



# An efficient neural network for solving convex optimization problems with a nonlinear complementarity problem function

M. Ranjbar<sup>1</sup> · S. Effati<sup>1</sup> · S. M. Miri<sup>1</sup>

Published online: 12 July 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

In this paper, we present a one-layer recurrent neural network (NN) for solving convex optimization problems by using the Mangasarian and Solodov (MS) implicit Lagrangian function. In this paper by using Krush–Kuhn–Tucker conditions and MS function the NN model was derived from an unconstrained minimization problem. The proposed NN model is one layer and compared to the available NNs for solving convex optimization problems, which has a better performance in convergence time. The proposed NN model is stable in the sense of Lyapunov and globally convergent to optimal solution of the original problem. Finally, simulation results on several numerical examples are presented and the validity of the proposed NN model is demonstrated.

**Keywords** One-layer neural networks · Convex programming · Nonlinear complementarity problem

## 1 Introduction

Mathematical programming problems have been widely applied in practically every area of production, physical sciences, many engineering problems and government planning. The nonlinear programming (NLP) problem plays an important part among them. Convex programming is a widespread class of NLP problems where the objective function and constraints are convex functions.

NNs are computing systems composed of a number of highly interconnected simple information processing units, and thus can usually solve optimization problems faster than most popular optimization algorithms. Also, the numerical ordinary differential equation techniques can be applied directly to the continuous-time NN for solving constrained optimization problems effectively. NN models are usually more competent than numerical optimization methods

because of their inherent parallel nature. First Tank and Hopfield (1986) proposed a NN model for linear programming (LP) problems. Then, Kennedy and Chua (1988) proposed a NN model with a finite penalty parameter for solving NLP problems. Rodriguez et al. (1990) proposed a switched capacitor NN for solving a class of constrained nonlinear convex optimization problems. Zhang and Constantinides (1992) proposed a two-layer NN model to solve some strictly convex programming problems. Bouzerdoum and Pattison (1993) presented a recurrent NN for solving convex quadratic optimization problems with bounded constraints. Zhang (1996), Zhang et al. (2002) proposed an adaptive NN model for NLP problems. Wang (1994), Xia (1996), Xia and Wang (2000, 2004, 2005) presented several NN models for solving LP and NLP problems, monotone variational inequality problems and monotone projection equations. Effati and Baymani (2005), Effati et al. (2011, 2015), Effati and Nazemi (2006), Effati and Ranjbar (2011), Ranjbar et al. (2017) proposed some NN models for solving LP, NLP and binary programming problems. Nazemi (2012, 2013, 2014) proposed some dynamic system models for solving convex NLP problems. Also, recently Huang and Cui (2016) proposed a NN model for solving convex quadratic programming (QP) problems.

The nonlinear complementarity problems (NCPs) attracted a lot of attention because of its wide applications in operations research, economics and engineering (Chen et al. 2010; Ferris et al. 2001). Liao et al. (2001) presented a NN

Communicated by V. Loia.

✉ M. Ranjbar  
m.ranjbartaghiabad@mail.um.ac.ir; m.ranjbart@gmail.com

S. Effati  
s-effati@um.ac.ir

S. M. Miri  
mohsenmiri@mail.um.ac.ir

<sup>1</sup> Department of Applied Mathematics, Ferdowsi University of Mashhad, Mashhad, Iran

approach for solving NCPs, which NN model has derived from an unconstrained minimization reformulation of the complementarity problem. A popular NCP function is the Fischer–Burmeister function (see Fischer 1992, 1997); also Chen and Pan (2008) proposed a family of NCP functions that subsumes the Fischer–Burmeister function as a special case. Chen et al. (2010) developed a NN for solving the NCPs based on the generalized Fischer–Burmeister.

In this paper by using Krush–Kuhn–Tucker (KKT) condition and by the MS implicit Lagrangian function (see Mangasarian and Soldov 1993) as NCP function, we presented a one-layer NN model for solving the convex optimization problems. The rest of the paper is organized as follows; in Sect. 2, some preliminary results are provided. In Sect. 3 are proposed an equivalent formulation for convex optimization problems and a NN model for it. Convergence and stability results are discussed in Sect. 4. Simulation results on several numerical examples of the new model are reported in Sect. 5. Finally, some concluding remarks are drawn in Sect. 6.

## 2 Preliminaries

In this section, we recall some necessary mathematical background concepts and materials which will play an important role for the desired NN and to study its stability. First, we recall some basic classes of functions and matrices and then introduce some properties of special NCP functions.

**Definition 2.1** A matrix  $A \in \mathbb{R}^{n \times n}$  is a

- (i)  $P_0$ -matrix if each of its principal minors is nonnegative.
- (ii)  $P$ -matrix if each of its principal minors is positive.

Obviously, a positive-definite matrix is a  $P$ -matrix and a semi-positive-definite matrix is a  $P_0$ -matrix. For more properties about  $P$ -matrix and  $P_0$ -matrix, please refer to Chen and Pan (2008).

**Definition 2.2** The function  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is said to be a

- (i) monotone if  $(x - y)(F(x) - F(y)) \geq 0$  for all  $x, y \in \mathbb{R}^n$ ;
- (ii)  $P_0$ -function if  $\max_{1 \leq i \leq n, x_i \neq y_i} (x_i - y_i)(F_i(x) - F_i(y)) \geq 0$  for all  $x, y \in \mathbb{R}^n$  with  $x \neq y$ ;
- (iii)  $P$ -function if  $\max_{1 \leq i \leq n} (x_i - y_i)(F_i(x) - F_i(y)) > 0$  for all  $x, y \in \mathbb{R}^n$  with  $x \neq y$ .

**Note 2.1** It is known that  $F(x)$  is a  $P_0$ -function if and only if  $\frac{dF}{dx}$  is a  $P_0$ -matrix for all  $x \in \mathbb{R}^n$ , and if  $\frac{dF}{dx}$  is a  $P$ -matrix for all  $x \in \mathbb{R}^n$ , then  $F$  must be a  $P$ -function.

The NCP is to find a point  $x \in \mathbb{R}^n$  such that

$$x \geq 0, \quad F(x) \geq 0, \quad \langle x, F(x) \rangle = 0, \tag{1}$$

where  $\langle \cdot, \cdot \rangle$  is the Euclidean inner product and  $F = (F_1, F_2, \dots, F_n)^T$  maps from  $\mathbb{R}^n$  to  $\mathbb{R}^n$ . There are many methods for solving the NCP, one of the most popular and powerful approaches that has been studied intensively is to reformulate the NCP as a system of nonlinear equations or as an unconstrained minimization problem. A function that can constitute an equivalent unconstrained minimization problem for the NCP is called a merit function. Many studies on NCP functions and applications have been achieved during the past three decades (see Chen 2007; Cottle et al. 1992; Ferris and Kanzow 2002; He et al. 2015; Hu et al. 2009; Kanzow et al. 1997; Mangasarian and Soldov 1993; Miri and Effati 2015). The class of NCP functions defined below is used to construct a merit function.

**Definition 2.3** A function  $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$  is called an NCP function if it satisfies,

$$\phi(a, b) = 0 \iff a \geq 0, \quad b \geq 0, \quad ab = 0. \tag{2}$$

For example, the following functions are NCP functions:

- (a)  $\phi(a, b) = \min\{a, b\}$
- (b)  $\phi(a, b) = (\frac{1}{2})((ab)^2 + \min^2\{0, a\} + \min^2\{0, b\})$
- (c)  $\phi(a, b) = \sqrt{a^2 + b^2} - a - b$
- (d)  $\phi(a, b) = ab + \frac{1}{2\alpha}(((a - \alpha b)_+)^2 - a^2 + ((b - \alpha a)_+)^2 - b^2)$ ,  $\alpha > 1$ ,

where  $(x)_+ = \max\{0, x\}$  and  $\alpha > 1$ . Some other NCP functions are listed in Chen (2007); Ferris and Kanzow (2002); Hu et al. (2009); Kanzow et al. (1997). Among introduced NCP functions in above, we use function (d), that is, the well-known MS NCP function, as follows:

$$M(x, \alpha) = xF(x) + \frac{1}{2\alpha} (\|(x - \alpha F(x))_+\|^2 - \|x\|^2 + \|(F(x) - \alpha x)_+\|^2 - \|F(x)\|^2).$$

This NCP function has some positive features as follows:

- $M(x, \alpha)$  is nonnegative on  $\mathbb{R}^n \times (1, \infty)$ . This is not true for NCP functions (a) and (c).
- $M(x, \alpha)$  is equal to zero if and only if  $x$  is a solution of the NCP, without regard to whether  $F$  is monotone or not.
- $M(x, \alpha)$  is continuously differentiable at all points. This is not true for NCP function (a).
- If  $F$  is differentiable on  $\mathbb{R}^n$ ,  $M(x, \alpha)$  satisfy in  $M(x, \alpha) = 0 \iff \nabla M(x, \alpha) = 0$  for  $\alpha > 1$ . This is not true for NCP function (b).
- It is a merit function. This is not true for NCP functions (a) and (c).

For to see other features and checking the validity of the above features, refer to Mangasarian and Soldov (1993). For  $F = (F_1, F_2, \dots, F_n)^T$ , NCP( $F$ ) can be equivalently reformulated as finding a solution of the following equation:

$$\begin{bmatrix} M_1(x, \alpha) \\ \vdots \\ M_n(x, \alpha) \end{bmatrix} = 0, \tag{3}$$

where  $M_i(x, \alpha) = x_i F_i(x) + \frac{1}{2\alpha}(((x_i - \alpha F_i(x))_+)^2 - x_i^2 + ((F_i(x) - \alpha x_i)_+)^2 - F_i^2(x))$  for  $i = 1, 2, \dots, n$ . For convenience, we define  $M_\alpha(x)$  as follows.

**Definition 2.4** We define  $M_\alpha(x) : \mathbb{R}^n \rightarrow \mathbb{R}_+$  by

$$M_\alpha(x) = \sum_{i=1}^n M_i(x, \alpha). \tag{4}$$

In follows, we will present some favorable properties of  $M_\alpha(x)$ , which their proofs are in Mangasarian and Soldov (1993).

**Theorem 2.1** For  $\alpha \in (1, \infty)$ ,  $M_\alpha(x) \geq 0$  for all  $x \in \mathbb{R}^n$  and  $M_\alpha(x) = 0$  if and only if  $x$  solves the NCP.

Theorem 2.1 establishes a one-to-one correspondence between solutions of the NCP and global unconstrained minima of the merit function  $M_\alpha(x)$ . As a consequence were obtained the following immediate results.

**Corollary 2.1** If  $F$  is differentiable at a solution  $\bar{x}$  of NCP, then  $\nabla M_\alpha(\bar{x}) = 0$  for  $\alpha \in (1, \infty)$ .

Theorem 2.2 shows that under certain assumptions, each stationary point of the unconstrained objective function  $M_\alpha(x)$  is already a global minimum and therefore a solution of problem (1).

**Theorem 2.2** Let  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be continuously differentiable having a positive-definite Jacobian  $\frac{dF}{dx}$  for all  $x \in \mathbb{R}^n$ . Assume that the complementarity problem is solvable, then  $x^*$  is a stationary point of  $M_\alpha(x)$  if and only if  $x^*$  solve NCP.

We also recall some materials about first-order differential equation and Lyapunov function. These materials can be found in ordinary differential equation and nonlinear control textbooks (see Miller and Michel 1982; Slotin and Li 1991).

**Definition 2.5** Let  $K \subseteq \mathbb{R}^n$  be an open neighborhood of  $x^*$ . A continuously differentiable function  $E : \mathbb{R}^n \rightarrow \mathbb{R}$  is said to be a Lyapunov function at the state  $x^*$  over the set  $K$  for  $\dot{x} = f(x(t))$  if

- (i)  $E(x^*) = 0$  and for all  $x \neq x^* \in K$ ,  $E(x) > 0$ .

- (ii)  $\frac{d(E(x(t)))}{dt} = (\nabla_x E(x(t)))^T \dot{x} = (\nabla_x E(x(t)))^T f(x(t)) \leq 0$ .

**Lemma 2.1** (i) An isolated equilibrium point  $x^*$  is Lyapunov stable if there exists a Lyapunov function over some neighborhood  $K$  of  $x^*$ .

(ii) An isolated equilibrium point  $x^*$  is asymptotically stable if there is a Lyapunov function over some neighborhood  $K$  of  $x^*$  such that  $\frac{dE}{dt} < 0$  for all  $x \neq x^* \in K$ .

### 3 Problem formulation and neural network design

Consider the following convex optimization problem:

$$\begin{aligned} \text{Min} \quad & f(x) \\ \text{s.t.} \quad & g(x) \leq 0 \\ & x \geq 0, \end{aligned} \tag{5}$$

where  $x \in \mathbb{R}^n$ ,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $g(x) = [g_1(x), \dots, g_m(x)]^T$  is an  $m$ -dimensional vector-valued continuous function of  $n$  variables and  $f, g_1, \dots, g_m$  are convex and twice differentiable.

It is well known (see Bazaraa and Shetty 1979) that by the KKT conditions,  $x \in \mathbb{R}^n$  is an optimal solution to (5) if and only if there exists  $u \in \mathbb{R}^m$  such that

$$\begin{aligned} \nabla f(x) + \sum_{i=1}^m u_i \nabla g_i(x) &\geq 0 \\ [\nabla f(x) + \sum_{i=1}^m u_i \nabla g_i(x)]^T x &= 0 \\ u_i g_i(x) &= 0, \quad i = 1, 2, \dots, m \\ u_i &\geq 0 \quad i = 1, 2, \dots, m \\ g_i(x) &\leq 0 \quad i = 1, 2, \dots, m \\ x &\geq 0. \end{aligned} \tag{6}$$

It is clear that the KKT optimality conditions of this problem lead to a complementarity problem as follows

$$z \geq 0, \quad F(z) \geq 0, \quad \langle z, F(z) \rangle = 0,$$

where  $z = (x, u)$  and  $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m}$  is defined by

$$F(z) = \begin{bmatrix} \nabla_{x_1} f(x) + \sum_{i=1}^m u_i \nabla_{x_1} g_i(x) \\ \vdots \\ \nabla_{x_n} f(x) + \sum_{i=1}^m u_i \nabla_{x_n} g_i(x) \\ -g_1(x) \\ \vdots \\ -g_m(x) \end{bmatrix}, \tag{7}$$

by considering:

$$F_k(z) = \nabla_{x_k} f(x) + \sum_{i=1}^m u_i \nabla_{x_k} g_i(x), \quad k = 1, 2, \dots, n,$$

$$F_{n+j}(z) = -g_j(x), \quad j = 1, 2, \dots, m.$$

Solving problem (5) is equivalent to find a solution for the following equation:

$$\begin{bmatrix} M_1(z, \alpha) \\ \vdots \\ M_{n+m}(z, \alpha) \end{bmatrix} = 0, \tag{8}$$

where  $M_i(z, \alpha) = z_i F_i(z) + \frac{1}{2\alpha} ((z_i - \alpha F_i(z))_+)^2 - z_i^2 + ((F_i(z) - \alpha z_i)_+)^2 - F_i^2(z)$  for  $i = 1, 2, \dots, n + m$ .

By Definition 2.4 and Theorem 2.1, we have  $M_\alpha(z) \geq 0$  for all  $z \in \mathbb{R}^{n+m}$  and  $z$  solves (8) if and only if  $M_\alpha(z) = 0$ . Hence, solving (8) is equivalent to finding the global minimizer of  $M_\alpha(z)$  if (8) has solution.

We utilize  $M_\alpha(z)$  as the traditional energy function. As mentioned above, problem (5) is equivalent to the unconstrained smooth minimization problem as follows

$$\min_{z \in \mathbb{R}^{n+m}} M_\alpha(z), \tag{9}$$

where  $M_\alpha(z) = \sum_{i=1}^{n+m} M_i(z, \alpha)$ . Hence, it is natural to adopt the following steepest descent-based NN for problem (5)

$$\frac{dz(t)}{dt} = -\eta \nabla M_\alpha(z), \tag{10}$$

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

Compared with the existent NNs for solving the such non-linear optimization problem, the proposed NN is one layer and its architecture is shown in Fig. 1. Also, it has a simple form and a better performance in convergence time that is shown in Sect. 5. Moreover, the proposed NN is stable in the sense of Lyapunov and has globally convergent to an exact optimal solution of the original problem.

### 4 Stability analysis

In this section, we will study the stability of the equilibrium point and the convergence of the optimal solution of NN (10).

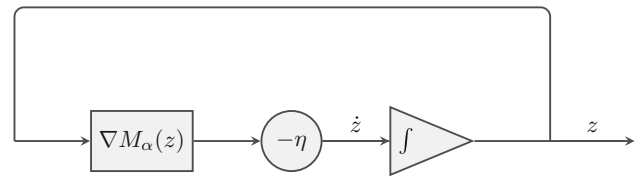


Fig. 1 Architecture of the proposed NN in (10)

We first state the relationships between an equilibrium point of (10) and a solution to the NCP.

**Lemma 4.1** *Let  $S$  be a nonempty open convex set in  $\mathbb{R}^n$ , and let  $f : S \rightarrow \mathbb{R}$  be twice differentiable on  $S$ . Then  $f$  is convex if and only if the Hessian matrix is semi-positive definite at each point in  $S$ .*

**Lemma 4.2** *If the objective function and all constraint functions are convex, then function  $F$  that has been defined in (7) is a  $P_0$ -function on  $\mathbb{R}^{n+m}$ .*

**Proof** For this purpose, it is sufficient that  $\frac{\partial F}{\partial z}$  is semi-positive-definite matrix on  $\mathbb{R}^{n+m}$ . We have

$$\frac{\partial F}{\partial z} = \begin{bmatrix} \nabla^2 f(x) + u \nabla^2 g(x) & \nabla g(x) \\ -\nabla g(x) & 0_{m \times m} \end{bmatrix}_{(n+m) \times (n+m)},$$

now, let  $z = (x, u) \neq 0$  is arbitrary vector in  $\mathbb{R}^{n+m}$ , since  $f(x)$  and  $g_i(x)$  are convex and  $u_i \geq 0$  for  $i = 1, 2, \dots, m$  by Lemma 2.1, we have

$$z \frac{\partial F}{\partial z} z^T = x(\nabla^2 f(x) + u \nabla^2 g(x))x^T \geq 0;$$

hence,  $F$  is an  $P_0$ -function. □

In the next theorem, we state the existence of the solution trajectory of (10).

**Theorem 4.1** *For any initial state  $z_0 = z(t_0)$ , there is exactly one unique solution  $z(t)$  with  $t \in [t_0, \tau(z_0))$  for NN (10).*

**Proof** Since  $\nabla M_\alpha(z)$  is continuous, there is a local solution  $z(t)$  for (10) with  $t \in [t_0, \tau)$  for some  $\tau \geq t_0$  and since  $\nabla M_\alpha(z)$  is locally Lipschitz continuous, the proof is completed. □

**Theorem 4.2** *Let  $z^*$  be an isolated equilibrium point of NN (10),  $z^*$  is globally asymptotically stable for (10).*

**Proof** Since  $z^*$  is a solution to the NCP,  $M_\alpha(z^*) = 0$ . In addition, since  $z^*$  is an isolated equilibrium point of (10), there is a neighborhood  $K \subseteq \mathbb{R}^{n+m}$  of  $z^*$  such that

$$\nabla M_\alpha(z^*) = 0, \quad \nabla M_\alpha(z) \neq 0 \quad \forall z \in K \setminus \{z^*\}.$$

Next, we define a Lyapunov function as

$$E(z(t)) = M_\alpha(z),$$

we show that  $E(z(t))$  is a Lyapunov function over the set  $K$  for NN (10). By Theorem 2.1, we have  $E(z) \geq 0$  over  $\mathbb{R}^{n+m}$ , since  $z^*$  is a solution to NCP( $F$ ), obviously  $E(z^*) = 0$ . If there is an  $z \in K \setminus \{z^*\}$  that satisfying  $E(z) = 0$ , then from Corollary 2.1  $\nabla M_\alpha(z) = 0$ ; hence,  $z$  is an equilibrium point of (10), and it is contradicting with the being isolated of  $z^*$  in  $K$ . On the other hand, since the  $F$  is  $P_0$  function, Theorem 3.2 in Facchinei (1998) shows that there is no solution other than  $z^*$  for NCP( $F$ ); therefore, for all  $z \neq z^*$  we have  $E(z) > 0$ .

Now it is sufficient that for all  $z \in K \setminus \{z^*\}$ , we have  $\frac{dE(z)}{dt} < 0$ . For this purpose, by taking the derivative of  $E(z)$  with respect to time  $t$ , since  $z^*$  is isolated equilibrium point for all  $z(t) \in K \setminus \{z^*\}$ , we have:

$$\begin{aligned} \frac{dE(z)}{dt} &= \frac{dE(z)}{dz} \frac{dz}{dt} = (\nabla M_\alpha(z))^T (-\eta \nabla M_\alpha(z)) \\ &= -\eta \|\nabla M_\alpha(z)\|^2 < 0, \end{aligned}$$

and thus, by Lemma 2.1(ii), it implies that  $z^*$  is globally asymptotically stable.  $\square$

### 5 Simulation results

In order to demonstrate the effectiveness and performance of the proposed NN model, we discuss several illustrative examples. Note, to solve the proposed NN in (10), we use the solver ode15s of MATLAB. The tolerances were chosen AbsTol = 1e-6 and RelTol = 1e-4.

**Example 5.1** Consider the following NLP problem:

$$\begin{aligned} \text{Min} \quad & f(x) = \frac{1}{4}x_1^4 + \frac{1}{2}x_1^2 + \frac{1}{4}x_2^4 + \frac{1}{2}x_2^2 - \frac{9}{10}x_1x_2 \\ \text{s.t.} \quad & x_1 + x_2 \leq 2 \\ & -x_1 + x_2 \leq 2 \\ & x_1 - 3x_2 \leq -2 \\ & x_i \geq 0, \quad i = 1, 2. \end{aligned}$$

In this problem,  $f(x)$  is strictly convex and the feasible region is a convex set, and the optimal solution of the NLP problem is  $x^* = [0.3461, 0.7820]^T$ . We apply the proposed NN in (10) to solve the above problem. Simulation results show the trajectory of (10) with 16 initial points is always convergent to

$$z^* = [0.3461, 0.7820, 0.0000, 0.0000, 0.3163]^T.$$

Figure 2 displays the transient behavior of  $z(t)$  with 16 initial points. Moreover, Fig. 3 shows the phase diagram of state

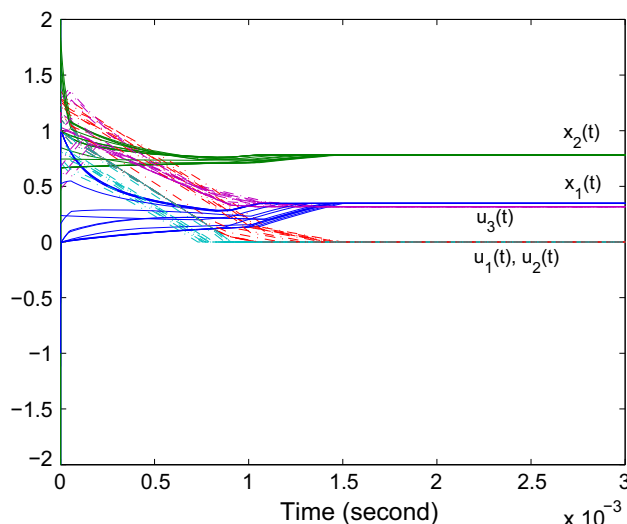


Fig. 2 State variables of Example 5.1 where obtained by the proposed NN of (10) with 16 initial points

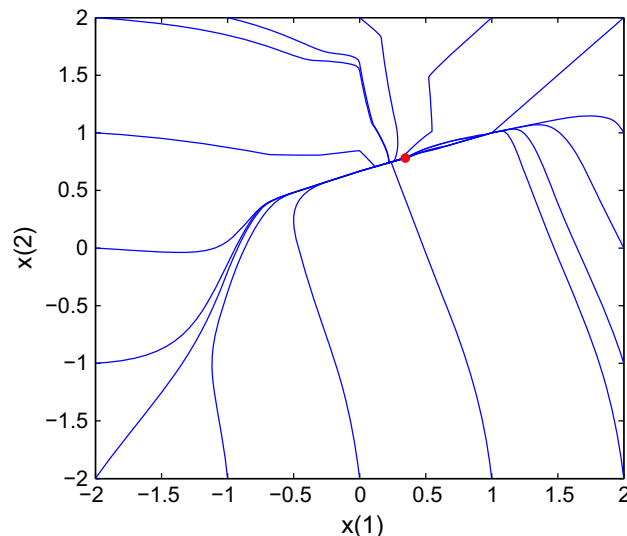
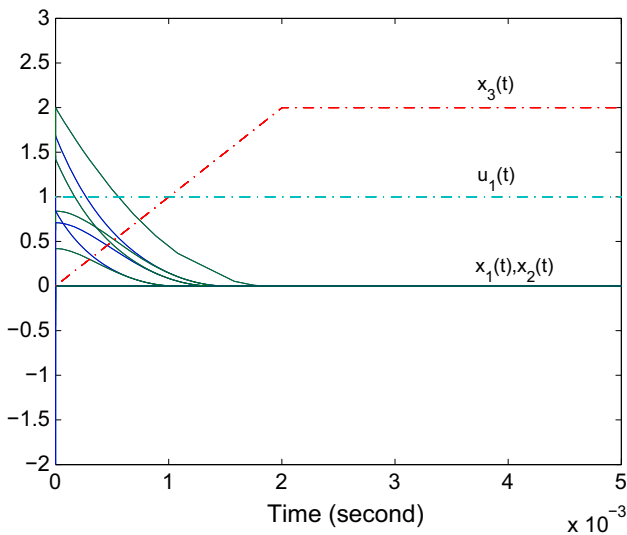


Fig. 3 Phase diagram of NN (10) with 16 different initial points in Example 5.1

variables  $(x_1(t), x_2(t))$  with 16 different initial points, which shows globally convergent to an exact optimal solution of Example 5.1.

**Example 5.2** Consider the following NLP problem (see Bazaraa and Shetty 1979):

$$\begin{aligned} \text{Min} \quad & f(x) = \frac{4}{3}(x_1^2 - x_1x_2 + x_2^2)^{\frac{3}{4}} - x_3 \\ \text{s.t.} \quad & x_3 \leq 2 \\ & x_i \geq 0, \quad i = 1, 2, 3. \end{aligned}$$



**Fig. 4** State variables of the Example 5.2 where obtained by the proposed NN of (10) with 20 initial points

In this problem, the objective function is convex and the feasible region is a convex set; also optimal solution is achieved at the unique point  $x^* = (0, 0, 2)$  (see Bazarraa and Shetty 1979). We apply the proposed NN in (10) to solve the above problem. Simulation results show the trajectory of (10) with 20 initial points is always convergent to

$$z^* = [0.0000, 0.0000, 2.0000, 1.0000]^T.$$

Figure 4 displays the transient behavior of  $z(t)$  with 20 initial points. Moreover, Fig. 5 shows the phase diagram of state variables  $(x_1(t), x_2(t), x_3(t))$  with 20 different initial points, which shows globally convergent to an exact optimal solution of Example 5.2.

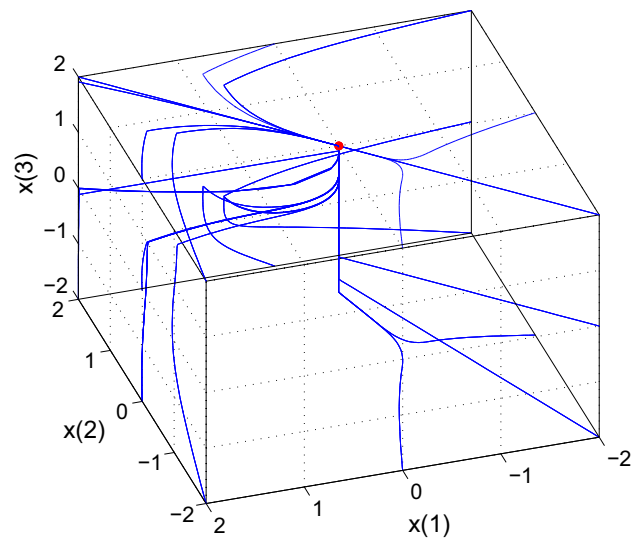
**Example 5.3** Consider the following NLP problem:

$$\begin{aligned} \text{Min} \quad & f(x) = (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_4)^4 \\ \text{s.t.} \quad & x_1^2 + x_2^2 + x_3^2 + x_4^2 \leq 10 \\ & (x_1 - 4)^2 + (x_2 + 4)^2 + (x_3 - 1)^2 + (x_4 + 1)^4 \leq 18 \\ & x_i \geq 0, \quad i = 1, 2, 3, 4. \end{aligned}$$

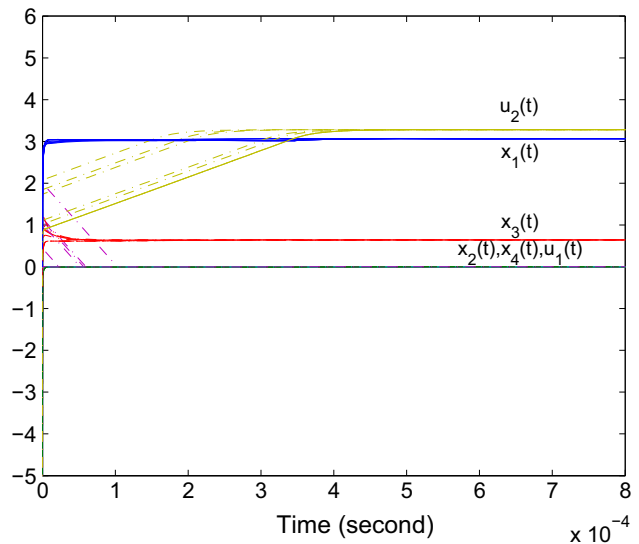
It is easy to see that the objective function is convex and constraint functions are strictly convex. The optimal solution is achieved at the unique point  $x^* = (3.0660, 0, 0.6426, 0)$ . We apply the proposed NN in (10) to solve the above problem. Simulation results show the trajectory of (10) with 10 initial points are always convergent to:

$$z^* = [3.0660, 0.0000, 0.6426, 0.0000, 0.0000, 3.2829]^T.$$

Figure 6 displays the transient behavior of  $z(t)$  with 20 initial points. Moreover, Fig. 7 shows the phase diagram of state



**Fig. 5** Phase diagram of NN (10) with 20 different initial points in Example 5.2



**Fig. 6** State variables of the Example 5.3 where obtained by the proposed NN of (10) with 20 initial points

variables  $(x_1(t), x_2(t), x_3(t))$  with 20 different initial points, which shows globally convergent to an exact optimal solution of Example 5.3.

**Example 5.4** Consider the following NLP problem:

$$\begin{aligned} \text{Min} \quad & f(x) = 0.4x_1 + x_1^2 + x_2^2 - x_1x_2 + 0.5x_3^2 + 0.5x_4^2 + \frac{1}{30}x_1^3 \\ \text{s.t.} \quad & -x_1 + x_2 - x_3 \leq 2 \\ & 3x_1 + x_2 - x_3 - x_4 \leq 18 \\ & \frac{1}{3}x_1 + x_2 - x_4 = 2 \\ & x_i \geq 0, \quad i = 1, 2, 3, 4. \end{aligned}$$

The optimal solution is achieved at the unique point  $x^* = (0.982, 1.672, 0, 0)$  (see Nazemi 2012). We apply the pro-



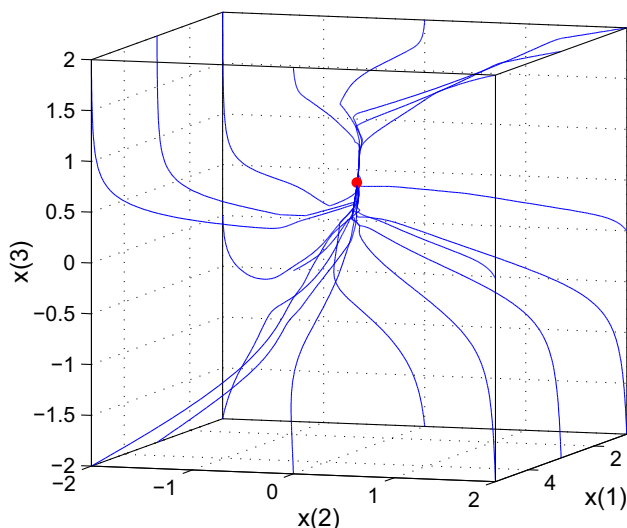


Fig. 7 Phase diagram of NN (10) with 20 different initial points in Example 5.3

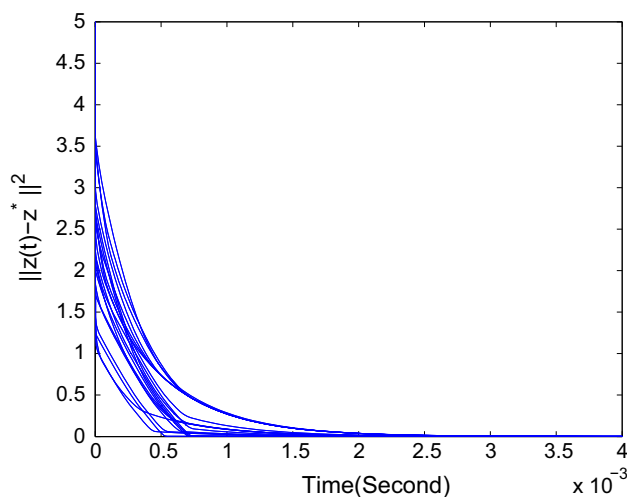


Fig. 9 Convergence behaviors of  $\|z(t) - z^*\|^2$  in Example 5.4 with 20 initial points

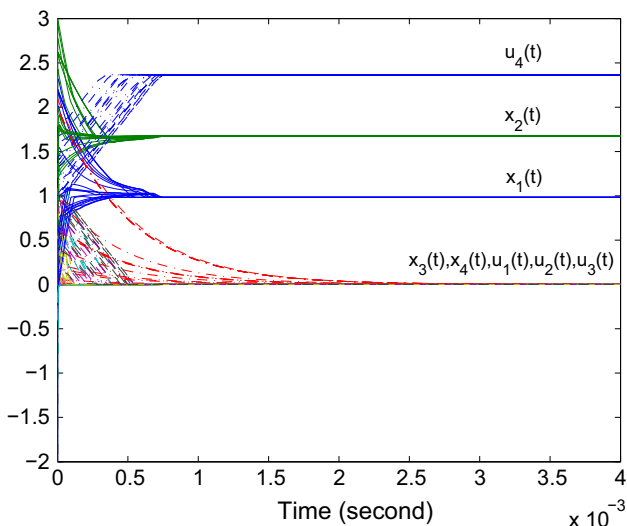


Fig. 8 State variables of the Example 5.4 where obtained by the proposed NN of (10) with 20 initial points

posed NN in (10) to solve the above problem. Simulation results show the trajectory of (10) with 20 initial points is always convergent to

$$z^* = [0.9820, 1.6727, 0.0000, 0.0000, 0.0000, 0.0000, 0.0000, 2.3635]^T.$$

Figure 8 displays the transient behavior of  $z(t)$  with 20 initial points. An  $l_2$  normal error between  $z(t)$  and  $z$  with 20 different initial points is also shown in Fig. 9.

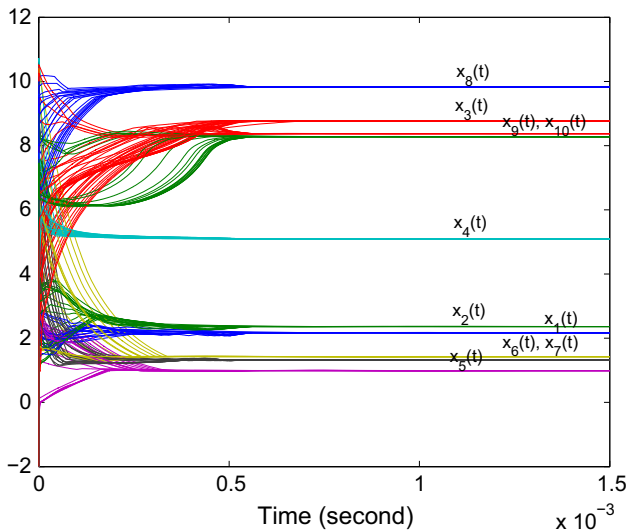
Example 5.5 Consider the following NLP problem:

$$\begin{aligned} \text{Min} \quad & f(x) = x_1^2 + x_2^2 + x_1x_2 - 14x_1 - 16x_2 + (x_3 - 10)^2 \\ & + 4(x_4 - 5)^2 + (x_5 - 3)^2 + 2(x_6 - 1)^2 + 5x_7^2 \\ & + 7(x_8 - 11)^2 + 2(x_9 - 10)^2 + (x_{10} - 7)^2 + 45 \\ \text{s.t.} \quad & 3(x_1 - 2)^2 + 4(x_2 - 3)^2 + 2x_3^2 - 7x_4 \leq 120 \\ & 5x_1^2 + 8x_2 + (x_3 - 6)^2 - 2x_4 \leq 40 \\ & \frac{1}{2}(x_1 - 8)^2 + 2(x_2 - 4)^2 + 3x_5^2 - x_6 \leq 30 \\ & x_1^2 + 2(x_2 - 2)^2 - 2x_1x_2 + 14x_5 - 6x_6 \leq 0 \\ & 4x_1 + 5x_2 - 3x_7 + 9x_8 \leq 105 \\ & 10x_1 - 8x_2 - 17x_7 + 2x_8 \leq 0 \\ & -3x_1 + 6x_2 + 12(x_9 - 8)^2 - 7x_{10} \leq 0 \\ & -8x_1 + 2x_2 + 5x_9 - 2x_{10} \leq 12 \\ & x_i \geq 0, \quad i = 1, 2, \dots, 10 \end{aligned}$$

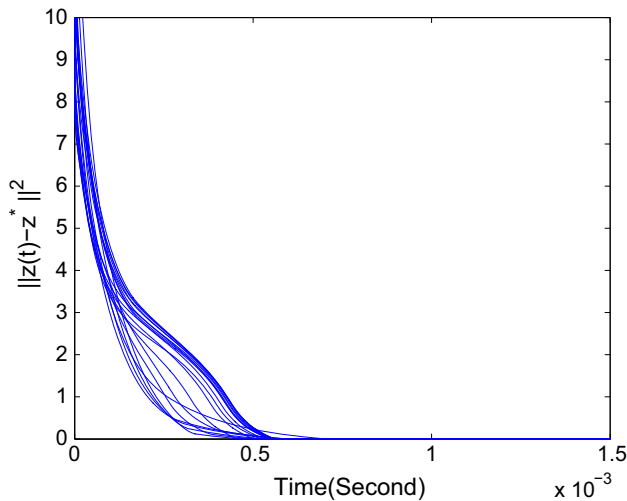
The optimal solution of this problem is given in Xia and Wang (2004), ( $x^* = [2.17199, 2.36368, 8.77392, 5.09598, 0.99065, 1.43057, 1.32164, 9.82872, 8.28009, 8.37592]^T$ ). We apply the proposed NN in (10) to solve the above problem. Simulation results show the trajectory of (10) with 20 initial points is always convergent to:

$$\begin{aligned} z^* = [x_1^* = 2.1720, x_2^* = 2.3637, x_3^* = 8.7739, \\ x_4^* = 5.0960, x_5^* = 0.9907, x_6^* = 1.4306 \\ x_7^* = 1.3216, x_8^* = 9.8287, x_9^* = 8.2801, \\ x_{10}^* = 8.3759, u_1^* = 0.0205, u_2^* = 0.3120 \\ u_3^* = 0.0000, u_4^* = 0.2870, u_5^* = 1.7165, \\ u_6^* = 0.4745, u_7^* = 0.0000, u_8^* = 1.3759]^T \end{aligned}$$

which corresponds to the optimal solution. Figure 10 shows the transient behavior of  $x(t)$  with 20 initial points. An  $l_2$  nor-



**Fig. 10** State variables of the Example 5.5 where obtained by the proposed NN of (10) with 20 initial points



**Fig. 11** Convergence behaviors of  $\|z(t) - z^*\|^2$  in Example 5.5 with 20 initial points

mal error between  $z(t)$  and  $z$  with 20 different initial points is also shown in Fig. 11.

In comparison with existing NN models, we use of the three NN models. First Kennedy and Chua (1988) proposed the following NN model for solving (5):

$$\frac{dx}{dt} = -\{\nabla f(x) + s(\nabla g(x)g(x)^+ - (-x)^+)\}, \quad (11)$$

where  $s$  is a penalty parameter. This NN has a low model complexity, but it only converges an approximate solution of (5) for any given finite penalty parameter. Afterward, Xia and Wang (2004) based on the projection formulation proposed a recurrent NN model for solving (5) with the following dynamical equation:

$$\frac{dx}{dt} = -x + (x - \alpha(\nabla f(x) + \nabla g(x)u))^+ \quad (12)$$

$$\frac{du}{dt} = -u + (u + \alpha g(x))^+, \quad (13)$$

where  $x \in \mathbb{R}^n$ ,  $u \in \mathbb{R}^m$  and  $\alpha > 0$ . This NN has a one-layer structure. Recently, Nazemi (2012) based on the Lagrange function proposed a NN model for solving the convex NLP problem, as the following form:

$$\begin{aligned} \text{Min} \quad & f(x) \\ \text{s.t.} \quad & g(x) \leq 0 \\ & h(x) = 0, \end{aligned} \quad (14)$$

with the following dynamical system:

$$\frac{dx}{dt} = -\left(\nabla f(x) + \frac{1}{2}\nabla g(x)^T u^2 + \nabla h(x)^T v\right) \quad (15)$$

$$\frac{du}{dt} = \text{diag}(u_1, \dots, u_m)g(x) \quad (16)$$

$$\frac{dv}{dt} = h(x). \quad (17)$$

Under the condition that the objective function is convex and constraint functions are strictly convex or that the objective function is strictly convex and the constraint functions are convex, Table 1 shows that for Examples 5.1–5.5, the proposed NN has a better performance in convergence time than the NNs introduced in (11), (12) and (13). Remark that the numerical implementation in all the models coded on MATLAB and the ordinary differential equation solver adopted ode15s. Note, initial states for all the implementations in Table 1 are equal. The stopping criterion in all the models is  $\|x(t_f) - x^*\| \leq 10^{-4}$  where  $t_f$  represents the final time when the stopping criterion is met.

Note that, in models (12) and (13), objective or constraint functions should be strictly convex. Therefore, since in Example 5.2 objective and constraint functions are convex, NNs (12) and (13) do not converge to the optimal solution.

It should be noted that the proposed model in (10) can solve LP and convex QP problems. In order to demonstrate, we give two examples in following, for LP and convex QP problems.

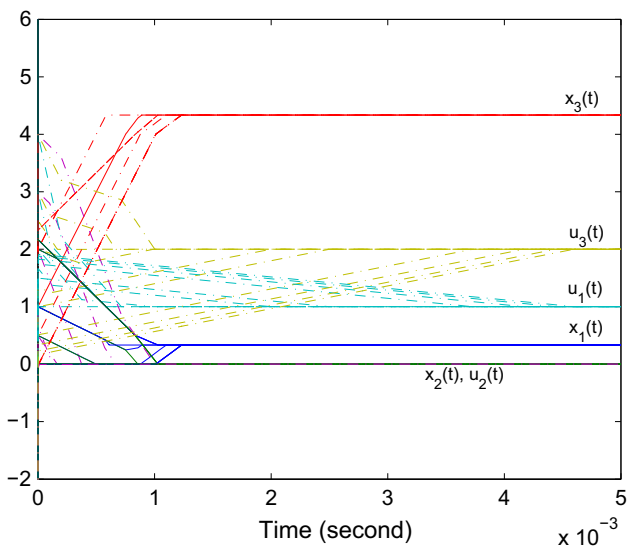
**Example 5.6** Consider the following LP problem (see Bazaraa et al. 1990):

$$\begin{aligned} \text{Min} \quad & f(x) = x_1 + x_2 - 4x_3 \\ \text{s.t.} \quad & x_1 + x_2 + 2x_3 \leq 9 \\ & x_1 + x_2 - x_3 \leq 2 \\ & -x_1 + x_2 + x_3 \leq 4 \\ & x_i \geq 0, \quad i = 1, 2, 3. \end{aligned}$$

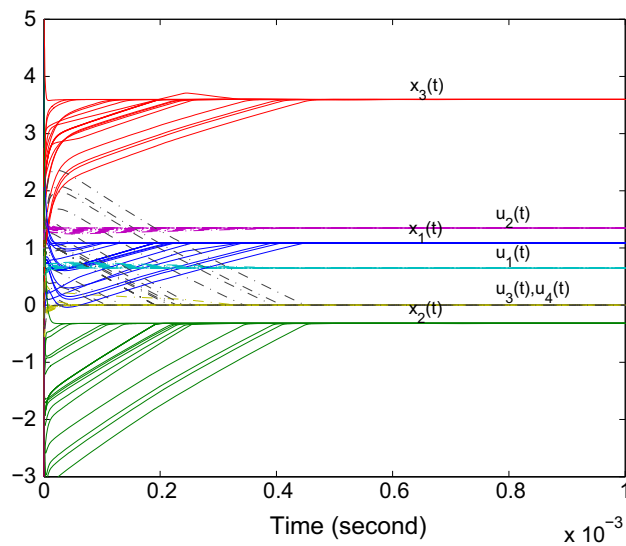


**Table 1** Final time (second) to stopping criterion for the proposed model in (10) compared to NN models in (11), (12) and (13) for Examples 5.1–5.5

Example/model	Proposed model (10)	Kennedy (11)	Xia (12)	Nazemi (13)
Time (s) in 5.1	0.002	10	70	20
Time (s) in 5.2	0.002	5	–	–
Time (s) in 5.3	0.0005	0.5	150	50
Time (s) in 5.4	0.003	10	40	40
Time (s) in 5.5	0.001	10	10	40



**Fig. 12** State variables of the Example 5.6 where obtained by the proposed NN of (10) with 10 initial points



**Fig. 13** State variables of the Example 5.7 where obtained by the proposed NN of (10) with 20 initial points

It is easy to see that objective function is convex and the feasible region is a convex set, and the optimal solution of LP problem is  $x^* = [\frac{1}{3}, 0, \frac{13}{3}]^T$  (see Bazaraa et al. 1990). We apply the proposed NN in (10) to solve the above problem. Simulation results show the trajectory of (10) with 10 initial points is always convergent to

$$z^* = [0.3333, 0.0000, 4.3333, 1.0000, 0.0000, 1.9999]^T.$$

Figure 12 displays the transient behavior of  $z(t)$  with 10 initial points.

**Example 5.7** Consider the following QP problem (see Huang and Cui 2016):

$$\begin{aligned} \text{Min} \quad & f(x) = 0.4x_1^2 + 0.3x_2^2 - 0.1x_1x_2 - 0.2x_1 \\ \text{s.t.} \quad & -0.4x_2 + 0.7x_3 \\ & x_1 - x_2 + x_3 = 5 \\ & 0.9x_1 + 0.2x_2 - 0.2x_3 \leq 4 \\ & 0.2x_1 + 0.7x_2 - 0.1x_3 \leq 10 \end{aligned}$$

This problem is a convex QP, and also optimal solution is achieved at the unique point  $x^* = (1.0851, -0.3191, 3.5957)$

(see Huang and Cui 2016). We apply the proposed NN in (10) to solve the above problem. Simulation results show the trajectory of (10) with 20 initial points is always convergent to

$$z^* = [1.0851, -0.3191, 3.5957, 0.6489, 1.3489, 0.0000, 0.0000]^T.$$

Figure 13 displays the transient behavior of  $z(t)$  with 20 initial points. Moreover, final time to stopping criterion for the proposed model in (10) is 0.002s and final time to stopping criterion for the proposed NN in Huang and Cui (2016) for  $k = 10$  is 2s, and it shows that the new model has a better performance in convergence time.

### 6 Conclusions

In this paper, we proposed a one-layer recurrent NN model for solving convex optimization problems. First by using KKT condition and with application MS merit function as a NCP function, we proposed a one-layer NN model, and in comparison with other existing NNs for solving such problems,

the proposed NN has a simple form and a better performance in convergence time. Moreover, the proposed NN is stable in the sense of Lyapunov and has globally convergent to optimal solution of the original problem.

## Compliance with ethical standards

**Conflict of interest** All authors have no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

- Bazaraa MS, Shetty CM (1979) Nonlinear programming. Theory and algorithms. Wiley, New York
- Bazaraa MS, Jarvis JJ, Sherali HD (1990) Linear programming and network flows, 2nd edn. Wiley, Hoboken
- Bouzerdoum A, Pattison TR (1993) Neural network for quadratic optimization with bound constraints. *IEEE Trans Neural Netw* 4(2):293–304
- Chen JS (2007) On some NCP-functions based on the generalized Fischer–Burmeister function. *Asia-Pac J Oper Res* 24:401–420
- Chen JS, Pan S (2008) A family of NCP functions and a descent method for the nonlinear complementarity problem. *Comput Optim Appl* 40:389–404
- Chen JS, Ko CH, Pan S (2010) A neural network based on the generalized Fischer–Burmeister function for nonlinear complementarity problems. *Inf Sci* 180:697–711
- Cottle RW, Pang JS, Stone RE (1992) The linear complementarity problem. Academic Press, New York
- Effati S, Baymani M (2005) A new nonlinear neural network for solving convex nonlinear programming problems. *Appl Math Comput* 168:1370–1379
- Effati S, Nazemi AR (2006) Neural network and its application for solving linear and quadratic programming problems. *Appl Math Comput* 172:305–331
- Effati S, Ranjbar M (2011) A novel recurrent nonlinear neural network for solving quadratic programming problems. *Appl Math Model* 35:1688–1695
- Effati S, Ghomashi A, Abbasi M (2011) A novel recurrent neural network for solving MLCPs and its application to linear and quadratic programming problems. *Asia-Pac J Oper Res* 28:523–541
- Effati S, Mansoori A, Eshaghnezhad M (2015) An efficient projection neural network for solving bilinear programming problems. *Neurocomputing* 168:1188–1197
- Facchinei F (1998) Structural and stability properties of  $P_0$  nonlinear complementarity problems. *Math Oper Res* 23(3):735–745
- Ferris MC, Kanzow C (2002) Complementarity and related problems: a survey. In: *Handbook of applied optimization*. Oxford University Press, New York, pp 514–530
- Ferris MC, Mangasarian OL, Pang JS (eds) (2001) Complementarity: applications, algorithms and extensions. Kluwer Academic Publishers, Dordrecht
- Fischer A (1992) A special Newton-type optimization method. *Optimization* 24:269–284
- Fischer A (1997) Solution of the monotone complementarity problem with locally Lipschitzian functions. *Math Program* 76:513–532
- He S, Zhang L, Zhang J (2015) The rate of convergence of a NLM based on F-B NCP for constrained optimization problems without strict complementarity. *Asia-Pac J Oper Res* 32(3):1550012
- Hu SL, Huang ZH, Chen JS (2009) Properties of a family of generalized NCP-functions and a derivative free algorithm for complementarity problems. *J Comput Appl Math* 230:69–82
- Huang X, Cui B (2016) A novel neural network for solving convex quadratic programming problems subject to equality and inequality constraints. *Neurocomputing* 214:23–31
- Kanzow C, Yamashita N, Fukushima M (1997) New NCP-functions and their properties. *J Optim Theory Appl* 94:115–135
- Kennedy MP, Chua LO (1988) Neural networks for nonlinear programming. *IEEE Trans Circuits Syst* 35:554–562
- Liao LZ, Qi H, Qi L (2001) Solving nonlinear complementarity problems with neural networks: a reformulation method approach. *J Comput Appl Math* 131:342–359
- Mangasarian OL, Sordov MV (1993) Nonlinear complementarity as unconstrained and constrained minimization. *Math Program* 62:277–297
- Miller RK, Michel AN (1982) Ordinary differential equations. Academic Press, New York
- Miri SM, Effati S (2015) On generalized convexity of nonlinear complementarity functions. *J Optim Theory Appl* 164(2):723–730
- Nazemi AR (2012) A dynamic system model for solving convex nonlinear optimization problems. *Commun Nonlinear Sci Numer Simul* 17:1696–1705
- Nazemi AR (2013) Solving general convex nonlinear optimization problems by an efficient neurodynamic model. *Eng Appl Artif Intell* 26(2):685–696
- Nazemi AR (2014) A neural network model for solving convex quadratic programming problems with some applications. *Eng Appl Artif Intell* 32:54–62
- Ranjbar M, Effati S, Miri SM (2017) An artificial neural network for solving quadratic zero-one programming problems. *Neurocomputing* 235:192–198
- Rodriguez VA, Domingues-Castro R, Rueda A, Huertas JL, Sanchez SE (1990) Nonlinear switched-capacitor neural networks for optimization problems. *IEEE Trans Circuits Syst* 37(3):384–397
- Slotin JJE, Li W (1991) Applied nonlinear control. Prentice Hall, Upper Saddle River
- Tank DW, Hopfield JJ (1986) Simple neural optimization networks: an A/D converter, signal decision network and a linear programming circuit. *IEEE Trans Circuits Syst* 33:533–541
- Wang J (1994) A deterministic annealing neural network for convex programming. *Neural Netw* 7(4):629–641
- Xia Y (1996) A new neural network for solving linear programming problems and its application. *IEEE Trans Neural Netw* 7:525–529
- Xia Y, Wang J (2000) A recurrent neural network for solving linear projection equations. *Neural Netw* 13:337–350
- Xia Y, Wang J (2004) A recurrent neural network for nonlinear convex optimization subject to nonlinear inequality constraints. *IEEE Trans Circuits Syst* 51(7):1385–1394
- Xia Y, Wang J (2005) A recurrent neural network for solving nonlinear convex programs subject to linear constraints. *IEEE Trans Neural Netw* 16(2):1045–1053
- Zhang XS (1996) Mathematical analysis of some neural networks for solving linear and quadratic programming. *Acta Mathematicae Applicatae Sinica* 12(1):1–10
- Zhang Z, Constantinides AG (1992) Lagrange programming neural networks. *IEEE Trans Circuits Syst* 39(7):441–452
- Zhang XS, Zhuo XJ, Jing ZJ (2002) An adaptive neural network model for nonlinear programming problems. *Acta Mathematicae Applicatae Sinica* 18(3):377–388

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.