

Neural networks based on three classes of NCP-functions for solving nonlinear complementarity problems[☆]

Jan Harold Alcantara, Jein-Shan Chen^{*}

Department of Mathematics National Taiwan Normal University Taipei 11677, Taiwan



ARTICLE INFO

Article history:

Received 5 February 2019

Revised 15 April 2019

Accepted 29 May 2019

Available online 31 May 2019

Communicated by Dr Q Wei

Keywords:

NCP-function

Natural residual function

Complementarity problem

Neural network

Stability

ABSTRACT

In this paper, we consider a family of neural networks for solving nonlinear complementarity problems (NCP). The neural networks are constructed from the merit functions based on three classes of NCP-functions: the generalized natural residual function and its two symmetrizations. In this paper, we first characterize the stationary points of the induced merit functions. Growth behavior of the complementarity functions is also described, as this will play an important role in describing the level sets of the merit functions. In addition, the stability of the steepest descent-based neural network model for NCP is analyzed. We provide numerical simulations to illustrate the theoretical results, and also compare the proposed neural networks with existing neural networks based on other well-known NCP-functions. Numerical results indicate that the performance of the neural network is better when the parameter p associated with the NCP-function is smaller. The efficiency of the neural networks in solving NCPs is also reported.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction and motivation

Given a function $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$, the nonlinear complementarity problem (NCP) is to find a point $x \in \mathbb{R}^n$ such that

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

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product and \geq means the component-wise order on \mathbb{R}^n . Throughout this paper, we assume that F is continuously differentiable, and let $F = (F_1, \dots, F_n)^T$ with $F_i: \mathbb{R}^n \rightarrow \mathbb{R}$ for $i = 1, \dots, n$.

For decades, substantial research efforts have been devoted in the study of nonlinear complementarity problems because of their wide range of applications in many areas such as optimization, operations research, engineering, and economics [8,9,12,48]. Some source problems of NCPs include models of equilibrium problems in the aforementioned fields and complementarity conditions in constrained optimization problems [9,12].

There are many methods in solving the NCP (1). In general, these solution methods may be categorized into two classes, depending on whether or not they make use of the so-called NCP-function (see Definition 2.1). Some techniques that usually exploit NCP-functions include merit function approach [11,19,26], nons-

mooth Newton method [10,45], smoothing methods [4,31], and regularization approach [17,37]. On the other hand, interior-point method [29,30] and proximal point algorithm [33] are some well-known approaches to solve (1) which do not utilize NCP-functions in general. The excellent monograph of Facchinei and Pang [9] provides a thorough survey and discussion of solution methods for complementarity problems and variational inequalities.

The above numerical approaches can efficiently solve the NCP; however, it is often desirable in scientific and engineering applications to obtain a real-time solution. One promising approach that can provide real-time solutions is the use of neural networks, which were first introduced in optimization by Hopfield and Tank in the 1980s [13,38]. Neural networks based on circuit implementation exhibit real-time processing. Furthermore, prior researches show that neural networks can be used efficiently in linear and nonlinear programming, variational inequalities and nonlinear complementarity problems [2,7,14,15,20,23,42–44,47,49] and as well as in other fields [25,28,34,36,39,40,46,50,51,55].

Motivated by the preceding discussion, we construct a new family of neural networks based on recently discovered discrete-type NCP-functions to solve NCPs. Neural networks based on the Fischer-Burmeister (FB) function [23] and the generalized Fischer-Burmeister function [2] have already been studied. The latter NCP-functions, which have been extensively used in the different solution methods, are strongly semismooth functions, which often provide efficient performance [9]. In this paper, we explore the use of smooth NCP-functions as building blocks of the proposed neural

[☆] The research is supported by Ministry of Science and Technology, Taiwan.

^{*} Corresponding author.

E-mail addresses: 80640005s@ntnu.edu.tw (J.H. Alcantara), jschen@math.ntnu.edu.tw (J.-S. Chen).

networks. Moreover, the NCP-functions we consider herein have piecewise-defined formulas, as opposed to the FB and generalized FB functions which have simple formulations. In turn, the subsequent analysis is more complicated. Nevertheless, we show that the proposed neural networks may offer promising results too. The analysis and numerical reports in this paper, on the other hand, pave the way for the use of piecewise-defined NCP-functions.

This paper is organized as follows: In Section 2, we revisit equivalent reformulations of the NCP (1) using NCP-functions. We also elaborate on the purpose and limitations of the paper. In Section 3, we review some mathematical preliminaries related to nonlinear mappings and stability analysis. We also summarize some important properties of the three classes of NCP-functions we used in constructing the neural networks. In Section 4, we describe the general properties of the neural networks, which include the characterization of stationary points of the induced merit functions. In Section 5, we look at the growth behavior of the three classes of NCP-functions considered. This result will be used to prove the boundedness of the level sets of the induced merit functions. We also prove some stability properties of the neural networks. In Section 6, we present the results of our numerical simulations. Conclusions and some recommendations for future studies are discussed in Section 7.

Throughout the paper, \mathbb{R}^n denotes the space of n -dimensional real column vectors, $\mathbb{R}^{m \times n}$ denotes the space of $m \times n$ real matrices, and A^T denotes the transpose of $A \in \mathbb{R}^{m \times n}$. For any differentiable function $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $\nabla f(x)$ means the gradient of f at x . For any differentiable mapping $F = (F_1, \dots, F_m)^T: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\nabla F(x) = [\nabla F_1(x) \ \dots \ \nabla F_m(x)] \in \mathbb{R}^{n \times m}$ denotes the transposed Jacobian of F at x . We assume that p is an odd integer greater than 1, unless otherwise specified.

2. Overview and contributions of the paper

In this section, we give an overview of this research. We begin by looking at equivalent reformulations of the nonlinear complementarity problem (1) using NCP-functions, which is defined as follows.

Definition 2.1. A function $\phi: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is called an NCP-function if it satisfies

$$\phi(a, b) = 0 \iff a \geq 0, \ b \geq 0, \ ab = 0.$$

The well-known natural-residual function given by

$$\phi_{\text{NR}}(a, b) = a - (a - b)_+ = \min\{a, b\}$$

is an example of an NCP-function, which is widely used in solving NCP. Recently, in [3], the discrete-type generalization of ϕ_{NR} is proposed and described by

$$\phi_{\text{NR}}^p(a, b) = a^p - (a - b)_+^p \quad \text{where } p > 1 \text{ is odd integer.} \quad (2)$$

It is shown in [3] that ϕ_{NR}^p is twice continuously differentiable. However, its surface is not symmetric, which may result to difficulties in designing and analyzing solution methods [16]. To conquer this, two symmetrizations of the ϕ_{NR}^p are presented in [1]. A natural symmetrization of ϕ_{NR}^p is given by

$$\phi_{\text{S-NR}}^p(a, b) = \begin{cases} a^p - (a - b)^p & \text{if } a > b, \\ a^p = b^p & \text{if } a = b, \\ b^p - (b - a)^p & \text{if } a < b. \end{cases} \quad (3)$$

The above NCP-function is symmetric, but is only differentiable on $\{(a, b) \mid a \neq b \text{ or } a = b = 0\}$. It was however shown in [16] that $\phi_{\text{S-NR}}^p$ is semismooth and is directionally differentiable. The second symmetrization of ϕ_{NR}^p is described by

$$\psi_{\text{S-NR}}^p(a, b) = \begin{cases} a^p b^p - (a - b)^p b^p & \text{if } a > b, \\ a^p b^p = a^{2p} & \text{if } a = b, \\ a^p b^p - (b - a)^p a^p & \text{if } a < b, \end{cases} \quad (4)$$

which possesses both differentiability and symmetry. The functions ϕ_{NR}^p , $\phi_{\text{S-NR}}^p$ and $\psi_{\text{S-NR}}^p$ are three classes of the four discrete-type families of NCP-functions which are recently discovered, together with the discrete-type generalization of the Fischer-Burmeister function given by

$$\phi_{\text{D-FB}}^p(a, b) = \left(\sqrt{x^2 + y^2} \right)^p - (x + y)^p. \quad (5)$$

A comprehensive discussion of their properties is presented in [16].

To see how an NCP-function ϕ can be useful in solving NCP (1), we define $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ by

$$\Phi(x) = \begin{pmatrix} \phi(x_1, F_1(x)) \\ \vdots \\ \phi(x_n, F_n(x)) \end{pmatrix}. \quad (6)$$

It is easy to see that x^* solves NCP (1) if and only if $\Phi(x^*) = 0$ (see also Proposition 4.1 (a)). Thus, the NCP is equivalent to the nonlinear system of equations $\Phi(x) = 0$. Meanwhile, if ϕ is an NCP-function, then $\psi: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_+$ given by

$$\psi(a, b) := \frac{1}{2} |\phi(a, b)|^2 \quad (7)$$

is also an NCP-function. Accordingly, if we define $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}_+$ by

$$\Psi(x) = \sum_{i=1}^n \psi(x_i, F_i(x)) = \frac{1}{2} \|\Phi(x)\|^2, \quad (8)$$

then the NCP can be reformulated as a minimization problem $\min_{x \in \mathbb{R}^n} \Psi(x)$. Hence, Ψ given by (8) is a merit function for the NCP, that is, its global minimizer coincides with the solution of the NCP. Consequently, it is only natural to consider the steepest descent-based neural network

$$\frac{dx(t)}{dt} = -\rho \nabla \Psi(x(t)), \quad x(t_0) = x^0, \quad (9)$$

where $\rho > 0$ is a time-scaling factor. The above neural network (9) is also motivated by the ones considered in [23] and in [2], where the NCP functions used are the Fischer-Burmeister (FB) function given by

$$\phi_{\text{FB}}(a, b) = \sqrt{a^2 + b^2} - (a + b), \quad (10)$$

and the generalized Fischer-Burmeister functions given by

$$\phi_{\text{FB}}^p(a, b) = \|(a, b)\|_p - (a + b) \quad \text{where } p \in (1, +\infty), \quad (11)$$

respectively. We aim to compare the neural networks based on the generalized natural-residual functions (2), (3) and (4) with the well-studied networks based on the FB functions (10) and (11).

One of the contributions of this paper lies on establishing the theoretical properties of the generalized natural residual functions. These are fundamental in designing NCP-based solution methods, and in this paper, we use the neural network approach. Basic properties of these functions are already presented in [16]. The purpose of this paper is to elaborate some more properties and applications of the newly discovered discrete-type classes of NCP-functions given by (2), (3) and (4). Specifically, we look at the properties of their induced merit functions given by (8). First, it is important for us to determine the correspondence between the solutions of NCP (1) and the stationary points of Ψ . From the above discussion (also see Proposition 4.1(d)), we already know that an NCP solution is a stationary point. On the other hand, we also want to determine which stationary points of Ψ are solutions to the NCP. For certain NCP functions such as the Mangasarian and

Solodov function [19], FB function [11] and generalized FB function [5], a stationary point of the merit function was shown to be a solution to the NCP when F is monotone or a P_0 -function. It should be pointed out that these NCP-functions possess the following nice properties:

- (P1) $\nabla_a \psi(a, b) \cdot \nabla_b \psi(a, b) \geq 0$ for all $(a, b) \in \mathbb{R}^2$; and
 (P2) For all $(a, b) \in \mathbb{R}^2$, $\nabla_a \psi(a, b) = 0 \iff \nabla_b \psi(a, b) = 0 \iff \phi(a, b) = 0$.

However, these properties are not possessed by ϕ_{NR}^p , $\phi_{\text{S-NR}}^p$ and $\psi_{\text{S-NR}}^p$, which leads to some difficulties in the subsequent analysis. Hence, we seek for other conditions which will guarantee that a stationary point is an NCP solution. Furthermore, we also want to look at the growth behavior of the functions (2), (3) and (4). This will play a key role in characterizing the level sets of the induced merit functions. It must be noted that since the NCP functions $\phi_{\text{S-NR}}^p$ and $\psi_{\text{S-NR}}^p$ are piecewise-defined functions, then the analyses of their growth behavior and the properties of their induced merit functions are more difficult, as compared with the commonly used FB functions (10) and (11) which have simple formulations.

Another purpose of this paper is to discuss the stability properties of the neural networks based on ϕ_{NR}^p , $\phi_{\text{S-NR}}^p$ and $\psi_{\text{S-NR}}^p$. We further look into different examples to see the influence of p on the convergence of trajectories of the neural network to the NCP solution. Finally, we compare the numerical performance of these three types of neural networks with two well-studied neural networks based on the FB function [23] and generalized FB function [2].

We recall that a solution x^* is said to be degenerate if $\{i | x_i^* = F_i(x^*) = 0\}$ is not empty. Note that if x^* is degenerate and ϕ is differentiable at x^* , then $\nabla \phi(x^*)$ is singular. Consequently, one should not expect a locally fast convergence of numerical methods based on smooth NCP-functions if the computed solution is degenerate [9,18]. Because of the differentiability of ϕ_{NR}^p , $\phi_{\text{S-NR}}^p$ and $\psi_{\text{S-NR}}^p$ on the feasible region of the NCP problem, it is also expected that the convergence of the trajectories of the neural network (9) to a degenerate solution could be slow. Hence, in this paper, we will give particular attention to nondegenerate NCPs.

3. Preliminaries

In this section, we review some special nonlinear mappings, some properties of ϕ_{NR}^p , $\phi_{\text{S-NR}}^p$ and $\psi_{\text{S-NR}}^p$, as well as some tools

$$\nabla \psi_{\text{S-NR}}^p(a, b) = \begin{cases} p[a^{p-1}b^p - (a-b)^{p-1}b^p, a^pb^{p-1} - (a-b)^pb^{p-1} + (a-b)^{p-1}b^p]^T & \text{if } a > b, \\ p[a^{p-1}b^p, a^pb^{p-1}]^T = pa^{p-1}[1, 1]^T & \text{if } a = b, \\ p[a^{p-1}b^p - (b-a)^pa^{p-1} + (b-a)^{p-1}a^p, a^pb^{p-1} - (b-a)^{p-1}a^p]^T & \text{if } a < b. \end{cases}$$

from stability theory in dynamical systems that will be crucial in our analysis. We begin with recalling concepts related to nonlinear mappings.

Definition 3.1. Let $F = (F_1, \dots, F_n)^T : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Then, the mapping F is said to be

- (a) monotone if $\langle x - y, F(x) - F(y) \rangle \geq 0$ for all $x, y \in \mathbb{R}^n$.
 (b) strictly monotone if $\langle x - y, F(x) - F(y) \rangle > 0$ for all $x, y \in \mathbb{R}^n$ and $x \neq y$.
 (c) strongly monotone with modulus $\mu > 0$ if $\langle x - y, F(x) - F(y) \rangle \geq \mu \|x - y\|^2$ for all $x, y \in \mathbb{R}^n$.
 (d) a P_0 -function if $\max_{\substack{1 \leq i \leq n \\ x_i \neq y_i}} (x_i - y_i)(F_i(x) - F_i(y)) \geq 0$ for all $x, y \in \mathbb{R}^n$ and $x \neq y$.

- (e) a P -function if $\max_{1 \leq i \leq n} (x_i - y_i)(F_i(x) - F_i(y)) > 0$ for all $x, y \in \mathbb{R}^n$ and $x \neq y$.
 (f) a uniform P -function with modulus $\kappa > 0$ if $\max_{1 \leq i \leq n} (x_i - y_i)(F_i(x) - F_i(y)) \geq \kappa \|x - y\|^2$, for all $x, y \in \mathbb{R}^n$.

From Definition 3.1, the following one-sided implications can be obtained:

F is strongly monotone $\implies F$ is a uniform P -function $\implies F$ is a P_0 -function.

It is known that F is monotone (resp. strictly monotone) if and only if $\nabla F(x)$ is positive semidefinite (resp. positive definite) for all $x \in \mathbb{R}^n$. In addition, F is a P_0 -function if and only if $\nabla F(x)$ is a P_0 -matrix for all $x \in \mathbb{R}^n$; that is, its principal minors are nonnegative. Further, if $\nabla F(x)$ is a P -matrix (that is, its principal minors are positive) for all $x \in \mathbb{R}^n$, then F is a P -function. However, we point out that a P -function does not necessarily have a Jacobian which is a P -matrix.

The following characterization of P -matrices and P_0 -matrices will be useful in our analysis.

Lemma 3.1. A matrix $M \in \mathbb{R}^{n \times n}$ is a P -matrix (resp. a P_0 -matrix) if and only if whenever $x_i(Mx)_i \leq 0$ (resp. $x_i(Mx)_i < 0$) for all i , then $x = 0$.

Proof. Please see [6]. \square

The following two lemmas summarize some properties of ϕ_{NR}^p , $\phi_{\text{S-NR}}^p$ and $\psi_{\text{S-NR}}^p$ that will be useful in our subsequent analysis.

Lemma 3.2. Let $p > 1$ be an odd integer. Then, the following hold.

- (a) The function ϕ_{NR}^p is twice continuously differentiable. Its gradient is given by

$$\nabla \phi_{\text{NR}}^p(a, b) = p \begin{bmatrix} a^{p-1} - (a-b)^{p-2}(a-b)_+ \\ (a-b)^{p-2}(a-b)_+ \end{bmatrix}.$$

- (b) The function $\phi_{\text{S-NR}}^p$ is twice continuously differentiable on the set $\Omega := \{(a, b) | a \neq b\}$. Its gradient is given by

$$\nabla \phi_{\text{S-NR}}^p(a, b) = \begin{cases} p[a^{p-1} - (a-b)^{p-1}, (a-b)^{p-1}]^T & \text{if } a > b, \\ p[(b-a)^{p-1}, b^{p-1} - (b-a)^{p-1}]^T & \text{if } a < b. \end{cases}$$

Further, $\phi_{\text{S-NR}}^p$ is differentiable at $(0,0)$ with $\nabla \phi_{\text{S-NR}}^p(0,0) = [0, 0]^T$.

- (c) The function $\psi_{\text{S-NR}}^p$ is twice continuously differentiable. Its gradient is given by

Proof. Please see [3, Proposition 2.2], [1, Propositions 2.2 and 3.2], and [16, Proposition 4.3]. \square

Lemma 3.3. Let $p > 1$ be a positive odd integer. Then, the following hold.

- (a) If $\phi \in \{\phi_{\text{NR}}^p, \phi_{\text{S-NR}}^p\}$, then $\phi(a, b) > 0 \iff a > 0, b > 0$. On the other hand, $\psi_{\text{S-NR}}^p(a, b) \geq 0$ on \mathbb{R}^2 .
 (b) $\nabla_a \phi_{\text{NR}}^p(a, b) \cdot \nabla_b \phi_{\text{NR}}^p(a, b) \begin{cases} > 0 & \text{on } \{(a, b) | a > b > 0 \text{ or } a > b > 2a\}, \\ = 0 & \text{on } \{(a, b) | a \leq b \text{ or } a > b = 2a \text{ or } a > b = 0\}, \\ < 0 & \text{otherwise,} \end{cases}$
 $\nabla_a \phi_{\text{S-NR}}^p(a, b) \cdot \nabla_b \phi_{\text{S-NR}}^p(a, b) > 0$ on $\{(a, b) | a > b > 0\} \cup \{(a, b) | b > a > 0\}$, and
 $\nabla_a \psi_{\text{S-NR}}^p(a, b) \cdot \nabla_b \psi_{\text{S-NR}}^p(a, b) > 0$ on the first quadrant \mathbb{R}_{++}^2 .

- (c) If $\phi \in \{\phi_{NR}^p, \phi_{S-NR}^p\}$, then $\nabla_a \phi(a, b) \cdot \nabla_b \phi(a, b) = 0$ provided that $\phi(a, b) = 0$. On the other hand, $\psi_{S-NR}^p(a, b) = 0 \iff \nabla \psi_{S-NR}^p(a, b) = 0$. In particular, we have $\nabla_a \psi_{S-NR}^p(a, b) \cdot \nabla_b \psi_{S-NR}^p(a, b) = 0$ provided that $\psi_{S-NR}^p(a, b) = 0$.

Proof. Please see [16, Propositions 3.4, 4.5, and 5.4]. \square

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

$$\dot{x}(t) = H(x(t)), \quad x(t_0) = x^0 \in \mathbb{R}^n \quad (12)$$

where $H: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a mapping. We also introduce three kinds of stability that we will consider later. These materials can be found in ODE textbooks; see [27].

Definition 3.2. A point $x^* = x(t^*)$ is called an equilibrium point or a steady state of the dynamic system (12) if $H(x^*) = 0$. If there is a neighborhood $\Omega^* \subseteq \mathbb{R}^n$ of x^* such that $H(x^*) = 0$ and $H(x) \neq 0 \forall x \in \Omega^* \setminus \{x^*\}$, then x^* is called an isolated equilibrium point.

Lemma 3.4. Assume that $H: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a continuous mapping. Then, for any $t_0 \geq 0$ and $x_0 \in \mathbb{R}^n$, there exists a local solution $x(t)$ for (12) with $t \in [t_0, \tau)$ for some $\tau > t_0$. If, in addition, H is locally Lipschitz continuous at x_0 , then the solution is unique; if H is Lipschitz continuous in \mathbb{R}^n , then τ can be extended to ∞ .

Definition 3.3. (Stability in the sense of Lyapunov) Let $x(t)$ be a solution for (12). An isolated equilibrium point x^* is Lyapunov stable if for any $x_0 = x(t_0)$ and any $\varepsilon > 0$, there exists a $\delta > 0$ such that $\|x(t) - x^*\| < \varepsilon$ for all $t \geq t_0$ and $\|x(t_0) - x^*\| < \delta$.

Definition 3.4. (Asymptotic stability) An isolated equilibrium point x^* is said to be asymptotically stable if in addition to being Lyapunov stable, it has the property that $x(t) \rightarrow x^*$ as $t \rightarrow \infty$ for all $\|x(t_0) - x^*\| < \delta$.

Definition 3.5. (Lyapunov function) Let $\Omega \subseteq \mathbb{R}^n$ be an open neighborhood of \bar{x} . A continuously differentiable function $W: \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be a Lyapunov function at the state \bar{x} over the set Ω for Eq. (12) if

$$\begin{cases} W(\bar{x}) = 0, & W(x) > 0, \quad \forall x \in \Omega \setminus \{\bar{x}\}. \\ \frac{dW(x(t))}{dt} = \nabla W(x(t))^T H(x(t)) \leq 0, & \forall x \in \Omega. \end{cases}$$

Lemma 3.5.

- (a) An isolated equilibrium point x^* is Lyapunov stable if there exists a Lyapunov function over some neighborhood Ω^* of x^* .
 (b) An isolated equilibrium point x^* is asymptotically stable if there is a Lyapunov function over some neighborhood Ω^* of x^* such that $\frac{dW(x(t))}{dt} < 0$ for all $x \in \Omega^* \setminus \{x^*\}$.

Definition 3.6. (Exponential stability) An isolated equilibrium point x^* is exponentially stable if there exists a $\delta > 0$ such that arbitrary point $x(t)$ of (12) with the initial condition $x(t_0) = x_0$ and $\|x(t_0) - x^*\| < \delta$ is well-defined on $[0, +\infty)$ and satisfies

$$\|x(t) - x^*\|_2 \leq ce^{-\omega t} \|x(t_0) - x^*\| \quad \forall t \geq t_0,$$

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

The following result will also be helpful in our stability analysis.

Lemma 3.6. Let F be locally Lipschitzian. If all $V \in \partial F(x)$ are nonsingular, then there is a neighborhood $N(x)$ of x and a constant C such that for any $y \in N(x)$ and any $V \in \partial F(y)$, V is nonsingular and $\|V^{-1}\| \leq C$

Proof. Please see [32, Propositions 3.1]. \square

4. Neural network model

In this section, we describe the properties of the neural network (9) based on the functions ϕ_{NR}^p , ϕ_{S-NR}^p and ψ_{S-NR}^p . Before this, we summarize first some important properties of Ψ as defined in (8) for general NCP-functions. Proposition 4.1 (a) is in fact Lemma 2.2 in [19]. On the other hand, Proposition 4.1(b) and (e) are true for all gradient systems (9).

Proposition 4.1. Let $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}_+$ be defined as in (8), with ϕ being any NCP-function, and let ψ be as in (7). Suppose that F is continuously differentiable. Then,

- (a) $\Psi(x) \geq 0$ for all $x \in \mathbb{R}^n$. If the NCP (1) has a solution, x is a global minimizer of $\Psi(x)$ if and only if x solves the NCP.
 (b) $\Psi(x(t))$ is a nonincreasing function of t , where $x(t)$ is a solution of (9).
 (c) Let $x \in \mathbb{R}^n$, and suppose that ϕ is differentiable at $(x_i, F_i(x))$ for each $i = 1, \dots, n$. Then

$$\nabla \Psi(x) = \nabla_a \psi(x, F(x)) + \nabla F(x) \nabla_b \psi(x, F(x)) \quad (13)$$

where

$$\begin{aligned} \nabla_a \psi(x, F(x)) &:= [\nabla_a \psi(x_1, F_1(x)), \dots, \nabla_a \psi(x_n, F_n(x))]^T, \\ \nabla_b \psi(x, F(x)) &:= [\nabla_b \psi(x_1, F_1(x)), \dots, \nabla_b \psi(x_n, F_n(x))]^T. \end{aligned}$$

- (d) Let x be a solution to the NCP such that ϕ is differentiable at $(x_i, F_i(x))$ for each $i = 1, \dots, n$. Then, x is a stationary point of Ψ .
 (e) Every accumulation point of a solution $x(t)$ of neural network (9) is an equilibrium point.

Proof. (a) It is clear that $\Psi \geq 0$. Notice that $\Psi(x) = 0$ if and only if $\Phi(x) = 0$, which occurs if and only if $\phi(x_i, F_i(x)) = 0$ for all i . Since ϕ is an NCP-function, this is equivalent to having $x_i \geq 0$, $F_i(x) \geq 0$ and $x_i F_i(x) = 0$. Thus, $\Psi(x) = 0$ if and only if $x \geq 0$, $F(x) \geq 0$ and $\langle x, F(x) \rangle = 0$. This proves part (a).

(b) The desired result follows from

$$\begin{aligned} \frac{d\Psi(x(t))}{dt} &= \nabla \Psi(x(t))^T \frac{dx}{dt} = \nabla \Psi(x(t))^T (-\rho \nabla \Psi(x(t))) \\ &= -\rho \|\nabla \Psi(x(t))\|^2 \leq 0 \end{aligned}$$

for all solutions $x(t)$.

(c) The formula is clear from chain rule.

(d) First, note that from Eq. (7), we have $\nabla \psi(a, b) = \phi(a, b) \cdot \nabla \phi(a, b)$. Thus, if x is a solution to the NCP, it gives $\nabla \psi(x_i, F_i(x)) = 0$ for all $i = 1, \dots, n$. Then, it follows from formula (13) in part(c) that $\nabla \Psi(x) = 0$. That is, x is a stationary point of Ψ .

(e) Please see page 232 of [41]. \square

We adopt the neural network (9) with $\Psi(x) = \frac{1}{2} \|\Phi(x)\|^2$, where Φ is given by (6) with $\phi \in \{\phi_{NR}^p, \phi_{S-NR}^p, \psi_{S-NR}^p\}$. The function Φ corresponding to ϕ_{NR}^p , ϕ_{S-NR}^p and ψ_{S-NR}^p is denoted, respectively, by Φ_{NR}^p , Φ_{S1-NR}^p and Φ_{S2-NR}^p . Their corresponding merit functions will be denoted by Ψ_{NR}^p , Ψ_{S1-NR}^p and Ψ_{S2-NR}^p , respectively. We note that by formula (13) and the differentiability of $\Psi \in \{\Psi_{NR}^p, \Psi_{S1-NR}^p, \Psi_{S2-NR}^p\}$ (see Proposition 4.2), the neural network (9) can be implemented on hardware as in Fig. 1.

We first establish the existence and uniqueness of the solutions of neural network (9).

Proposition 4.2. Let $p > 1$ be an odd integer. Then, the following hold.

- (a) Ψ_{NR}^p and Ψ_{S2-NR}^p are both continuously differentiable on \mathbb{R}^n .
 (b) Ψ_{S1-NR}^p is continuously differentiable on the open set $\Omega = \{x \in \mathbb{R}^n \mid x_i \neq F_i(x), \forall i = 1, 2, \dots, n\}$.

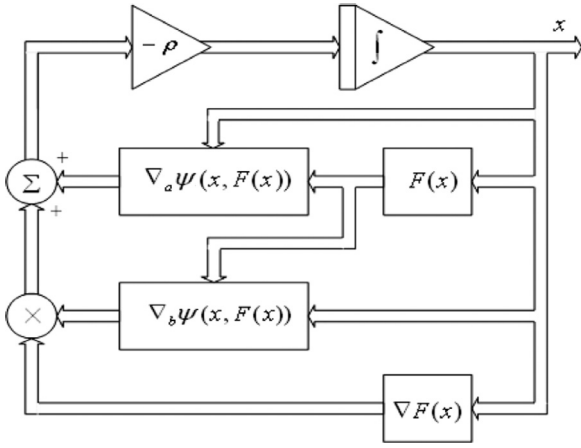


Fig. 1. Simplified block diagram for neural network (9). This figure is lifted from Chen et al. [2].

Consequently, the neural network (9) with Ψ_{NR}^p or $\Psi_{\text{S2-NR}}^p$ has a unique solution for all $x^0 \in \mathbb{R}^n$. The neural network (9) with $\Psi_{\text{S1-NR}}^p$ has a unique solution for all $x^0 \in \Omega$.

Proof. Part (a) and (b) directly follow from Proposition 4.1(c) and Lemma 3.2. The existence and uniqueness of the solutions follows from Lemma 3.4, noting the continuous differentiability of F and Ψ_{NR}^p , $\Psi_{\text{S1-NR}}^p$ (on Ω), and $\Psi_{\text{S2-NR}}^p$. \square

We note that because of Proposition 4.2(b), we only consider the neural network (9) with $\Psi = \Psi_{\text{S1-NR}}^p$ as a dynamical system defined on the set Ω . Our next goal is to determine the conditions such that equilibrium points of (9) are global minimizers of Ψ . When an NCP-function has properties (P1) and (P2) (see Introduction), an equilibrium point is a global minimizer when F is a P_0 -function. However, these properties only hold on a proper subset of \mathbb{R}^n for the functions ϕ_{NR}^p , $\phi_{\text{S-NR}}^p$ and $\psi_{\text{S-NR}}^p$. Thus, we seek for other conditions to achieve the goal. We start with the merit function Ψ_{NR}^p .

Proposition 4.3. If F is strongly monotone with modulus $\mu > 1$, then every stationary point of Ψ_{NR}^p is a global minimizer.

Proof. Let x^* be a stationary point of Ψ_{NR}^p , that is, $\nabla \Psi_{\text{NR}}^p(x^*) = 0$. For convenience, we denote by $A(x^*)$ and $B(x^*)$ the diagonal matrices such that for each $i = 1, \dots, n$,

$$A_{ii}(x^*) = (x_i^*)^{p-1} \quad \text{and} \quad B_{ii}(x^*) = (x_i^* - F_i(x^*))^{p-2} (x_i^* - F_i(x^*))_+.$$

Then, by formula (13) and Lemma 3.2 (b), we have

$$p[A(x^*) - B(x^*)]\Phi_{\text{NR}}^p(x^*) + p\nabla F(x^*)B(x^*)\Phi_{\text{NR}}^p(x^*) = 0,$$

which yields

$$A(x^*)\Phi_{\text{NR}}^p(x^*) + (\nabla F(x^*) - I)B(x^*)\Phi_{\text{NR}}^p(x^*) = 0. \quad (14)$$

Analogous to the technique in [11], pre-multiplying both sides of (14) by $(B(x^*)\Phi_{\text{NR}}^p(x^*))^T$ leads to

$$\begin{aligned} & \Phi_{\text{NR}}^p(x^*)^T [B(x^*)A(x^*)]\Phi_{\text{NR}}^p(x^*) \\ & + (B(x^*)\Phi_{\text{NR}}^p(x^*))^T (\nabla F(x^*) - I)B(x^*)\Phi_{\text{NR}}^p(x^*) = 0. \end{aligned} \quad (15)$$

Since p is odd integer, we have $A(x^*) \geq 0$ and $B(x^*) \geq 0$; and hence,

$$\Phi_{\text{NR}}^p(x^*)^T [B(x^*)A(x^*)]\Phi_{\text{NR}}^p(x^*) \geq 0.$$

On the other hand, since F is strongly monotone with modulus $\mu > 1$, defining $G(x) := F(x) - x$ gives

$$\begin{aligned} \langle x - y, G(x) - G(y) \rangle &= \langle x - y, F(x) - x - F(y) + y \rangle \\ &= \langle x - y, F(x) - F(y) \rangle - \|x - y\|^2 \\ &\geq (\mu - 1)\|x - y\|^2 \\ &> 0, \end{aligned}$$

for all $x, y \in \mathbb{R}^n$. Note then that $\nabla G(x) = \nabla F(x) - I$ is positive definite. Consequently, each term of the left-hand side of (15) is non-negative. With $(\nabla F(x^*) - I)$ being positive definite, it yields $B(x^*)\Phi_{\text{NR}}^p(x^*) = 0$. In addition, from (14), we have $A(x^*)\Phi_{\text{NR}}^p(x^*) = 0$. To sum up, we have proved that $A_{ii}(x^*)\phi_{\text{NR}}^p(x_i^*, F_i(x^*)) = 0$ and $B_{ii}(x^*)\phi_{\text{NR}}^p(x_i^*, F_i(x^*)) = 0$ for all i .

Now, if $\phi_{\text{NR}}^p(x_i^*, F_i(x^*)) \neq 0$ for some i , then we must have $A_{ii}(x^*) = B_{ii}(x^*) = 0$. Thus, $(x_i^*)^{p-1} = 0$ (i.e., $x_i^* = 0$), and $x_i^* \leq F_i(x^*)$. Since ϕ_{NR}^p is an NCP-function, the latter implies that $\phi_{\text{NR}}^p(x_i, F_i(x^*)) = 0$ for all i , that is, x^* is a global minimizer of Ψ_{NR}^p . This completes the proof. \square

The following proposition provides a weaker condition on F to guarantee that a stationary point of Ψ_{NR}^p is a global minimizer.

Proposition 4.4. If $(\nabla F - I)$ is a P -matrix, then every stationary point of Ψ_{NR}^p is a global minimizer.

Proof. Suppose that $\nabla \Psi_{\text{NR}}^p(x^*) = 0$. If $B(x^*)\Phi_{\text{NR}}^p(x^*) = 0$, then $A(x^*)\Phi_{\text{NR}}^p(x^*) = 0$ by Eq. (14). As in the preceding proof, we obtain $\Phi_{\text{NR}}^p(x^*) = 0$, and hence we are done. It remains to consider another case that $B(x^*)\Phi_{\text{NR}}^p(x^*) \neq 0$. Note that

$$\begin{aligned} (B(x^*)\Phi_{\text{NR}}^p(x^*))_i &= (x_i^* - F_i(x^*))^{p-2} (x_i^* - F_i(x^*))_+ \phi_{\text{NR}}^p(x_i^*, F_i(x^*)) \\ &= \begin{cases} 0 & \text{if } x_i^* \leq F_i(x^*) \text{ or } x_i^* > F_i(x^*) = 0, \\ (x_i^* - F_i(x^*))^{p-1} \phi_{\text{NR}}^p(x_i^*, F_i(x^*)) & \text{if } x_i^* > F_i(x^*) \text{ and } F_i(x^*) \neq 0. \end{cases} \end{aligned}$$

Thus, the nonzero entries of $B(x^*)\Phi_{\text{NR}}^p(x^*)$ appear at indices i where $x_i^* > F_i(x^*)$ and $F_i(x^*) \neq 0$. To proceed, we denote

$$I_1 = \{i \mid x_i^* \neq 0 \text{ and } (B(x^*)\Phi_{\text{NR}}^p(x^*))_i \neq 0\},$$

$$I_2 = \{i \mid x_i^* = 0 \text{ and } (B(x^*)\Phi_{\text{NR}}^p(x^*))_i \neq 0\}.$$

With these notations, we observe the following facts.

- (i) For $i \in I_1$, since p is odd, it is clear that the i th entry of $A(x^*)\Phi_{\text{NR}}^p(x^*)$ and $B(x^*)\Phi_{\text{NR}}^p(x^*)$ are both nonzero and have the same sign.
- (ii) For $i \in I_2$, then $(B(x^*)\Phi_{\text{NR}}^p(x^*))_i \neq 0$ and $(A(x^*)\Phi_{\text{NR}}^p(x^*))_i = 0$.

Because $(\nabla F - I)$ is a P -matrix, it follows from Lemma 3.1 that there exists an index j such that

$$(B(x^*)\Phi_{\text{NR}}^p(x^*))_j [(\nabla F(x^*) - I)(B(x^*)\Phi_{\text{NR}}^p(x^*))]_j > 0.$$

This says that $(B(x^*)\Phi_{\text{NR}}^p(x^*))_j \neq 0$ and therefore $j \in I_1 \cup I_2$. Note that by (i) above, $(A(x^*)\Phi_{\text{NR}}^p(x^*))_i$ and $(B(x^*)\Phi_{\text{NR}}^p(x^*))_i$ have the same sign if $j \in I_1$ which will contradict Eq. (14). On the other hand, if $j \in I_2$, we have from fact (ii) that $(A(x^*)\Phi_{\text{NR}}^p(x^*))_j = 0$. However, we also have that $[(\nabla F(x^*) - I)(B(x^*)\Phi_{\text{NR}}^p(x^*))]_j \neq 0$. This certainly violates Eq. (14). Thus, we conclude that $B(x^*)\Phi_{\text{NR}}^p(x^*) = 0$, and hence $\Phi_{\text{NR}}^p(x^*) = 0$. Then, the proof is complete. \square

Remark 4.1. In fact, if the function F is nonnegative (or if we at least have $F(x^*) \geq 0$ for an equilibrium point x^*), then case (ii) in the above proof cannot happen. Thus, the above theorem is valid even when $(\nabla F - I)$ is a P_0 -matrix by Lemma 3.1.

From Lemma 2.2(b) and Lemma 2.2(c), we see that the structures of $\nabla \Phi_{\text{NR}}^p$ and $\nabla \Psi_{\text{S2-NR}}^p$ corresponding to the NCP-functions ϕ_{NR}^p and $\psi_{\text{S-NR}}^p$ are complex because of the piecewise nature of ϕ_{NR}^p and $\psi_{\text{S-NR}}^p$. This makes it difficult to find conditions on F so

that a stationary point of Ψ_{S1-NR}^p or Ψ_{S2-NR}^p is also a global minimizer. However, if F is a nonnegative function, we have the following proposition.

Proposition 4.5. Suppose that F is a nonnegative P_0 -function and $x^* \geq 0$. If x^* is a stationary point of Ψ_{S1-NR}^p or Ψ_{S2-NR}^p , then it is a global minimizer.

Proof. If we can show that properties (P1) and (P2) mentioned in the Introduction Section hold for ϕ_{S-NR}^p and ψ_{S-NR}^p on the non-negative quadrant \mathbb{R}_+^2 , then we can proceed as in the proof of [5, Proposition 3.4]. Thus, it is enough to show that (P1) and (P2) hold on \mathbb{R}_+^2 . To simplify our notations, we denote $\phi_1 = \phi_{S-NR}^p$, $\phi_2 = \psi_{S-NR}^p$, and $\psi_i = \frac{1}{2}|\phi_i|^2$ ($i = 1, 2$). Note that the domain of $\nabla \Psi_{S1-NR}^p$ is $\{x | x_i \neq F_i(x) \text{ or } x_i = F_i(x) = 0\}$. Thus, for ψ_1 , it suffices to check that it has properties (P1) and (P2) only on the set $\{(a, b) \in \mathbb{R}_+^2 | a \neq b \text{ or } a = b = 0\}$.

To proceed, we observe that

$$\nabla_a \psi_i(a, b) = \phi_i(a, b) \nabla_a \phi_i(a, b) \quad \text{and}$$

$$\nabla_b \psi_i(a, b) = \phi_i(a, b) \nabla_b \phi_i(a, b),$$

which imply

$$\nabla_a \psi_i(a, b) \cdot \nabla_b \psi_i(a, b) = (\phi_i(a, b))^2 \cdot \nabla_a \phi_i(a, b) \cdot \nabla_b \phi_i(a, b),$$

$$i = 1, 2.$$

If $a \geq b = 0$ or $b \geq a = 0$, then $\phi_i(a, b) = 0$; and thus, the above product is zero. Otherwise, the above product is positive by Lemma 3.3(b). This asserts (P1).

To show (P2), note that it is obvious that $\nabla_a \psi_i(a, b) = \nabla_b \psi_i(a, b) = 0$ if $\phi_i(a, b) = 0$ for $i = 1, 2$.

To show the converse, it is enough to argue that if $\nabla_a \phi_i(a, b) = 0$ or $\nabla_b \phi_i(a, b) = 0$, then $\phi_i(a, b) = 0$. First, we analyze the case for ϕ_1 . Suppose that $\nabla_a \phi_1(a, b) = 0$. From Lemma 2.2(c),

$$\frac{1}{p} \nabla_a \phi_1(a, b) = \begin{cases} a^{p-1} - (a-b)^{p-1} & \text{if } a > b \\ 0 & \text{if } a = b = 0 \\ (b-a)^{p-1} & \text{if } a < b \end{cases}$$

For $a = b = 0$, then $\phi_1(a, b) = 0$. For $a > b$, then $a = |a-b| = a-b$ since p is an odd integer. Thus, $b = 0$ and because $a > b$, we obtain $\phi_1(a, b) = 0$. For $a < b$, we have from (4) that $(b-a)^{p-1} = 0$, which is impossible. This proves that $\nabla_a \phi_1(a, b) = 0$ implies that $\phi_1(a, b) = 0$. Similarly, we can show that $\nabla_b \phi_1(a, b) = 0$ implies that $\phi_1(a, b) = 0$. This asserts (P2) for the function ψ_1 .

Analogously, for ψ_2 , assume that $\nabla_a \phi_2(a, b) = 0$. From Lemma 2.2(d), we have

$$\frac{1}{p} \nabla_a \phi_2(a, b) = \begin{cases} a^{p-1}b^p - (a-b)^{p-1}b^p & \text{if } a > b, \\ a^{2p-1} & \text{if } a = b, \\ a^{p-1}b^p - (b-a)^pa^{p-1} + (b-a)^{p-1}a^p & \text{if } a < b. \end{cases}$$

For $a = b$, then $a^{2p-1} = 0$, and hence $a = 0$ and $\phi_2(a, b) = 0$. For $a > b$, then $a^{p-1}b^p - (a-b)^{p-1}b^p = 0$. For $b = 0$, we obtain $\phi_2(a, b) = 0$ by using $a > b$. Otherwise, $a^{p-1} - (a-b)^{p-1} = 0$. Because p is odd and $a > b$, we have $a = |a-b| = a-b$. Consequently, $b = 0$ and $\phi_2(a, b) = 0$. For $a < b$, then we have from the above formula for $\nabla_a \phi_2$ that $a^{p-1}b^p - (b-a)^pa^{p-1} + (b-a)^{p-1}a^p = 0$. For $a = 0$, then $\phi_2(a, b) = 0$ due to $a < b$. Otherwise, $a > 0$ and

$$\begin{aligned} 0 &= b^p - (b-a)^p + (b-a)^{p-1}a \\ &= b^p - (b-a)^{p-1}(b-2a) \\ &= (a+k)^p - k^{p-1}(k-a) \quad \text{where } k = b-a > 0 \\ &= \sum_{i=0}^{p-1} \binom{p}{i} a^{p-i} k^i + ak^{p-1} \\ &> 0 \end{aligned}$$

which is a contradiction. To sum up, we have shown that $\nabla_a \phi_2(a, b) = 0$ implies that $\phi_2(a, b) = 0$. Similarly, it can be verified $\phi_2(a, b) = 0$ provided $\nabla_b \phi_2(a, b) = 0$. Thus, ψ_2 possesses the property (P2). This completes the proof. \square

5. Stability analysis

We now look at the properties of the neural network (9) related to the behavior of its solutions. We have the following consequences, which easily follow from Propositions 4.1(a), 4.1(d), 4.4, and 4.5.

Proposition 5.1. Consider the neural network (9) with $\Psi \in \{\Psi_{NR}^p, \Psi_{S1-NR}^p, \Psi_{S2-NR}^p\}$.

- (a) Every solution of the NCP is an equilibrium point.
- (b) If $(\nabla F - I)$ is a P -matrix, then every equilibrium point of (9) with $\Psi = \Psi_{NR}^p$ solves the NCP.
- (c) If F is a nonnegative P_0 -function, every equilibrium point $x^* \geq 0$ of (9) with $\Psi \in \{\Psi_{S1-NR}^p, \Psi_{S2-NR}^p\}$ solves the NCP.

Theorem 5.1 below addresses the boundedness of the level sets of Ψ and convergence of the trajectories of the neural network. Before we state this theorem, we need the following lemma.

Lemma 5.1. Let $\{(a^k, b^k)\}_{k=1}^\infty \subseteq \mathbb{R}^2$ such that $|a^k| \rightarrow \infty$ and $|b^k| \rightarrow \infty$ as $k \rightarrow \infty$. Then, $|\phi_{NR}^p(a^k, b^k)| \rightarrow \infty$, $|\phi_{S-NR}^p(a^k, b^k)| \rightarrow \infty$, and $|\psi_{S-NR}^p(a^k, b^k)| \rightarrow \infty$.

Proof. (a) First, we verify that $|\phi_{S-NR}^p(a^k, b^k)| \rightarrow \infty$. To proceed, we consider three cases.

(i) Suppose $a^k \rightarrow \infty$ and $b^k \rightarrow \infty$. Note that for all $x \in [-1, 0]$ and $n \in \mathbb{N}$, there holds

$$(1+x)^n \leq (1-nx)^{-1}$$

which is a useful inequality. Thus, when $a > b > 0$, we have

$$\begin{aligned} \phi_{S-NR}^p(a, b) &= a^p - (a-b)^p = a^p - a^p \left(1 - \frac{b}{a}\right)^p \\ &\geq a^p - a^p \left(1 - p\left(-\frac{b}{a}\right)\right)^{-1} \\ &= a^p - a^p \left(\frac{a}{a+pb}\right) \\ &= \frac{pa^p b}{a+pb} \\ &\geq \frac{pa^{p-1}b}{1+p} \\ &\geq \frac{pb^p}{1+p}. \end{aligned}$$

Similarly, $\phi_{S-NR}^p(a, b) \geq \frac{pa^p}{1+p}$ for $b > a > 0$. Thus, $\phi_{S-NR}^p(a^k, b^k) \rightarrow \infty$ as $k \rightarrow \infty$.

(ii) Suppose $a^k \rightarrow -\infty$ and $b^k \rightarrow -\infty$. Observe that $\phi_{S-NR}^p(a, b) \leq a^p$ when $a > b$, and $\phi_{S-NR}^p(a, b) \leq b^p$ when $a < b$. Thus, $\phi_{S-NR}^p(a^k, b^k) \rightarrow -\infty$ as $k \rightarrow \infty$.

(iii) Suppose $a^k \rightarrow \infty$ and $b^k \rightarrow -\infty$. For $a > 0$ and $b < 0$, we have

$$(a-b)^p \geq a^p + (-b)^p = a^p - b^p.$$

Thus, $\phi_{S-NR}^p(a, b) = a^p - (a-b)^p \leq b^p$ and we conclude that $\phi_{S-NR}^p(a^k, b^k) \rightarrow -\infty$ as $k \rightarrow \infty$. In the case that $a^k \rightarrow -\infty$ and $b^k \rightarrow \infty$, we also have $\phi_{S-NR}^p(a^k, b^k) \rightarrow -\infty$ as $k \rightarrow \infty$ by symmetry of ϕ_{S-NR}^p .

(b) Next, we show that $|\phi_{NR}^p(a^k, b^k)| \rightarrow \infty$. Again, we consider three cases.

(i) Suppose that $a^k \rightarrow -\infty$. Since $\phi_{\text{NR}}^p(a, b) = a^p - (a - b)_+^p \leq a^p$ for all $(a, b) \in \mathbb{R}^2$, it is trivial to see that $\phi_{\text{NR}}^p(a^k, b^k) \rightarrow -\infty$.

(ii) Suppose that $a^k \rightarrow \infty$ and $b^k \rightarrow \infty$. For $a > b > 0$, then we have

$$\phi_{\text{NR}}^p(a, b) = \phi_{\text{S-NR}}^p(a, b) \geq \frac{pb^p}{1+p}.$$

For $0 \leq a < b$, it is clear that $\phi_{\text{NR}}^p(a, b) = a^p$. Then, we conclude that $\phi_{\text{NR}}^p(a^k, b^k) \rightarrow \infty$.

(iii) Suppose that $a^k \rightarrow \infty$ and $b^k \rightarrow -\infty$. For $a > 0$ and $b < 0$, we have

$$\phi_{\text{NR}}^p(a, b) = \phi_{\text{S-NR}}^p(a, b) \leq b^p$$

and so $\phi_{\text{NR}}^p(a^k, b^k) \rightarrow -\infty$. Thus, we have proved that $|\phi_{\text{NR}}^p(a^k, b^k)| \rightarrow \infty$.

(c) The last limit, $|\psi_{\text{S-NR}}^p(a^k, b^k)| \rightarrow \infty$, follows from the fact that

$$\psi_{\text{S-NR}}^p(a, b) = \begin{cases} \phi_{\text{S-NR}}^p(a, b)b^p & \text{if } a > b, \\ a^p b^p = a^{2p} & \text{if } a = b, \\ \phi_{\text{S-NR}}^p(b, a)a^p & \text{if } a < b. \end{cases}$$

and the inequalities obtained above for $\phi_{\text{S-NR}}^p$. \square

Theorem 5.1. Let F be a uniform P -function and let $\Psi \in \{\Psi_{\text{NR}}^p, \Psi_{\text{S1-NR}}^p, \Psi_{\text{S2-NR}}^p\}$.

- (a) The level sets $L(\Psi, \gamma) := \{x \in \mathbb{R}^n \mid \Psi(x) \leq \gamma\}$ of Ψ are bounded for any $\gamma \geq 0$. Consequently, the trajectory $x(t)$ through any initial condition $x^0 \in \mathbb{R}^n$ is defined for all $t \geq 0$.
- (b) The trajectory $x(t)$ of (9) through any $x^0 \in \mathbb{R}^n$ converges to an equilibrium point.

Proof. (a) Suppose otherwise. Then, there exists a sequence $\{x^k\}_{k=1}^\infty \subseteq L(\Psi, \gamma)$ such that $\|x^k\| \rightarrow \infty$ as $k \rightarrow \infty$. A similar argument as in [10] shows that there exists an index i such that $|x_i^k| \rightarrow \infty$ and $|F_i(x^k)| \rightarrow \infty$ as $k \rightarrow \infty$. By Lemma 5.1, we have $|\phi(x_i^k, F_i(x^k))| \rightarrow \infty$, where $\phi \in \{\phi_{\text{NR}}^p, \phi_{\text{S-NR}}^p, \psi_{\text{S-NR}}^p\}$. But, this is impossible since $\Psi(x^k) \leq \gamma$ for all k . Thus, the level set $L(\Psi, \gamma)$ is bounded. The remaining part of the theorem can be proved similar to Proposition 4.2(b) in [2].

(b) From part(a), the level sets of Ψ are compact and so by LaSalle's Invariance Principle [22], we reach the desired conclusion. \square

Theorem 5.2. Suppose x^* is an isolated equilibrium point of (9). Then, x^* is asymptotically stable provided that either

- (i) $\Psi = \Psi_{\text{NR}}^p$ and $(\nabla F - I)$ is a P -matrix; or
- (ii) $\Psi \in \{\Psi_{\text{S1-NR}}^p, \Psi_{\text{S2-NR}}^p\}$, F is a nonnegative P_0 -function, and the equilibrium point is nonnegative.

Proof. Let x^* be an isolated equilibrium point of (9). Then, it has a neighborhood O such that

$$\nabla \Psi(x^*) = 0 \quad \text{and} \quad \nabla \Psi(x) \neq 0 \quad \text{for all } x \in O \setminus \{x^*\}.$$

We claim that Ψ is a Lyapunov function at x^* over Ω . To proceed, we note first that $\Psi(x) \geq 0$. By Proposition 5.1(b) and (c), $\Psi(x^*) = 0$. Further, if $\Psi(x) = 0$ for some $x \in O \setminus \{x^*\}$, then x solves the NCP and by Proposition 5.1(a), it is an equilibrium point. This contradicts the isolation of x^* . Thus, $\Psi(x) > 0$ for all $x \in O \setminus \{x^*\}$. Finally, it is clear that

$$\frac{d\Psi(x(t))}{dt} = -\rho \|\nabla \Psi(x(t))\|^2 < 0$$

over the set $O \setminus \{x^*\}$. Then, applying Lemma 3.5 yields that x^* is asymptotically stable. \square

We look now at the exponential stability of the neural network.

Theorem 5.3. Consider the neural network (9) with $\Psi \in \{\Psi_{\text{NR}}^p, \Psi_{\text{S1-NR}}^p, \Psi_{\text{S2-NR}}^p\}$. If $\nabla \Phi(x^*)$ is nonsingular for some isolated equilibrium point x^* , then x^* solves the NCP and x^* is exponentially stable.

Proof. Let x^* be an equilibrium point such that $\nabla \Phi(x^*)$ is nonsingular. Note that $\nabla \Psi(x^*) = \nabla \Phi(x^*)\Phi(x^*)$, and so $\nabla \Psi(x^*) = 0$ implies that $\Phi(x^*) = 0$. This proves the first claim of the theorem. Further, using Ψ as a Lyapunov function as in the preceding theorem, x^* is asymptotically stable.

Note that since Φ is differentiable at x^* , we have

$$\Phi(x) = \nabla \Phi(x)^T(x - x^*) + o(\|x - x^*\|) \quad \text{as } x \rightarrow x^* \quad (16)$$

By Lemma 3.6, there exists $\delta > 0$ and a constant C such that $\nabla \Phi(x)$ is nonsingular for all x with $\|x - x^*\| < \delta$, and $\|\nabla \Phi(x)^{-1}\| \leq C$. Then, it gives

$$\kappa \|y\|^2 \leq \|\nabla \Phi(x)y\|^2 \quad (17)$$

for any x in the δ -neighborhood (call it N_δ) and any $y \in \mathbb{R}^n$, where $\kappa = 1/C^2$.

Let $\varepsilon < 2\rho\kappa$. Since x^* is asymptotically stable, we may choose δ small enough so that $o(\|x - x^*\|^2) < \varepsilon\|x - x^*\|^2$ and $x(t) \rightarrow x^*$ as $t \rightarrow \infty$ for any initial condition $x(0) \in N_\delta$. Now, define $g: [0, \infty] \rightarrow \mathbb{R}$ by

$$g(t) := \|x(t) - x^*\|^2$$

where $x(t)$ is the unique solution through $x(0) \in N_\delta$. Using Eqs. (16) and (17), we obtain

$$\begin{aligned} \frac{dg(t)}{dt} &= 2(x(t) - x^*)^T \frac{dx(t)}{dt} \\ &= -2\rho(x(t) - x^*)^T \nabla \Psi(x(t)) \\ &= -2\rho(x(t) - x^*)^T \nabla \Phi(x(t))\Phi(x(t)) \\ &= -2\rho(x(t) - x^*)^T \nabla \Phi(x(t))\nabla \Phi(x)^T(x(t) - x^*) \\ &\quad + o(\|x(t) - x^*\|^2) \\ &\leq (-2\rho\kappa + \varepsilon)\|x(t) - x^*\|^2 \\ &= (-2\rho\kappa + \varepsilon)g(t). \end{aligned}$$

Then, it follows that $g(t) \leq e^{(-2\rho\kappa + \varepsilon)t}g(0)$, which says

$$\|x(t) - x^*\| \leq e^{(-\rho\kappa + \varepsilon/2)t}\|x(0) - x^*\|,$$

where $-\rho\kappa + \varepsilon/2 < 0$. This proves that x^* is exponentially stable. \square

6. Simulation results

In this section, we look at some nonlinear complementarity problems and test them using the neural network (9) with $\Psi \in \{\Psi_{\text{NR}}^p, \Psi_{\text{S1-NR}}^p, \Psi_{\text{S2-NR}}^p\}$. We also compare the rate of convergence of each network for different values of p . Further, we compare the numerical performance of these networks with the neural network based on the Fischer-Burmeister (FB) function [23] given by (10) and the neural network based on the generalized Fischer-Burmeister function [2] given by (11).

In the following simulations, we use the Matlab ordinary differential equation solver *ode23s*. Recall that ρ is a time-scaling parameter. In particular, if we wish to achieve faster convergence, a higher value of ρ can be used. In our simulations, the values of ρ used are 10^3 , 10^6 or 10^9 , as indicated in the figures. The stopping criterion in simulating the trajectories is $\|\nabla \Psi(x(t))\| \leq 10^{-5}$.

Example 6.1. [21, Kojima-Shindo] Consider the NCP, where $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ is given by

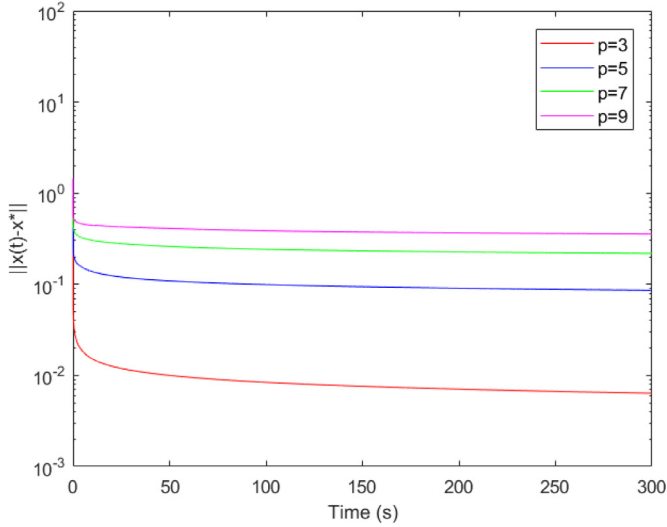


Fig. 2. Convergence behavior of the error $\|x(t) - x^*\|$ in Example 5.1 using the neural network with ϕ_{NR}^p for different values of p , where $x^0 = (2, 0.5, 0.5, 1.5)^T$ and $\rho = 10^6$.

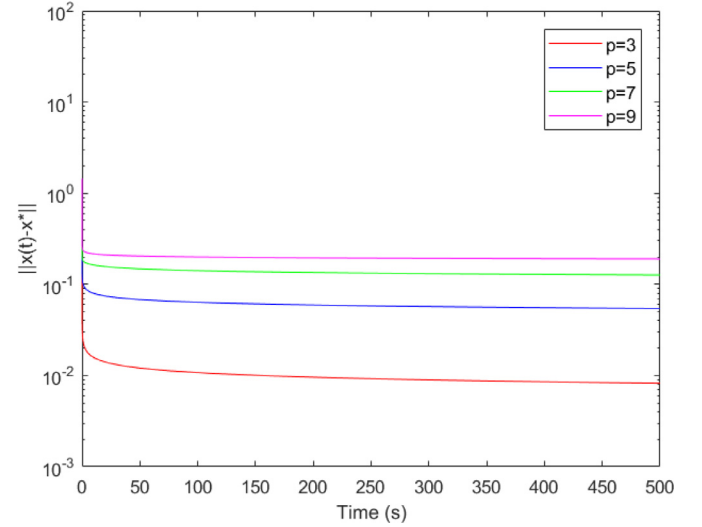


Fig. 4. Convergence behavior of the error $\|x(t) - x^*\|$ in Example 5.1 using the neural network with ψ_{S-NR}^p for different values of p , where $x^0 = (2, 0.5, 0.5, 1.5)^T$ and $\rho = 10^9$.

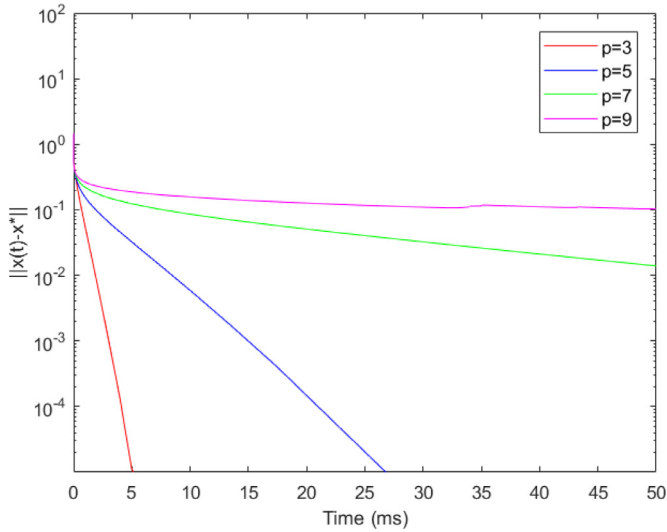


Fig. 3. Convergence behavior of the error $\|x(t) - x^*\|$ in Example 5.1 using the neural network with ϕ_{NR}^p for different values of p , where $x^0 = (2, 0.5, 0.5, 1.5)^T$ and $\rho = 10^3$.

$$F(x) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 3x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 3x_4 - 1 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix}.$$

This is a non-degenerate NCP and the solution is $x^* = (\sqrt{6}/2, 0, 0, 1/2)$.

We simulate the network (9) with different $\Psi \in \{\Psi_{NR}^p, \Psi_{S1-NR}^p, \Psi_{S2-NR}^p\}$ for various values of p to see the influence of p on convergence of trajectories to the NCP solution. From Figs. 2–4, we see that a smaller value of p yields a faster convergence when the initial condition is $x^0 = (2, 0.5, 0.5, 1.5)^T$. Fig. 5 depicts the comparison of the different NCP-functions with $p = 3$, together with the FB and generalized FB functions. Among these five classes of NCP-functions, we see that the neural network based on ϕ_{S-NR}^p has the best numerical performance. In Fig. 6, we simulate the neural network based on ϕ_{S-NR}^p using 6 random initial points, and the trajectories converges to x^* at around $t = 5.5$ ms.

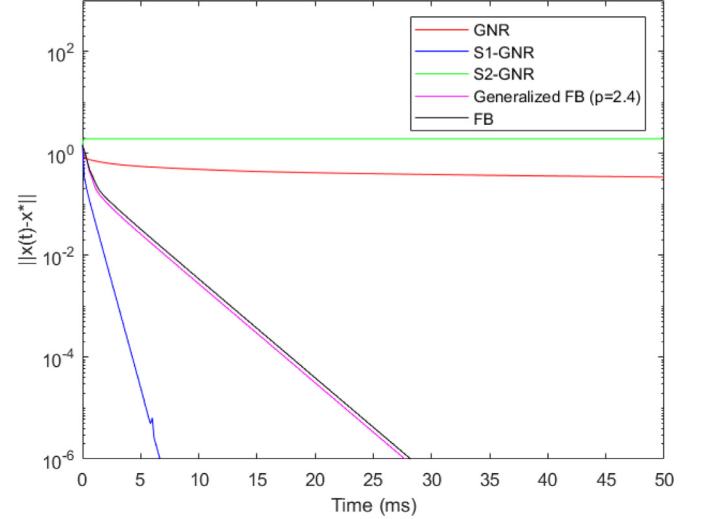


Fig. 5. Comparison of convergence rates of $\|x(t) - x^*\|$ in Example 5.1 using the neural network with different NCP functions, where $x^0 = (2, 0.5, 0.5, 1.5)^T$ and $\rho = 10^3$.

One can also observe from Fig. 6 that the convergence of $x_2(t)$ and $x_3(t)$ is very fast. We note that $\nabla \Phi_{S1-NR}^p(x^*)$ is non-singular, which leads to the exponential stability of x^* by Theorem 5.3. This particular problem was also simulated using neural networks based on the FB function [23] and the generalized FB function with $p = 2.4$ [2], with convergence time at around $t = 22$ ms and $t = 18$ ms, respectively. It can be verified that the neural networks based on ϕ_{NR}^p and ψ_{S-NR}^p also converge to the solution, but the convergence rate is very slow. This can be observed in Figs. 2 and 4, in which the convergence will even be worse if a smaller value of ρ is chosen.

Example 6.2. [21, Kojima-Shindo] Consider the NCP, where $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ is given by

$$F(x) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix}.$$

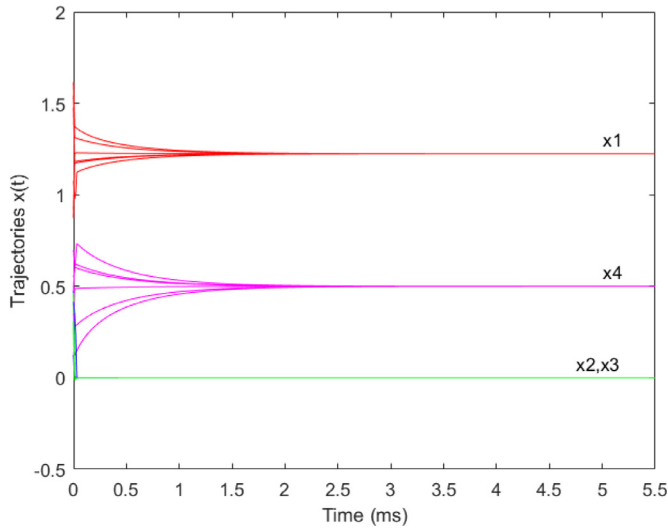


Fig. 6. Transient behavior of $x(t)$ Example 5.1 of the neural network with ϕ_{S-NR}^p ($p = 3$) with 6 random initial points, where $\rho = 10^3$.

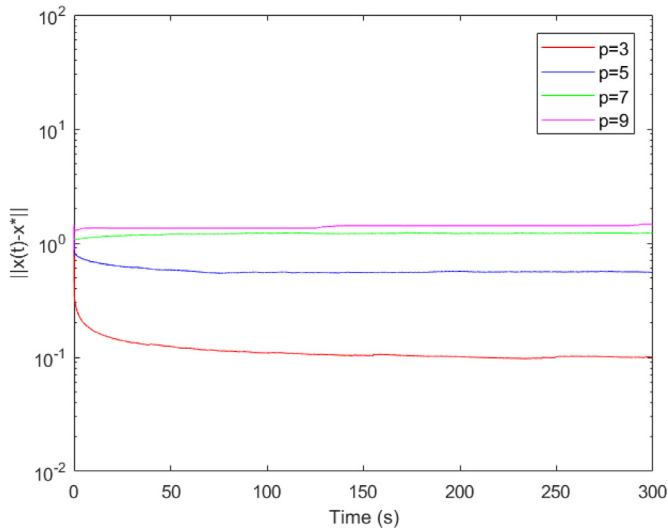


Fig. 7. Convergence behavior of the error $\|x(t) - x^*\|$ in Example 5.2 using the neural network with ϕ_{NR}^p for different values of p , where $x^0 = (0.5, 0.5, 3.5, 0.5)^T$ and $\rho = 10^6$.

The corresponding NCP problem has a degenerate solution $x^* = (\sqrt{6}/2, 0, 0, 1/2)$ and a non-degenerate solution $x^* = (1, 0, 3, 0)$.

Figs. 7 and 8 show the convergence behavior of the error $\|x(t) - x^*\|$ for different values of p using the neural network (9) with ϕ_{NR}^p and ψ_{S-NR}^p , where x^* is the non-degenerate solution. As one may observe, a smaller value of p leads to a faster convergence of the error. Unfortunately, it is computationally expensive to simulate (9) with $p \geq 5$. Nevertheless, we provide a comparison of the performance of the three discrete-type NCP functions (with $p = 3$) and the FB functions as shown in Fig. 9. We can see the same performance as in Example 5.1, with the neural network based on ϕ_{S-NR}^p as the one with the fastest convergence, whereas the one based on ϕ_{NR}^p is also the slowest. From Fig. 10, we see that the solution converges to the NCP solution at around $t = 12$ ms. From this figure, we can also see that the solutions $x_2(t)$ and $x_4(t)$ converges to 0 at a very fast rate. It should be noted that $\nabla \Phi_{S1-NR}^p(x^*)$ is nonsingular, and therefore x^* is exponentially stable by Theorem 5.3. On the other hand, the neural network based on

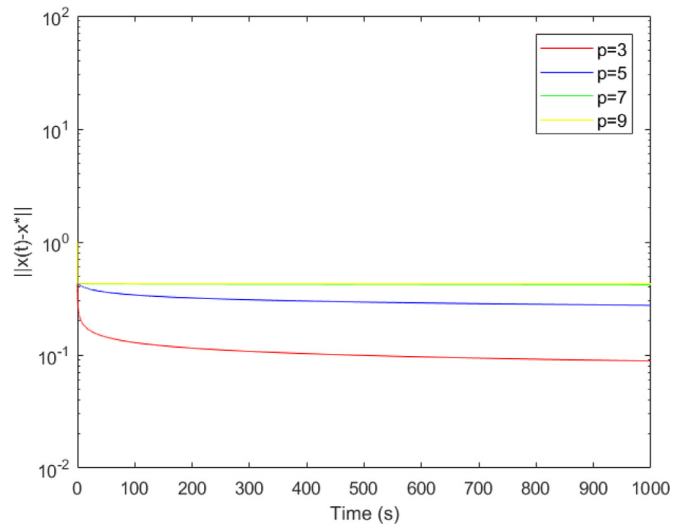


Fig. 8. Convergence behavior of the error $\|x(t) - x^*\|$ in Example 5.2 using the neural network with ψ_{S-NR}^p for different values of p , where $x^0 = (0.5, 0.5, 3.5, 0.5)^T$ and $\rho = 10^6$.

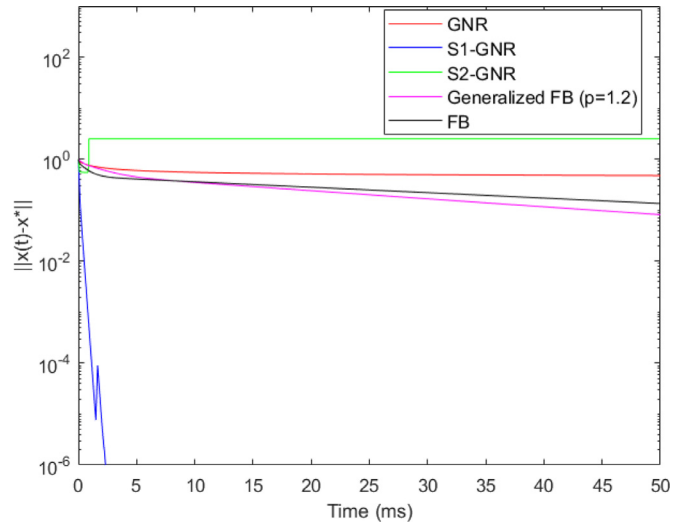


Fig. 9. Comparison of convergence rates of $\|x(t) - x^*\|$ in Example 5.2 using the neural network with different NCP functions, where $x^0 = (0.5, 0.5, 3.5, 0.5)^T$ and $\rho = 10^3$.

the generalized FB function (with $p = 1.2$) converged to the equilibrium point at around $t = 800$ ms.

Example 6.3. Consider the NCP, where $F: \mathbb{R}^5 \rightarrow \mathbb{R}^5$ is given by

$$F(x) = \begin{pmatrix} x_1 + x_2 x_3 x_4 x_5 / 50 + 5 \\ x_2 + x_1 x_3 x_4 x_5 / 50 - 3 \\ x_3 + x_1 x_2 x_4 x_5 / 50 - 1 \\ x_4 + x_1 x_2 x_3 x_5 / 50 + 1/2 \\ x_5 + x_1 x_2 x_3 x_4 / 50 + 1 \end{pmatrix}.$$

The above function is a slight modification of [42, Example 2]. The NCP has only one solution $x^* = (0, 3, 1, 0, 0)$. It can be verified that similar to the preceding examples, a smaller value of p yields a faster convergence for all the three discrete-type classes of NCP-functions. Among all the five neural networks, the neural network with the fastest convergence is the one based on the FB function, followed by the generalized FB function (with $p = 1.8$) and by ϕ_{NR}^p

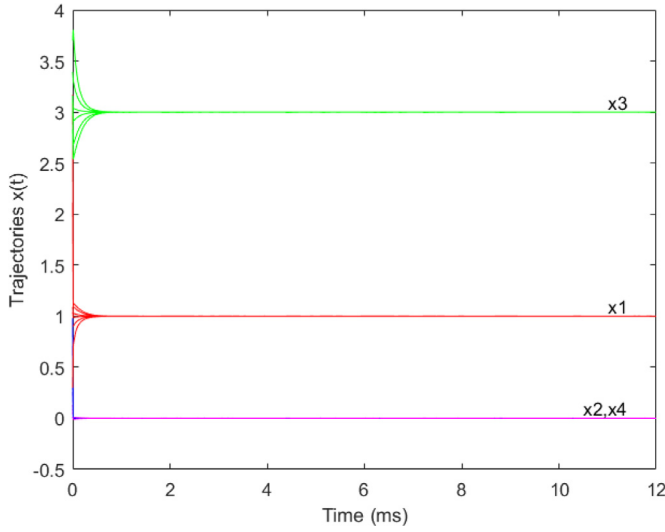


Fig. 10. Transient behavior of $x(t)$ in Example 5.2 of the neural network with ϕ_{S-NR}^p ($p = 3$) with 6 random initial points, where $\rho = 10^3$.

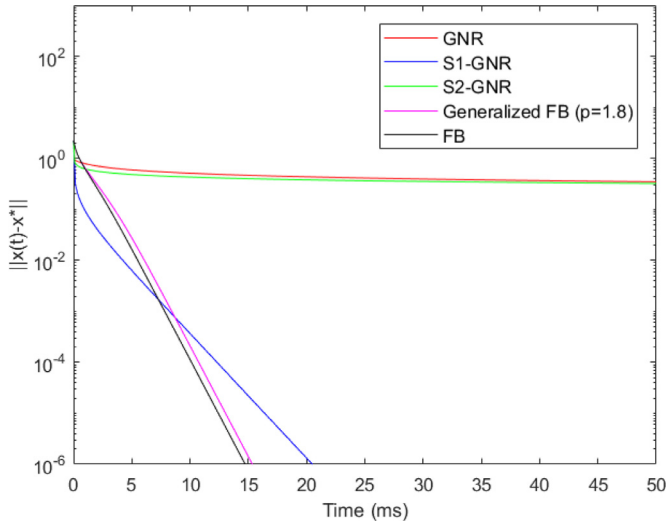


Fig. 11. Comparison of convergence rates of $\|x(t) - x^*\|$ in Example 5.3 using the neural network with different NCP functions, where $x^0 = (0.5, 1, 1.5, 0.5, 0.5)^T$ and $\rho = 10^3$.

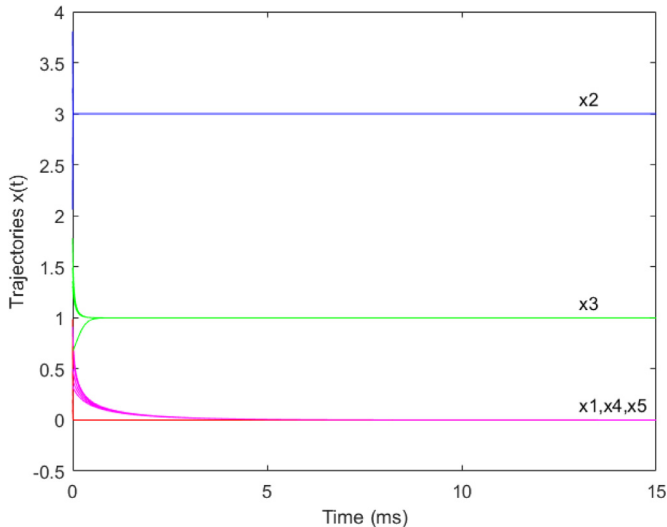


Fig. 12. Transient behavior of $x(t)$ in Example 5.3 of the neural network with ϕ_{S-NR}^p ($p = 3$) with 6 random initial points, where $\rho = 10^3$.

(with $p = 3$); see Fig. 11. Nevertheless, the neural network based on ϕ_{NR}^p is also efficient in solving the NCP as shown in Fig. 12.

To close this section, we summarize our findings based from the preceding simulations. We also discuss some advantages of the neural networks based on the three discrete-type NCP-functions over other existing models.

- A smaller value of p yields faster convergence of trajectories for the three neural networks based on ϕ_{NR}^p , ϕ_{S-NR}^p and ψ_{S-NR}^p .
- Among these three classes of NCP-functions, simulations show that the neural network based on ψ_{S-NR}^p has the worst performance in all the examples. Note that this function possesses both differentiability and symmetry. Meanwhile, the first symmetrization ϕ_{S-NR}^p which is symmetric but not differentiable on the line $a = b$ ($a, b \neq 0$) has the best numerical performance.
- Although we only consider nondegenerate NCPs in the above examples, the three neural networks based on the discrete-type classes of NCP-functions can also be used in degenerate cases. It can be verified, for instance, that the trajectories in Example 5.2 can converge to the degenerate solution $x^* = (\sqrt{6}/2, 0, 0, 1/2)$. However, the convergence is very slow. As remarked in the Introduction, this phenomenon is expected since these three classes of NCP-functions are differentiable at all points in the feasible region $x \geq 0, F(x) \geq 0$.
- Interestingly, the neural network based on ϕ_{S-NR}^p has the potential to outperform the neural networks based on the FB function ϕ_{FB}^p and generalized FB function ϕ_{FB}^p , as illustrated in Examples 5.1 and 5.2. The difference in the numerical performance, in fact, is very large. In Example 5.1, the convergence time is only $t = 5.5$ ms when ϕ_{S-NR}^p is used. There is a significant margin when FB or generalized FB is used, where the convergence time is $t = 22$ ms and $t = 18$ ms, respectively.
- In particular, in Example 5.2, the convergence of trajectories for the network based on ϕ_{S-NR}^p is achieved at around $t = 12$ ms only. This is a significant improvement over the convergence time of $t = 800$ ms when the neural network based on ϕ_{FB}^p was used. Looking at Fig. 9, the performance of ϕ_{FB}^p is even slower.
- In Example 5.3, despite the better performance of the FB and generalized FB function over ϕ_{S-NR}^p , the difference is only marginal.
- One can check that the functions considered above are not monotone. This is a requirement for Lyapunov stability and exponential stability for the networks used in [7,42,43]. Nonetheless, it is not needed for our experiments.

7. Conclusions and future directions

In this paper, we established the properties of the induced merit function of three classes of discrete-type NCP-functions ϕ_{NR}^p , ϕ_{S-NR}^p and ψ_{S-NR}^p . This includes the characterization of stationary points and level sets of the corresponding merit functions. We have also described the growth behavior of these three families of NCP-functions. Moreover, we proved the Lyapunov and asymptotic stability properties of the proposed steepest descent-based neural networks. Finally, numerical simulations indicate that for all three families of neural networks, a smaller value of p always yields faster convergence to the NCP solution. We leave it for future research to determine under what conditions on F does this behavior occur.

Numerical reports suggest that the neural network based on ϕ_{S-NR}^p is capable of outperforming the neural network based on the extensively used FB function, as well as the network based on the generalized FB function. However, it must be noted that as in other neural networks, results may be dependent on the choice of initial conditions. Nevertheless, it would be an interesting future

research endeavor to find out when ϕ_{S-NR}^p can be best used and whether it is capable of outperforming other NCP-functions when other solution methods are employed. Exploring this could possibly provide an alternative NCP-function to use when designing algorithms. Furthermore, the effect of varying the values of p when other approaches are used can also be considered. That is, one could find out if smaller values of p will also yield faster convergence when other solution methods are used. Theoretical proof for performance-dependence on p can also be investigated. Exploring deeper the properties of ϕ_{S-NR}^p deserves particular attention because of its numerical performance, despite its complexity.

A lot of future research directions can be taken on from here since many algorithms in optimization rely on NCP-functions. In the case of nonlinear complementarity problems, the three classes of NCP-functions used in this paper can be exploited to design other solution methods such as merit function approach, nonsmooth Newton method, smoothing methods, and regularization approach, among others. This paper serves as a starting point in designing such approaches, since we have established herein the properties of level sets and stationary points of the induced merit functions. Indeed, these properties are fundamental in constructing other NCP-based methods. Of course, the NCP-functions studied in this paper can be adopted to formulate neural network approaches to other complementarity problems, as well as variational inequalities and linear and nonlinear programming problems. For instance, delayed neural networks for optimization is an interesting direction [24,34,35,39,40,51–55]. Stochastic nonlinear complementarity problems [48] can also be explored. In [48], ϕ_{NR} , ϕ_{FB} and the penalized FB functions are used to solve stochastic NCPs. A numerical comparison of the performance of the newly discovered NCP-functions considered in this paper with old NCP-functions used in [48] can be done. Stochastic neural networks for these types of NCPs can also be investigated [51–55].

Declarations of interest

None.

References

- [1] Y.L. Chang, J.S. Chen, C.Y. Yang, Symmetrization of generalized natural residual function for NCP, *Operat. Res. Lett.* 43 (2015) 354–358.
- [2] J.S. Chen, C.H. Ko, S.H. Pan, A neural network based on generalized Fischer-Burmeister function for nonlinear complementarity problems, *Inf. Sci.* 180 (2010) 697–711.
- [3] J.S. Chen, C.H. Ko, X.R. Wu, What is the generalization of natural residual function for NCP, *Pacif. J. Optim.* 12 (2016) 19–27.
- [4] C. Chen, O.L. Mangasarian, A class of smoothing functions for nonlinear and mixed complementarity problems, *Comput. Optim. Appl.* 5 (1996) 97–138.
- [5] J.S. Chen, S.H. Pan, A family of NCP functions and a descent method for the nonlinear complementarity problem, *Comput. Optim. Appl.* 40 (2008) 389–404.
- [6] R.W. Cottle, J.S. Pang, R.E. Stone, *The Linear Complementarity Problem*, Academic Press, New York, 1992.
- [7] C. Dang, Y. Leung, X. Gao, K. Chen, Neural networks for nonlinear and mixed complementarity problems and their applications, *Neural Netw.* 17 (2004) 271–283.
- [8] M.C. Ferris, O.L. Mangasarian, J.S. Pang (Eds.), *Complementarity: Applications, Algorithms and Extensions*, Kluwer Academic Publishers, Dordrecht, 2001.
- [9] F. Facchinei, J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer-Verlag, New York, 2003. Volumes I and II.
- [10] F. Facchinei, J. Soares, A new merit function for nonlinear complementarity problems and a related algorithm, *SIAM J. Optim.* 7 (1997) 225–247.
- [11] C. Geiger, C. Kanzow, On the resolution of monotone complementarity problems, *Comput. Optim. Appl.* 5 (1996) 155–173.
- [12] P.T. Harker, J.S. Pang, Finite dimensional variational inequality and nonlinear complementarity problem: a survey of theory, algorithms and applications, *Math. Program.* 48 (1990) 161–220.
- [13] J.J. Hopfield, D.W. Tank, Neural computation of decision in optimization problems, *Biol. Cybern.* 52 (1985) 141–152.
- [14] X. Hu, J. Wang, Solving pseudomonotone variational inequalities and pseudoconvex optimization problems using the projection neural network, *IEEE Trans. Neural Netw.* 17 (2006) 1487–1499.
- [15] X. Hu, J. Wang, A recurrent neural network for solving a class of general variational inequalities, *IEEE Trans. Syst. Man Cybern. B* 37 (2007) 528–539.
- [16] C.H. Huang, K.J. Weng, J.S. Chen, H.W. Chu, M.Y. Li, On four discrete-type families of NCP functions, *J. Nonlinear Convex Anal.* 20 (2) (2019) 215–228.
- [17] Z.H. Huang, Locating a maximally complementary solutions of the monotone NCP by using non-interior-point smoothing algorithm, *Math. Methods Operat. Res.* 61 (2005) 41–55.
- [18] C. Kanzow, H. Kleinmichel, A class of newton-type methods for equality and inequality constrained optimization, *Optim. Methods Softw.* 5 (1995) 173–198.
- [19] C. Kanzow, Nonlinear complementarity as unconstrained optimization, *J. Optim. Theory Appl.* 88 (1996) 139–155.
- [20] M.P. Kennedy, L.O. Chua, Neural network for nonlinear programming, *IEEE Trans. Circuits Syst.* 35 (1988) 554–562.
- [21] M. Kojima, S. Shindo, Extensions of newton and quasi-newton methods to systems of PC^1 equations, *J. Operat. Res. Soc. Jpn.* 29 (1986) 352–374.
- [22] J.P. La Salle, Stability theory for ordinary differential equations, *J. Differ. Equ.* 4 (1968) 57–65.
- [23] L.Z. Liao, H. Qi, L. Qi, Solving nonlinear complementarity problems with neural networks: a reformulation method approach, *J. Comput. Appl. Math.* 131 (2001) 342–359.
- [24] Q. Liu, J. Cao, Y. Xia, A delayed neural network for solving linear projection equations and its analysis, *IEEE Trans. Neural Netw.* 16 (4) (2005) 834–843.
- [25] L. Liu, Y.J. Liu, S. Tong, Neural networks-based adaptive finite-time fault-tolerant control for a class of strict-feedback switched nonlinear systems, *IEEE Trans. Cybern.* (2018) 1–10.
- [26] O.L. Mangasarian, M.V. Solodov, Nonlinear complementarity as unconstrained and constrained minimization, *Math. Program.* 62 (1993) 277–297.
- [27] R.K. Miller, A.N. Michel, *Ordinary Differential Equations*, Academic Press, 1982.
- [28] S.-K. Oh, W. Pedrycz, S.-B. Roh, Genetically optimized fuzzy polynomial neural networks with fuzzy set-based polynomial neurons, *Inf. Sci.* 176 (2006) 3490–3519.
- [29] F.A. Potra, S.J. Wright, Interior-point methods, *J. Comput. Appl. Math.* 124 (2000) 255–281.
- [30] F.A. Potra, Y. Ye, Interior-point methods for nonlinear complementarity problems, *J. Optim. Theory Appl.* 88 (3) (1996) 617–642.
- [31] H.D. Qi, L.Z. Liao, A smoothing newton method for general nonlinear complementarity problems, *Comput. Optim. Appl.* 17 (2000) 231–253.
- [32] L. Qi, J. Sun, A nonsmooth version of Newton's method, *Math. Program.* 58 (1993) 353–368.
- [33] R.T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.* 14 (1976) 877–898.
- [34] K. Shi, J. Wang, Y. Tang, S. Zhong, Reliable asynchronous sampled-data filtering of T-S fuzzy uncertain delayed neural networks with stochastic switched topologies, *Fuzzy Sets Syst.* (2018), doi:10.1016/j.fss.2018.11.017.
- [35] K. Shi, J. Wang, S. Zhong, X. Zhang, Y. Liu, J. Cheng, New reliable nonuniform sampling control for uncertain chaotic neural networks under Markov switching topologies, *Appl. Math. Comput.* 347 (2019) 169–193.
- [36] A. Shortt, J. Keating, L. Monlinier, C. Pannell, Optical implementation of the KAK neural network, *Inf. Sci.* 171 (2005) 273–287.
- [37] D. Sun, A regularization newton method for solving nonlinear complementarity problems, *Appl. Math. Optim.* 40 (1999) 315–339.
- [38] D.W. Tank, J.J. Hopfield, Simple neural optimization networks: an A/D converter, signal decision circuit, and a linear programming circuit, *IEEE Trans. Circuits Syst.* 33 (1986) 533–541.
- [39] Z. Wang, S. Ding, Q. Shan, H. Zhang, Stability of recurrent neural networks with time-varying delay via flexible terminal method, *IEEE Trans. Neural Netw. Learn. Syst.* 28 (10) (2017) 2456–2463.
- [40] J. Wang, K. Shi, Q. Huang, S. Zhong, D. Zhang, Stochastic switched sampled-data control for synchronization of delayed chaotic neural networks with packet dropout, *Appl. Math. Comput.* 335 (2018) 211–230.
- [41] S. Wiggins, *Introduction to Applied and Nonlinear Dynamical Systems and Chaos*, Springer-Verlag, New York, 2003. Inc.
- [42] Y. Xia, H. Leung, J. Wang, A projection neural network and its application to constrained optimization problems, *IEEE Trans. Circuits Syst. I* 49 (2002) 447–458.
- [43] Y. Xia, H. Leung, J. Wang, A general projection neural network for solving monotone variational inequalities and related optimization problems, *IEEE Trans. Neural Netw.* 15 (2004) 318–328.
- [44] Y. Xia, H. Leung, J. Wang, A recurrent neural network for solving nonlinear convex programs subject to linear constraints, *IEEE Trans. Neural Netw.* 16 (2005) 379–386.
- [45] N. Yamashita, M. Fukushima, Modified newton methods for solving a semismooth reformulation of monotone complementarity problems, *Math. Program.* 76 (1997) 469–491.
- [46] X. Yao, Z. Wang, H. Zhang, Identification method for a class of periodic discrete-time dynamic nonlinear systems based on sinusoidal ESN, *Neurocomputing* 275 (2017) 1511–1521.
- [47] M. Yashtini, A. Malek, Solving complementarity and variational inequalities problems using neural networks, *Appl. Math. Comput.* 190 (2007) 216–230.
- [48] C. Zhang, X. Chen, Stochastic nonlinear complementarity problem and applications to traffic equilibrium under uncertainty, *J. Optim. Theory Appl.* 137 (2008) 277–295.
- [49] S.H. Zak, V. Upatizing, S. Hui, Solving linear programming problems with neural networks: a comparative study, *IEEE Trans. Neural Netw.* 6 (1995) 94–104.
- [50] G. Zhang, A neural network ensemble method with jittered training data for time series forecasting, *Inf. Sci.* 177 (2007) 5329–5340.

- [51] Q. Zhu, J. Cao, Stability analysis of Markovian jump stochastic BAM neural networks with impulse control and mixed time delays, *IEEE Trans. Neural Netw. Learn. Syst.* 23 (3) (2012a) 467–479.
- [52] Q. Zhu, J. Cao, pth moment exponential synchronization for stochastic delayed Cohen-Grossberg neural networks with Markovian switching, *Nonlinear Dyn.* 67 (1) (2012b) 829–845.
- [53] Q. Zhu, J. Cao, R. Rakkiyappan, Exponential input-to-state stability of stochastic Cohen-Grossberg neural networks with mixed delays, *Nonlinear Dyn.* 79 (2) (2015) 1085–1098.
- [54] Q. Zhu, X. Li, Exponential and almost sure exponential stability of stochastic fuzzy delayed Cohen-Grossberg neural networks, *Fuzzy Sets Syst.* 203 (2012) 79–94.
- [55] Q. Zhu, R. Rakkiyappan, A. Chandrasekar, Stochastic stability of Markovian jump BAM neural networks with leakage delays and impulse control, *Neurocomputing* 136 (2014) 136–151.



Jein-Shan Chen is a Professor at Mathematics Department of National Taiwan Normal University. His research area is mainly in continuous optimization with side interests in nonsmooth analysis and operations research. His publications include around 110 journal papers and one monograph, titled *OC Functions and Their Applications* Springer Optimization and Its Applications 143, Springer, 2019.



Jan Harold Alcantara received the MS degree in Mathematics from De La Salle University. He is currently pursuing the Ph.D. degree with the Department of Mathematics, National Taiwan Normal University. His research interests are continuous optimization and neural network approach for optimization.