*, \lambda^*)$ is a solution to (15).

It is well known that the nonlinear complementarity problem, which is denoted by NCP$(F, K)$ and to find an $x \in \mathbb{R}^n$ such that

$$x \in K, \quad F(x) \in K \quad \text{and} \quad \langle F(x), x \rangle = 0$$

where $K$ is a closed convex set of $\mathbb{R}^n$, is equivalent to the following VI$(F, K)$ problem: finding an $x \in K$ such that

$$\langle F(x), y - x \rangle \ge 0 \quad \forall y \in K.$$

Furthermore, if $K = \mathbb{R}^n$, then NCP$(F, K)$ becomes the system of nonlinear equations

$$F(x) = 0.$$

Based on the above, solution of (15) is equivalent to solution of the following VI problem: find $w \in \mathbb{K}$ such that

$$\langle U(w), v - w \rangle \ge 0, \quad \forall v \in \mathbb{K}, \tag{16}$$

where $\mathbb{K} = \mathbb{R}^{n+l} \times \mathcal{K}$. In addition, by applying the Property 2.1, its solution is equivalent to solution of below projection formulation

$$P_{\mathbb{K}}(w - U(w)) = w \quad \text{with } \mathbb{K} = \mathbb{R}^{n+l} \times \mathcal{K}, \tag{17}$$

where function $U$ and vector $w$ are defined in (14). Now, according to (17), we give the following neural network:

$$\frac{dw}{dt} = \rho\{P_{\mathbb{K}}(w - U(w)) - w\}, \tag{18}$$

where $\rho > 0$. Note that $\mathbb{K}$ is a closed and convex set. For any $w \in \mathbb{R}^{n+l+m}$, $P_{\mathbb{K}}$ means

$$\begin{aligned} P_{\mathbb{K}}(w) = [&P_{\mathbb{K}}(w_1), P_{\mathbb{K}}(w_2), \ldots, P_{\mathbb{K}}(w_{n+l}), P_{\mathbb{K}}(w_{n+l+1}), P_{\mathbb{K}}(w_{n+l+2}), \ldots, \\ &P_{\mathbb{K}}(w_{n+l+p})], \end{aligned}$$

where

$$\begin{aligned} &P_{\mathbb{K}}(w_i) = w_i, \quad i = 1, \ldots, n+l, \\ &P_{\mathbb{K}}(w_{n+l+j}) = [\lambda_1(w_{n+l+j})]_+ \cdot u^{(1)}_{w_{n+l+j}} + [\lambda_2(w_{n+l+j})]_+ \cdot u^{(2)}_{w_{n+l+j}}, \quad j = 1, \ldots, p. \end{aligned}$$

Here, for the sake of simplicity, we denote the vector $w_{n+l+j}$ by $v$ for the moment, and $[\cdot]_+$ is the scalar projection, $\lambda_1(v)$, $\lambda_2(v)$ and $u_v^{(1)}$, $u_v^{(2)}$ are the spectral values and the associated spectral vectors of $v = (v_1; v_2) \in \mathbb{R} \times \mathbb{R}^{m_j - 1}$, respectively, given by

$$\begin{cases} \lambda_i(v) = v_1 + (-1)^i \|v_2\|, \\ u_v^{(i)} = \frac{1}{2}\big(1, (-1)^i \frac{v_2}{\|v_2\|}\big), \end{cases}$$

for $i = 1, 2$, see [7, 33].

The dynamic system described by (18) can be recognized as a recurrent neural network with a single-layer structure. To analyze the stability conditions of (18), we need the following lemmas and proposition.

**Lemma 4.1** *If the gradient of $L(x, \mu, \lambda)$ is positive semi-definite (respectively, positive definite), then the gradient of $U$ in (14) is positive semi-definite (respectively, positive definite).*

*Proof* Since we have

$$\nabla U(x, \mu, \lambda) = \begin{bmatrix} \nabla_x L^T(x, \mu, \lambda) & -\nabla h(x) & -\nabla g(x) \\ \nabla^T h(x) & 0 & 0 \\ \nabla^T g(x) & 0 & 0 \end{bmatrix},$$

for any nonzero vector $d = (p^T, q^T, r^T)^T \in \mathbb{R}^{n+l+m}$, we can obtain that

$$d^T \nabla U(x, \mu, \lambda) d = \begin{pmatrix} p^T & q^T & r^T \end{pmatrix} \begin{bmatrix} \nabla_x L^T(x, \mu, \lambda) & -\nabla h(x) & -\nabla g(x) \\ \nabla^T h(x) & 0 & 0 \\ \nabla^T g(x) & 0 & 0 \end{bmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix}$$

$$= p^T \nabla_x L(x, \mu, \lambda) p.$$

This leads to the desired results. □

**Proposition 4.1** *For any initial point $w_0 = (x_0, \mu_0, \lambda_0)$ with $\lambda_0 := \lambda(t_0) \in \mathcal{K}$, there exist a unique solution $w(t) = (x(t), \mu(t), \lambda(t))$ for neural network* (18), *Moreover, $\lambda(t) \in \mathcal{K}$.*

*Proof* For simplicity, we assume $\mathcal{K} = \mathcal{K}^m$. The analysis can be carried over to the general case where $\mathcal{K}$ is the Cartesian product of second-order cones. Since $F, h, g$ are continuous differentiable, the function

$$F(w) := P_{\mathbb{K}}(w - U(w)) - w \quad \text{with } \mathbb{K} = \mathbb{R}^{n+l} \times \mathcal{K}^m \tag{19}$$

is semi-smooth and Lipschitz continuous. Thus, there exists a unique solution $w(t) = (x(t), \mu(t), \lambda(t))$ for neural network (18). Therefore, it remains to show that $\lambda(t) \in \mathcal{K}^m$. For convenience, we denote $\lambda(t) := (\lambda_1(t), \lambda_2(t)) \in \mathbb{R} \times \mathbb{R}^{m-1}$. To complete the proof, we need to verify two things: (i) $\lambda_1(t) \geq 0$ and (ii) $\|\lambda_2(t)\| \leq \lambda_1(t)$. First, from (18), we have

$$\frac{d\lambda}{dt} + \rho\lambda(t) = \rho P_{\mathcal{K}^m}(\lambda + g(x)).$$

The solution of the above first-order ordinary differential equation is

$$\lambda(t) = e^{-\rho(t-t_0)}\lambda(t_0) + \rho e^{-\rho t} \int_{t_0}^{t} \rho e^{\rho s} P_{\mathcal{K}^m}(\lambda + g(x)) ds. \tag{20}$$

If we let $\lambda(t_0) := (\lambda_1(t_0), \lambda_2(t_0)) \in \mathbb{R} \times \mathbb{R}^{m-1}$ and denote $P_{\mathcal{K}^m}(\lambda + g(x))$ as $z(t_0) := (z_1(t_0), z_2(t_0))$, then (20) leads to

$$\lambda_1(t) = e^{-\rho(t-t_0)}\lambda_1(t_0) + \rho e^{-\rho t} \int_{t_0}^{t} \rho e^{\rho s} z_1(s) ds, \tag{21}$$

$$\lambda_2(t) = e^{-\rho(t-t_0)}\lambda_2(t_0) + \rho e^{-\rho t} \int_{t_0}^{t} \rho e^{\rho s} z_2(s) ds. \tag{22}$$

Due to both $\lambda(t_0)$ and $z(t)$ belong to $\mathcal{K}^m$, there have $\lambda_1(t_0) \geq 0$, $\|\lambda_2(t_0)\| \leq \lambda_1(t_0)$ and $\|z_2(t)\| \leq z_1(t)$. Therefore, $\lambda_1(t) \geq 0$ since both terms in the right-hand side of (21) are nonnegative. In addition,

$$\begin{aligned}
\|\lambda_2(t)\| &\leq e^{-\rho(t-t_0)}\|\lambda_2(t_0)\| + \rho e^{-\rho t} \int_{t_0}^{t} \rho e^{\rho s} \|z_2(s)\| ds \\
&\leq e^{-\rho(t-t_0)}\lambda_1(t_0) + \rho e^{-\rho t} \int_{t_0}^{t} \rho e^{\rho s} z_1(s) ds \\
&= \lambda_1(t),
\end{aligned}$$

which implies that $\lambda(t) \in \mathcal{K}^m$. □

**Lemma 4.2** *Let $U(w)$, $F(w)$ be defined as in (14) and (19), respectively. Suppose $w^* = (x^*, \mu^*, \lambda^*)$ is an equilibrium point of neural network (18) with $(x^*, \mu^*, \lambda^*)$ being an KKT triple of SOCCVI problem. Then, the following inequality holds*:

$$(F(w) + w - w^*)^T(-F(w) - U(w)) \geq 0. \tag{23}$$

*Proof* Notice that

$$\begin{aligned}
&(F(w) + w - w^*)^T(-F(w) - U(w)) \\
&\quad = [-w + P_{\mathbb{K}}(w - U(w)) + w - w^*]^T[w - P_{\mathbb{K}}(w - U(w)) - U(w)] \\
&\quad = [-w^* + P_{\mathbb{K}}(w - U(w))]^T[w - P_{\mathbb{K}}(w - U(w)) - U(w)] \\
&\quad = -[w^* - P_{\mathbb{K}}(w - U(w))]^T[w - U(w) - P_{\mathbb{K}}(w - U(w))].
\end{aligned}$$

Since $w^* \in \mathbb{K}$, applying Property 2.1 gives

$$[w^* - P_{\mathbb{K}}(w - U(w))]^T[w - U(w) - P_{\mathbb{K}}(w - U(w))] \leq 0.$$

Thus, we have

$$(F(w) + w - w^*)^T(-F(w) - U(w)) \geq 0.$$

This completes the proof. □

We now show the stability and convergence issues regarding neural network (18).

**Theorem 4.1** *If $\nabla_x L(w)$ is positive semi-definite (respectively, positive definite), the solution of neural network (18) with initial point $w_0 = (x_0, \mu_0, \lambda_0)$ where $\lambda_0 \in \mathcal{K}$ is Lyapunov stable (respectively, asymptotically stable). Moreover, the solution trajectory of neural network (18) is extendable to the global existence.*

*Proof* Again, for simplicity, we assume $\mathcal{K} = \mathcal{K}^m$. From Proposition 4.1, there exists a unique solution $w(t) = (x(t), \mu(t), \lambda(t))$ for neural network (18) and $\lambda(t) \in \mathcal{K}^m$. Let $w^* = (x^*, \mu^*, \lambda^*)$ be an equilibrium point of neural network (18). We define a Lyapunov function as below:

$$V(w) := V(x, \mu, \lambda) := -U(w)^T F(w) - \frac{1}{2}\|F(w)\|^2 + \frac{1}{2}\|w - w^*\|^2. \tag{24}$$

From [12, Theorem 3.2], we know that $V$ is continuously differentiable with

$$\nabla V(w) = U(w) - [\nabla U(w) - I]F(w) + (w - w^*).$$

It is also trivial that $V(w^*) = 0$. Then, we have

$$\frac{dV(w(t))}{dt} = \nabla V(w(t))^T \frac{dw}{dt}$$

$$
\begin{aligned}
&= \{U(w) - [\nabla U(w) - I]F(w) + (w - w^*)\}^T \rho F(w) \\
&= \rho\{[U(w) + (w - w^*)]^T F(w) + \|F(w)\|^2 - F(w)^T \nabla U(w) F(w)\}.
\end{aligned}
$$

Inequality (23) in Lemma 4.2 implies

$$
[U(w) + (w - w^*)]^T F(w) \le -U(w)^T (w - w^*) - \|F(w)\|^2,
$$

which yields

$$
\begin{aligned}
&\frac{dV(w(t))}{dt} \\
&\quad \le \rho\{-U(w)^T (w - w^*) - F(w)^T \nabla U(w) F(w)\} \\
&\quad = \rho\{-U(w^*)^T (w - w^*) - (U(w) - U(w^*))^T (w - w^*) \\
&\qquad - F(w)^T \nabla U(w) F(w)\}. 
\end{aligned} \tag{25}
$$

Note that $w^*$ is the solution of the variational inequality (16). Since $w \in \mathbb{K}$, we therefore obtain $-U(w^*)^T(w - w^*) \le 0$. Because $U(w)$ is continuous differentiable and $\nabla U(w)$ is positive semi-definite, by [32, Theorem 5.4.3], we obtain that $U(w)$ is monotone. Hence, we have $-(U(w) - U(w^*))^T(w - w^*) \le 0$ and $-F(w)^T \nabla U(w) F(w) \le 0$. The above discussions lead to $\frac{dV(w(t))}{dt} \le 0$. Also, by [32, Theorem 5.4.3], we know that if $\nabla U(w)$ is positive definite, then $U(w)$ is strictly monotone, which implies $\frac{dV(w(t))}{dt} < 0$ in this case.

In order to obtain $V(w)$ is a Lyapunov function and $w^*$ is Lyapunov stable, we will show the following inequality:

$$
-U(w)^T F(w) \ge \|F(w)\|^2. \tag{26}
$$

To see this, we first observe that

$$
\|F(w)\|^2 + U(w)^T F(w) = [w - P_{\mathbb{K}}(w - U(w))]^T [w - U(w) - P_{\mathbb{K}}(w - U(w))].
$$

Since $w \in \mathbb{K}$, applying Property 2.1 again, there holds

$$
[w - P_{\mathbb{K}}(w - U(w))]^T [w - U(w) - P_{\mathbb{K}}(w - U(w))] \le 0,
$$

which yields the desired inequality (26). By combining (24) and (26), we have

$$
V(w) \ge \frac{1}{2}\|F(w)\|^2 + \frac{1}{2}\|w - w^*\|^2,
$$

which says $V(w) > 0$ if $w \ne w^*$. Hence $V(w)$ is indeed a Lyapunov function and $w^*$ is Lyapunov stable. Furthermore, if $\nabla_x L(w)$ is positive definite, we have $w^*$ is asymptotically stable. Moreover, it holds that

$$
V(w_0) \ge V(w) \ge \frac{1}{2}\|w - w^*\|^2 \quad \text{for } t \ge t_0, \tag{27}
$$

which tells us the solution trajectory $w(t)$ is bounded. Hence, it can be extended to global existence. □

**Theorem 4.2** *Let $w^* = (x^*, \mu^*, \lambda^*)$ be an equilibrium point of (18). If $\nabla_x L(w)$ is positive definite, the solution of neural network (18) with initial point $w_0 = (x_0, \mu_0, \lambda_0)$ where $\lambda_0 \in \mathcal{K}$ is globally convergent to $w^*$ and has finite convergence time.*

*Proof* From (27), the level set

$$\mathcal{L}(w_0) := \{w | V(w) \le V(w_0)\}$$

is bounded. Then, the Invariant Set Theorem [15] implies the solution trajectory $w(t)$ converges to $\theta$ as $t \to +\infty$ where $\theta$ is the largest invariant set in

$$\Lambda = \left\{ w \in \mathcal{L}(w_0) \;\middle|\; \frac{dV(w(t))}{dt} = 0 \right\}.$$

We will show that $dw/dt = 0$ if and only if $dV(w(t))/dt = 0$ which yields that $w(t)$ converges globally to the equilibrium point $w^* = (x^*, \mu^*, \lambda^*)$. Suppose $dw/dt = 0$, then it is clear that $dV(w(t))/dt = \nabla V(w)^T(dw/dt) = 0$. Let $\hat{w} = (\hat{x}, \hat{\mu}, \hat{\lambda}) \in \Lambda$ which says $dV(\hat{w}(t))/dt = 0$. From (23), we know that

$$dV(\hat{w}(t))/dt \le \rho\{(-U(\hat{w}) - U(w^*))^T(\hat{w} - w^*) - F(\hat{w})^T \nabla U(\hat{w}) F(\hat{w})\}.$$

Both terms inside the big parenthesis are nonpositive as shown in Theorem 4.1, so $(U(\hat{w}) - U(w^*))^T(\hat{w} - w^*) = 0$, $F(\hat{w})^T \nabla U(\hat{w}) F(\hat{w}) = 0$. The condition $\nabla_x L(w)$ being positive definite leads to $\nabla U(\hat{w})$ being positive definite. Hence,

$$F(\hat{w}) = -\hat{w} + P_{\mathbb{K}}(\hat{w} - U(\hat{w})) = 0,$$

which is equivalent to $d\hat{w}/dt = 0$. From the above, $w(t)$ converges globally to the equilibrium point $w^* = (x^*, \mu^*, \lambda^*)$. Moreover, with Theorem 4.1 and following the same argument as in [42, Theorem 2], the neural network (18) has finite convergence time. □

## 5 Simulations

To demonstrate effectiveness of the proposed neural networks, some illustrative SOC-CVI problems are tested. The numerical implementation is coded by Matlab 7.0 and the ordinary differential equation solver adopted is *ode23*, which uses Runge-Kutta (2), (3) formula. In the following tests, the parameter $\rho$ in both neural networks is set to be 1000.

*Example 5.1* Consider the SOCCVI problem (1)–(2) where

$$F(x)=\begin{pmatrix} 2x_1+x_2+1\\ x_1+6x_2-x_3-2\\ -x_2+3x_3-\frac{6}{5}x_4+3\\ -\frac{6}{5}x_3+2x_4+\frac{1}{2}\sin x_4\cos x_5\sin x_6+6\\ \frac{1}{2}\cos x_4\sin x_5\sin x_6+2x_5-\frac{5}{2}\\ -\frac{1}{2}\cos x_4\cos x_5\cos x_6+2x_6+\frac{1}{4}\cos x_6\sin x_7\cos x_8+1\\ \frac{1}{4}\sin x_6\cos x_7\cos x_8+4x_7-2\\ -\frac{1}{4}\sin x_6\sin x_7\sin x_8+2x_8+\frac{1}{2} \end{pmatrix}$$

and

$$C=\{x\in\mathbb{R}^8:-g(x)=x\in\mathcal{K}^3\times\mathcal{K}^3\times\mathcal{K}^2\}.$$

This problem has an approximate solution

$$x^*=(0.3820, 0.1148, -0.3644, 0.0000, 0.0000, 0.0000, 0.5000, -0.2500)^T.$$

It can be verified that the Lagrangian function for this example is

$$L(x,\mu,\lambda)=F(x)-\lambda$$

and the gradient of the Lagrangian function is

$$\nabla L(x,\mu,\lambda)=\begin{bmatrix}\nabla F(x)\\ I_{8\times 8}\end{bmatrix},$$

where $I$ is the identity mapping and $\nabla F(x)$ means the gradient of $F(x)$. We use the proposed neural networks with smoothed FB and projection functions, respectively, to solve the problem whose trajectories are depicted in Figs. 1 and 2. The simulation results show that both trajectories are globally convergent to $x^*$ and the neural network with projection function converges to $x^*$ quicker than that with smoothed FB function.

*Example 5.2* [37, Example 5.1] We consider the following SOCCVI problem:

$$\left\langle\frac{1}{2}Dx, y-x\right\rangle\geq 0,\quad \forall y\in C$$

where

$$C=\{x\in\mathbb{R}^n: Ax-a=0,\ Bx-b\preceq 0\},$$

$D$ is an $n\times n$ symmetric matrix, $A$ and $B$ are $l\times n$ and $m\times n$ matrices, respectively, $d$ is an $n\times 1$ vector, $a$ and $b$ are $l\times 1$ and $m\times 1$ vectors with $l+m\leq n$, respectively.

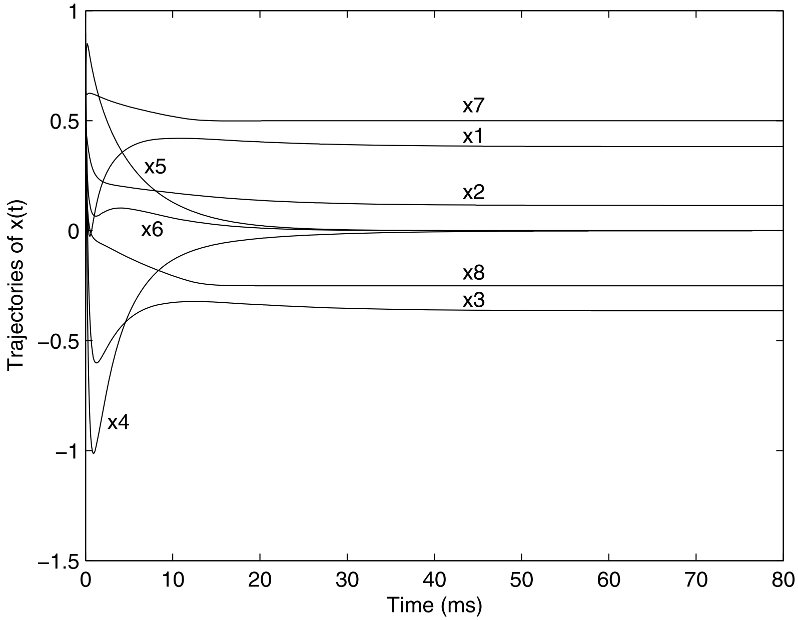


**Fig. 1** Transient behavior of neural network with smoothed FB function in Example 5.1

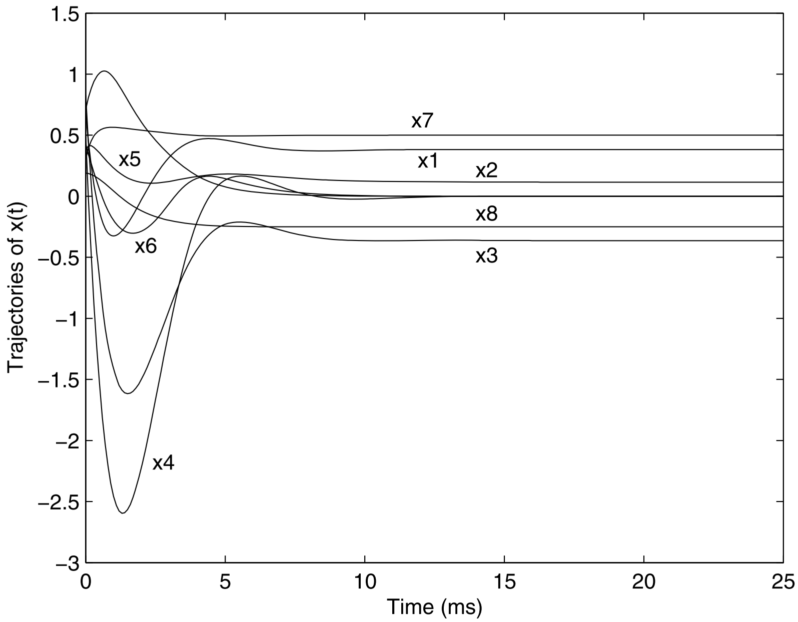


**Fig. 2** Transient behavior of the neural network with projection function in Example 5.1

In fact, we can determine the data $a, b, A, B$ and $D$ randomly. However, as in [37, Example 5.1], we set the data as follows:

$$D=[D_{ij}]_{n\times n},\quad \text{where } D_{ij}=\begin{cases}2, & i=j,\\ 1, & |i-j|=1,\\ 0, & \text{otherwise},\end{cases}$$

$A=[I_{l\times l}\ \ 0_{l\times(n-l)}]_{l\times n}$, $B=[0_{m\times(n-m)}\ \ I_{m\times m}]_{m\times n}$, $a=0_{l\times 1}$, $b=(e_{m_1}, e_{m_2}, \dots, e_{m_p})^T$, where $e_{m_i}=(1,0,\dots,0)^T\in\mathbb{R}^{m_i}$ and $l+m\le n$. Clearly, $A$ and $B$ are full row rank and $\operatorname{rank}([A^T\ \ B^T])=l+m$.

In the simulation, the parameters $l, m$, and $n$ are set to be 3, 3, and 6, respectively. The problem has an solution $x^*=(0,0,0,0,0,0)^T$. It can be verified that the Lagrangian function for this example is

$$L(x,\mu,\lambda)=\frac{1}{2}Dx+A^T\mu+B^T\lambda.$$

Note that $\nabla_x L(x,\mu,\lambda)$ is positive definite. We know from Theorems 3.1 and 4.2 that both proposed neural networks are globally convergent to the KKT triple of the SOCCVI problem. Figures 3 and 4 depict the trajectories of Example 5.2 obtained using the proposed neural networks. The simulation results show that both neural networks are effective in the SOCCVI problem and the neural network with projection function converges to $x^*$ quicker than that with smoothed FB function.

*Example 5.3* Consider the SOCCVI problem (1)–(2) where

$$F(x)=\begin{pmatrix} x_3\exp(x_1x_3)+6(x_1+x_2)\\ 6(x_1+x_2)+\frac{2(2x_2-x_3)}{\sqrt{1+(2x_2-x_3)^2}}\\ x_1\exp(x_1x_3)-\frac{2x_2-x_3}{\sqrt{1+(2x_2-x_3)^2}}\\ x_4\\ x_5\end{pmatrix}$$

and

$$C=\{x\in\mathbb{R}^5 : h(x)=0,\ -g(x)\in\mathcal{K}^3\times\mathcal{K}^2\},$$

with

$$h(x)=-62x_1^3+58x_2+167x_3^3-29x_3-x_4-3x_5+11,$$

$$g(x)=\begin{pmatrix}-3x_1^3-2x_2+x_3-5x_3^3\\ 5x_1^3-4x_2+2x_3-10x_3^3\\ -x_3\\ -x_4\\ -3x_5\end{pmatrix}.$$

This problem has an approximate solution $x^*=(0.6287, 0.0039, -0.2717, 0.1761, 0.0587)^T$.

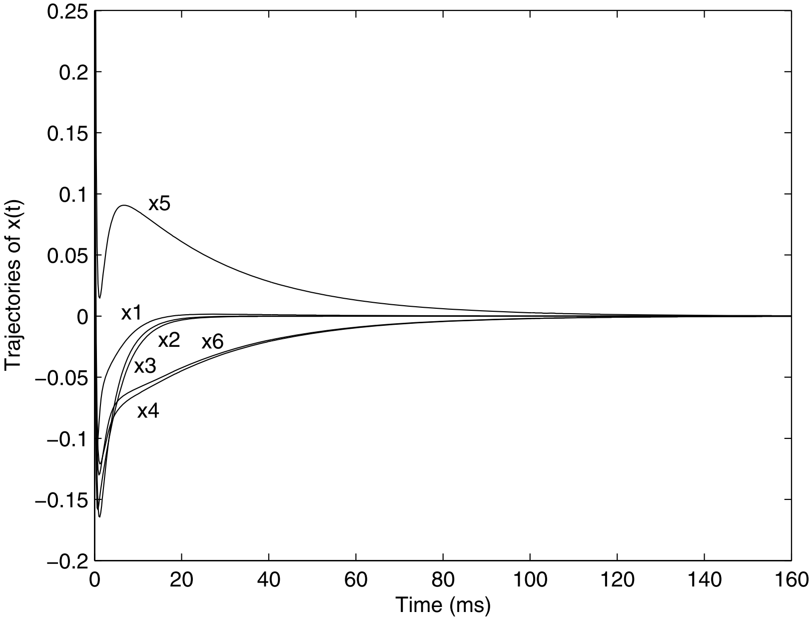


**Fig. 3** Transient behavior of neural network with smoothed FB function in Example 5.2

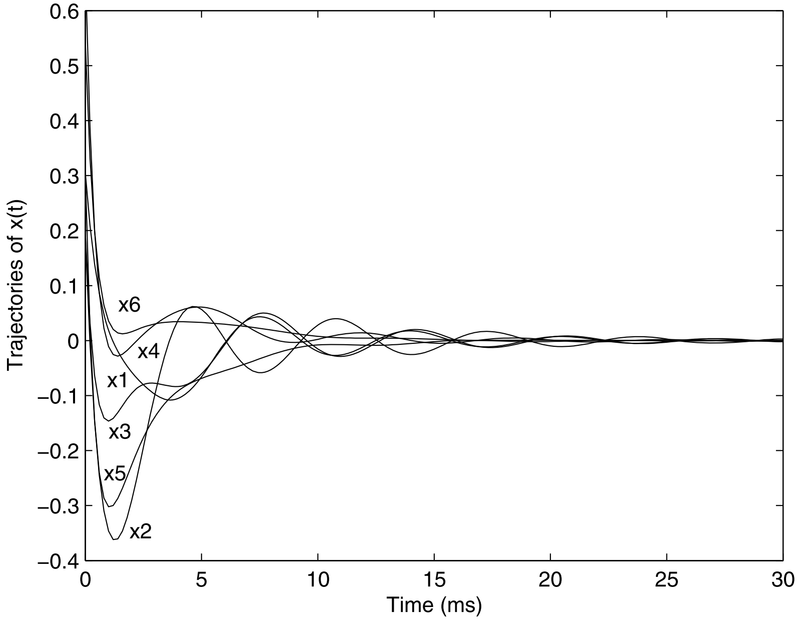


**Fig. 4** Transient behavior of neural network with projection function in Example 5.2

*Example 5.4* Consider the SOCCVI problem (1)–(2) where

$$F(x) = \begin{pmatrix} 4x_1 - \sin x_1 \cos x_2 + 1 \\ -\cos x_1 \sin x_2 + 6x_2 + \frac{9}{5}x_3 + 2 \\ \frac{9}{5}x_2 + 8x_3 + 3 \\ 2x_4 + 1 \end{pmatrix}$$

and

$$C = \left\{ x \in \mathbb{R}^4 : h(x) = \begin{pmatrix} x_1^2 - \frac{1}{10}x_2x_3 + x_3 \\ x_3^2 + x_4 \end{pmatrix} = 0,\ -g(x) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathcal{K}^2 \right\}.$$

This problem has an approximate solution $x^* = (0.2391, -0.2391, -0.0558, -0.0031)^T$.

The neural network (10) based on smoothed FB function can solve Examples 5.3–5.4 successfully, see Figs. 5, 6, whereas the neural network (18) based on projection function fails to solve them. This is because that $\nabla_x L(x, \mu, \lambda)$ is not always positive definite in Examples 5.3 and 5.4. Hence, the neural network with projection function is not effective in these two problems. To the contrast, though there is no guarantee that the Assumption 3.1(b) holds, the neural network with smoothed FB function is asymptotically stable from Theorem 3.2. Figures 5 and 6 depict the trajectories obtained using the neural network with the smoothed FB function for Examples 5.3 and 5.4, respectively. The simulation results show that each trajectory converges to the desired isolated equilibrium point which is exactly the approximate solution of Examples 5.3 and 5.4, respectively.

*Example 5.5* Consider the nonlinear convex SOCP [23] given by

$$\begin{aligned} \min \quad & \exp(x_1 - x_3) + 3(2x_1 - x_2)^4 + \sqrt{1 + (3x_2 + 5x_3)^2} \\ \text{s.t.} \quad & -g(x) = \begin{pmatrix} 4x_1 + 6x_2 + 3x_3 - 1 \\ -x_1 + 7x_2 - 5x_3 + 2 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathcal{K}^2 \times \mathcal{K}^3. \end{aligned}$$

The approximate solution of this problem is $x^* = (0.2324, -0.07309, 0.2206)^T$. As mentioned in Sect. 1, the CSOCP in Example 5.5 can be transformed into an equivalent SOCCVI problem. There are neural network models proposed for CSOCP in [27]. We here try a different approach for it. In other words, we use the proposed neural networks with the smoothed FB and projection functions, respectively, to solve the problem whose trajectories are depicted in Figs. 7 and 8. From the simulation results, we see that the neural network with smoothed FB function converges very slowly and it is not clear whether it converges to the solution in finite time or not. In view of these, to solve CSOCP, it seems better to apply the models introduced in [27] directly instead of transforming CSOCP into SOCCVI problem.

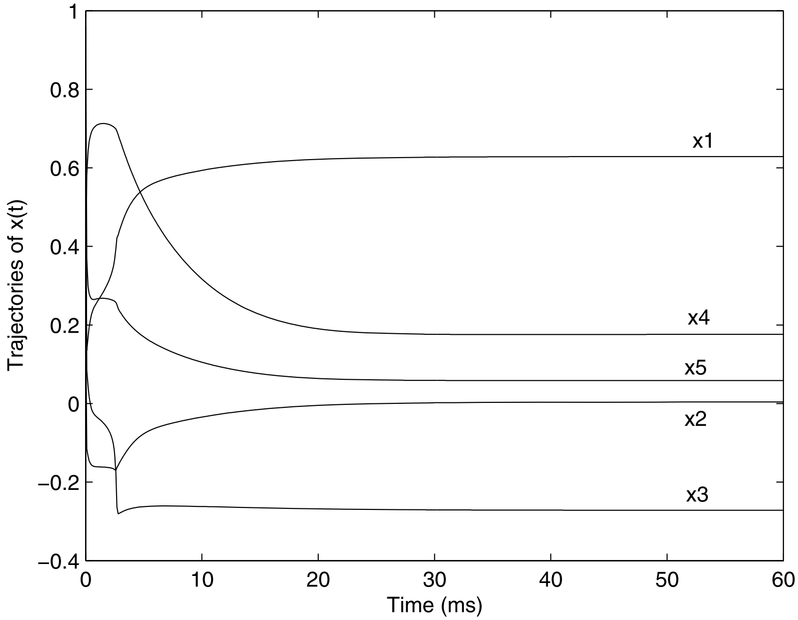


**Fig. 5** Transient behavior of neural network with smoothed FB function in Example 5.3

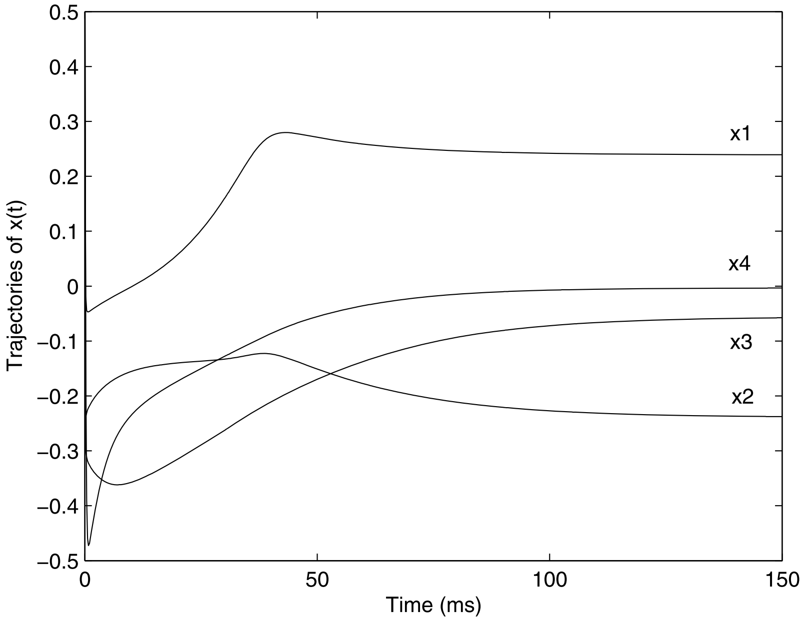


**Fig. 6** Transient behavior of neural network with smoothed FB function in Example 5.4

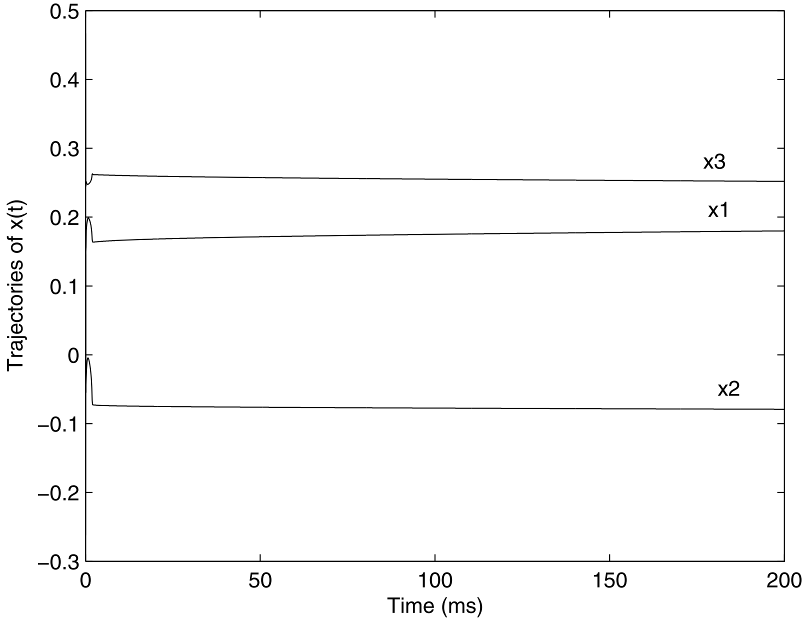


**Fig. 7** Transient behavior of neural network with smoothed FB function in Example 5.5

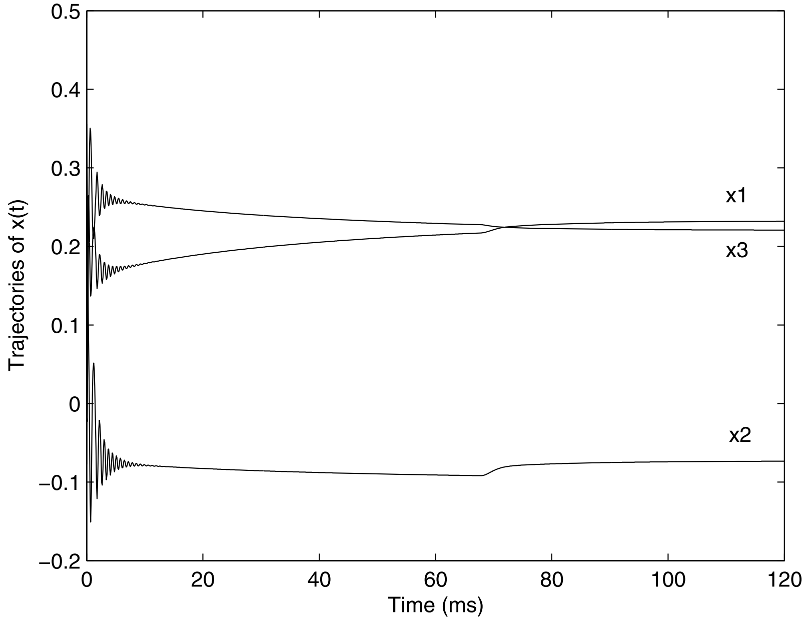


**Fig. 8** Transient behavior of neural network with projection function in Example 5.5

The simulation results of Examples 5.1, 5.2 and 5.5 tell us that the neural network with projection function converges to $x^*$ quicker than that with the smoothed FB function. In general, the neural network with projection function has lower model complexity than that with the smoothed FB function. Hence, the neural network with projection function is preferable to the neural network with the smoothed FB function when both can globally converge to the solution of SOCCVI problem. On the other hand, from Examples 5.3 and 5.4, the neural network with smoothed FB function seems better for use when the positive semidefinite condition of $\nabla_x L(x, \mu, \lambda)$ is not satisfied.

## 6 Conclusions

In this paper, We use the proposed neural networks with smoothed Fischer-Burmeister and projection functions to solve the SOCCVI problems. The first neural network uses the Fischer-Burmeister (FB) function to achieve an unconstrained minimization which is a merit function of the Karush-Kuhn-Tucker equation. We show that the merit function is a Lyapunov function and this neural network is asymptotically stable. Under Assumption 3.1, we prove that if $(\varepsilon, x, \mu, \lambda) \in \mathbb{R}^{1+n+l+m}$ is the accumulation point of the trajectory, then $(x, \mu, \lambda) \in \mathbb{R}^{n+l+m}$ is a KKT triple of SOCCVI problem and the neural network is exponentially stable. The second neural network is introduced for solving a projection formulation whose solutions coincide with the KKT triples of SOCCVI problem under the positive semidefinite condition of $\nabla_x L(x, \mu, \lambda)$. Its Lyapunov stability and global convergence are proved. Simulations show that both neural networks have merits of their own.

**Acknowledgement** The authors thank the referees for their carefully reading this paper and helpful suggestions.

## References

1. Aghassi, M., Bertsimas, D., Perakis, G.: Solving asymmetric variational inequalities via convex optimization. Oper. Res. Lett. **34**, 481–490 (2006)
2. Bertsekas, D.P.: Nonlinear Programming. Athena Scientific, Nashua (1995)
3. Chen, J.-S.: The semismooth-related properties of a merit function and a decent method for the nonlinear complementarity problem. J. Glob. Optim. **36**, 565–580 (2006)
4. Chen, J.-S., Gao, H.-T., Pan, S.-H.: An R-linearly convergent derivative-free algorithm for nonlinear complementarity problems based on the generalized Fischer-Burmeister merit function. J. Comput. Appl. Math. **232**, 455–471 (2009)
5. Chen, J.-S., Ko, C.-H., Pan, S.-H.: A neural network based on the generalized Fischer-Burmeister function for nonlinear complementarity problems. Inf. Sci. **180**, 697–711 (2010)
6. Chen, J.-S., Pan, S.-H.: A family of NCP functions and a descent method for the nonlinear complementarity problem. Comput. Optim. Appl. **40**, 389–404 (2008)
7. Chen, J.-S., Tseng, P.: An unconstrained smooth minimization reformulation of the second-order cone complementarity problem. Math. Program. **104**, 293–327 (2005)
8. Chen, X., Qi, L., Sun, D.: Global and superlinear convergence of the smoothing Newton method and its application to general box constrained variational inequalities. Math. Comput. **67**(222), 519–540 (1998)
9. Dang, C., Leung, Y., Gao, X., Chen, K.: Neural networks for nonlinear and mixed complementarity problems and their applications. Neural Netw. **17**, 271–283 (2004)

10. Effati, S., Ghomashi, A., Nazemi, A.R.: Applocation of projection neural network in solving convex programming problems. Appl. Math. Comput. **188**, 1103–1114 (2007)
11. Effati, S., Nazemi, A.R.: Neural network and its application for solving linear and quadratic programming problems. Appl. Math. Comput. **172**, 305–331 (2006)
12. Fukushima, M.: Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. Math. Program. **53**(1), 99–110 (1992)
13. Facchinei, F., Pang, J.: Finite-dimensional Variational Inequalities and Complementarity Problems. Springer, New York (2003)
14. Fukushima, M., Luo, Z.-Q., Tseng, P.: Smoothing functions for second-order-cone complimentarity problems. SIAM J. Optim. **12**, 436–460 (2002)
15. Golden, R.: Mathematical Methods for Neural Network Analysis and Design. MIT Press, Cambridge (1996)
16. Hammond, J.: Solving asymmetric variational inequality problems and systems of equations with generalized nonlinear programming algorithms. Ph.D. Dissertation, Department of Mathematics, MIT, Cambridge (1984)
17. Han, Q., Liao, L.-Z., Qi, H., Qi, L.: Stability analysis of gradient-based neural networks for optimization problems. J. Glob. Optim. **19**, 363–381 (2001)
18. Harker, P., Pang, J.-S.: Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. Math. Program. **48**, 161–220 (1990)
19. Hartman, P., Stampacchia, G.: On some nonlinear elliptic differential functional equations. Acta Math. **115**, 271–310 (1966)
20. Hopfield, J.J., Tank, D.W.: Neural computation of decision in optimization problems. Biol. Cybern. **52**, 141–152 (1985)
21. Hu, X., Wang, J.: Solving pseudomonotone variational inequalities and pseudoconvex optimization problems using the projection neural network. IEEE Trans. Neural Netw. **17**, 1487–1499 (2006)
22. Hu, X., Wang, J.: A recurrent neural network for solving a class of general variational inequalities. IEEE Trans. Syst. Man Cybern. B **37**, 528–539 (2007)
23. Kanzow, C., Ferenczi, I., Fukushima, M.: On the local convergence of semismooth Newton methods for linear and nonlinear second-order cone programs without strict complementarity. SIAM J. Optim. **20**, 297–320 (2009)
24. Kennedy, M.P., Chua, L.O.: Neural network for nonlinear programming. IEEE Trans. Circuits Syst. **35**, 554–562 (1988)
25. Kanno, Y., Martins, J.A.C., Pinto da Costa, A.: Three-dimensional quasi-static frictional contact by using second-order cone linear complementarity problem. Int. J. Numer. Methods Eng. **65**, 62–83 (2006)
26. Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications. Academic Press, San Diego (1980)
27. Ko, C.-H., Chen, J.-S., Yang, C.-Y.: A recurrent neural network for solving nonlinear second-order cone programs. Submitted manuscript (2009)
28. Liao, L.-Z., Qi, H., Qi, L.: Solving nonlinear complementarity problems with neural networks: a reformulation method approach. J. Comput. Appl. Math. **131**, 342–359 (2001)
29. Lions, J., Stampacchia, G.: Variational inequalities. Commun. Pure Appl. Math. **20**, 493–519 (1967)
30. Mancino, O., Stampacchia, G.: Convex programming and variational inequalities. J. Optim. Theory Appl. **9**, 3–23 (1972)
31. Miller, R.K., Michel, A.N.: Ordinary Differential Equations. Academic Press, San Diego (1982)
32. Ortega, J.M., Rheinboldt, W.C.: Iterative solution of nonlinear equations in several variables. Academic Press, San Diego (1970)
33. Pan, S.-H., Chen, J.-S.: A semismooth Newton method for the SOCCP based on a one-parametric class of SOC complementarity functions. Comput. Optim. Appl. **45**, 59–88 (2010)
34. Patriksson, M.: Nonlinear Programming and Variational Inequality Problems: A Unified Approach. Applied Optimization, vol. 23. Dordrecht, Kluwer (1998)
35. Stampacchia, G.: Formes bilineares coercives sur les ensembles convexes. C. R. Acad. Sci. Paris **258**, 4413–4416 (1964)
36. Stampacchia, G.: Variational inequalities. Theory and applications of monotone operators. In: Proceedings of the NATO Advanced Study Institute, pp. 101–192, Venice (1968)
37. Sun, J.-H., Zhang, L.-W.: A globally convergent method based on Fischer-Burmeister operators for solving second-order-cone constrained variational inequality problems. Comput. Math. Appl. **58**, 1936–1946 (2009)

38. Tank, D.W., Hopfield, J.J.: Simple neural optimization network: an A/D converter, signal decision circuit, and a linear programming circuit. IEEE Trans. Circuits Syst. **33**, 533–541 (1986)
39. Wright, S.J.: An infeasible-interior-point algorithm for linear complementarity problems. Math. Program. **67**, 29–51 (1994)
40. Xia, Y., Leung, H., Wang, J.: A projection neural network and its application to constrained optimization problems. IEEE Trans. Circuits Syst. I **49**, 447–458 (2002)
41. Xia, Y., Leung, H., Wang, J.: A general projection neural network for solving monotone variational inequalities and related optimization problems. IEEE Trans. Neural Netw. **15**, 318–328 (2004)
42. Xia, Y., Wang, J.: A recurrent neural network for solving nonlinear convex programs subject to linear constraints. IEEE Trans. Neural Netw. **16**, 379–386 (2005)
43. Yashtini, M., Malek, A.: Solving complementarity and variational inequalities problems using neural networks. Appl. Math. Comput. **190**, 216–230 (2007)
44. Zak, S.H., Upatising, V., Hui, S.: Solving linear programming problems with neural networks: a comparative study. IEEE Trans. Neural Netw. **6**, 94–104 (1995)