J. Aust. Math. Soc. **99** (2015), 364–379
doi:10.1017/S1446788715000300

A NEURAL NETWORK FOR THE GENERALIZED NONLINEAR COMPLEMENTARITY PROBLEM OVER A POLYHEDRAL CONE

YIFEN KE and CHANGFENG MA✉

(Received 26 November 2014; accepted 8 July 2015)

Communicated by V. Jeyakumar

Abstract

In this paper, we consider a neural network model for solving the generalized nonlinear complementarity problem (denoted by GNCP) over a polyhedral cone. The neural network is derived from an equivalent unconstrained minimization reformulation of the GNCP, which is based on the penalized Fischer–Burmeister function $\phi_\mu(a, b) = \mu\phi_{FB}(a, b) + (1 - \mu)a_+b_+$. We establish the existence and the convergence of the trajectory of the neural network, and study its Lyapunov stability, asymptotic stability and exponential stability. It is found that a larger μ leads to a better convergence rate of the trajectory. Simulation results are also reported.

2010 *Mathematics subject classification*: primary 90C33; secondary 65K10.

Keywords and phrases: generalized nonlinear complementarity problem, polyhedral cone, neural network, convergence analysis, simulation results.

1. Introduction

The generalized nonlinear complementarity problem, denoted by GNCP (F, G, K), is to find a vector $x^* \in R^n$ such that

$$F(x^*) \in K, \quad G(x^*) \in K^*, \quad F(x^*)^{\mathrm{T}} G(x^*) = 0,$$

where F and G are continuous functions from R^n to R^m, K is a nonempty closed convex cone in R^m, and K^* is the polar cone of K.

In this paper, we consider the GNCP (F, G, K) for the case where F and G are both continuously differentiable from R^n to R^m and K is a polyhedral cone in R^n: that is, there exist $A \in R^{s\times m}$ and $B \in R^{t\times m}$ such that

$$K = \{v \in R^m \mid Av \geq 0, Bv = 0\}.$$

The work is supported by the National Natural Science Foundation of China (11071041, 11201074), Fujian Natural Science Foundation (2013J01006, 2015J01578) and R&D of Key Instruments and Technologies for Deep Resources Prospecting (the National R&D Projects for Key Scientific Instruments) under grant number ZDYZ2012-1-02-04.

It is easy to verify that its polar cone K^* has the representation

$$K^* = \{u \in R^m \mid u = A^{\mathrm{T}}\lambda + B^{\mathrm{T}}\omega, \lambda \geq 0, \lambda \in R^s, \omega \in R^t\}.$$

Obviously, if A is an identity matrix, $B = 0$ and $G(x) = x$, then the GNCP reduces to the classical nonlinear complementarity problem.

To solve the GNCP, one usually reformulates it as a minimization problem over a simple set or as an unconstrained optimization problem. Here we reformulate the GNCP as a system of equations via the penalized Fischer–Burmeister function (see [2]), which is defined as

$$\phi_\mu(a, b) = \mu\phi_{FB}(a, b) + (1 - \mu)a_+b_+,$$

where $\mu \in (0, 1)$ is an arbitrary but fixed parameter, $a, b \in R$, $\phi_{FB}(a, b) = a + b - \sqrt{a^2 + b^2}$, $a_+ = \max\{0, a\}$ and $b_+ = \max\{0, b\}$. The basic property of the penalized Fischer–Burmeister function is that

$$\phi_\mu(a, b) = 0 \Longleftrightarrow a \geq 0, b \geq 0, ab = 0.$$

For arbitrary vectors $a = (a_1, a_2, \ldots, a_n)^{\mathrm{T}}$ and $b = (b_1, b_2, \ldots, b_n)^{\mathrm{T}} \in R^n$, we define a vector-valued function

$$\Phi_\mu(a, b) = \begin{pmatrix} \phi_\mu(a_1, b_1) \\ \phi_\mu(a_2, b_2) \\ \vdots \\ \phi_\mu(a_n, b_n) \end{pmatrix}.$$

Obviously,

$$\begin{aligned} \Phi_\mu(a, b) = 0 &\Longleftrightarrow \phi_\mu(a_i, b_i) = 0, \quad i = 1, 2, \ldots, n \\ &\Longleftrightarrow a_i \geq 0,\ b_i \geq 0,\ a_ib_i = 0, \quad i = 1, 2, \ldots, n \\ &\Longleftrightarrow a \geq 0,\ b \geq 0,\ ab = 0. \end{aligned}$$

Now, we give some equivalent statements relative to the solution of the GNCP.

LEMMA 1.1 [14].

$$\begin{aligned} &\textit{The vector } x^* \textit{ is a solution of GNCP } (F, G, K) \\ &\Longleftrightarrow \begin{cases} F(x^*) \in K = \{v \in R^m \mid Av \geq 0, Bv = 0\}, \\ G(x^*) \in K^* = \{u \in R^m \mid u = A^{\mathrm{T}}\lambda + B^{\mathrm{T}}\omega, \lambda \geq 0, \lambda \in R^s, \omega \in R^t\}, \\ F(x^*)^{\mathrm{T}}G(x^*) = 0, \end{cases} \\ &\Longleftrightarrow \textit{there exist } \lambda^* \in R^s, \omega^* \in R^t \textit{ such that } \begin{cases} AF(x^*) \geq 0, \\ BF(x^*) = 0, \\ G(x^*) = A^{\mathrm{T}}\lambda^* + B^{\mathrm{T}}\omega^*, \\ \lambda^* \geq 0, \\ F(x^*)^{\mathrm{T}}G(x^*) = 0, \end{cases} \end{aligned}$$

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

$$\Longleftrightarrow \textit{there exist } \lambda^* \in R^s, \omega^* \in R^t \textit{ such that } \begin{cases} AF(x^*) \geq 0, \\ \lambda^* \geq 0, \\ (\lambda^*)^{\mathrm{T}} AF(x^*) = 0, \\ BF(x^*) = 0, \\ G(x^*) = A^{\mathrm{T}}\lambda^* + B^{\mathrm{T}}\omega^*, \end{cases}$$

$$\Longleftrightarrow \textit{there exist } \lambda^* \in R^s, \omega^* \in R^t \textit{ such that } \begin{cases} \Phi_\mu(AF(x^*), \lambda^*) = 0, \\ BF(x^*) = 0, \\ G(x^*) = A^{\mathrm{T}}\lambda^* + B^{\mathrm{T}}\omega^*. \end{cases}$$

Moreover, let $z := (x, \lambda, \omega) \in R^{n+s+t}$ and

$$\Theta(z) := \begin{pmatrix} \Phi_\mu(AF(x), \lambda) \\ BF(x) \\ G(x) - A^{\mathrm{T}}\lambda - B^{\mathrm{T}}\omega \end{pmatrix}. \tag{1.1}$$

If we define the function $f : R^{n+s+t} \to R$ as

$$f(z) := \tfrac{1}{2}\|\Theta(z)\|^2, \tag{1.2}$$

then the GNCP can be reformulated into the smooth minimization problem

$$\min_{z:=(x,\lambda,\omega)\in R^{n+s+t}} f(z). \tag{1.3}$$

From Lemma 1.1, (1.1) and (1.2), we can obtain the following theorem.

Theorem 1.2. *The vector $x^* \in R^n$ is a solution of the* GNCP (F, G, K) *if and only if there exist $\lambda^* \in R^s, \omega^* \in R^t$ such that $f(z) = 0$.*

There are many algorithms that can be applied to the unconstrained smooth minimization problem (1.3), such as effective gradient-type methods. However, in many scientific and engineering applications, it is desirable to have a real-time solution of the GNCP. Traditional unconstrained optimization algorithms may not be suitable for real-time implementation because the computing time required for a solution depends, to a large extent, on the dimension and structure of the problems. To overcome this difficulty, we consider an Artificial Neural Network (ANN) (see [9]).

ANNs for optimization were first introduced in the 1980s by Hopfield and Tank (see [6, 13]). Generally speaking, ANNs provide an alternative and attractive method for solving optimization problems, and applications include linear programming, nonlinear programming, quadratic programming, variational inequalities, and solving linear and nonlinear complementarity problems (see [1, 3–5, 8, 10–12, 16]).

The main idea of the neural network approach for optimization is to construct a nonnegative energy function and establish a dynamic system that can be represented by an ANN. Moreover, the function defining the dynamic system is continuous only, and not necessarily differentiable. The dynamic system is usually in the form of first-order ordinary differential equations. Furthermore, it is expected that, for an initial point, the dynamic system will approach its static state (or an equilibrium point), which corresponds to the solution of the underlying optimization problem.

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

In this paper, we apply a first-order ANN for the GNCP, which is based on the steepest descent method, to (1.2): that is,

$$\frac{dz(t)}{dt} = -\rho \nabla f(z), \quad z(0) = z_0, \tag{1.4}$$

where $z := (x, \lambda, \omega) \in R^{n+s+t}$, $\rho > 0$ is a scaling factor.

Throughout this paper, R^n denotes the space of n-dimensional real column vectors. The inner product of vectors $x, y \in R^n$ is denoted by $x^{\mathrm{T}} y$ and $^{\mathrm{T}}$ denotes the transpose. Let $\|\cdot\|$ denote the 2-norm or the Euclidean norm. For any differentiable function $f : R^n \to R$, $\nabla f(x)$ means the gradient of f at x. For any differentiable mapping $F = (F_1, F_2, \ldots, F_m)^{\mathrm{T}} : R^n \to R^m$, $\nabla F(x) = [\nabla F_1(x), \nabla F_2(x), \ldots, \nabla F_m(x)] \in R^{n\times m}$ denotes the transposed Jacobian of F at x. For vector $a \in R^n$, $D_a = \mathrm{diag}(a)$ denotes the diagonal matrix in which the ith diagonal element is a_i.

2. Preliminaries

In this section, we will recall several definitions and results.

2.1. Properties of ϕ_μ and f. In this subsection, we will study the properties of ϕ_μ and f.

LEMMA 2.1 [2]. *The function $\phi_\mu : R^2 \to R$ satisfies the following properties:*

(a) $\phi_\mu(a, b) = 0 \Longleftrightarrow a \geq 0,\ b \geq 0,\ ab = 0$;

(b) *$\phi_\mu(a, b)$ is continuously differentiable on*

$$R^2 \backslash \{(a, b) \in R^2 \mid a \geq 0, b \geq 0, ab = 0\};$$

(c) *ϕ_μ is strong semismooth on R^2;*

(d) *the generalized gradient $\partial\phi_\mu(a, b)$ of ϕ_μ at a point $(a, b) \in R^2$ is equal to the set of all (v_a, v_b) such that*

$$(v_a, v_b) = \begin{cases} \mu(1-\xi, 1-\eta) & \text{if } (a, b) = (0, 0), \\ \mu\left(1 - \dfrac{a}{\|(a, b)\|}, 1 - \dfrac{b}{\|(a, b)\|}\right) + (1-\mu)(b_+\partial a_+, a_+\partial b_+) & \text{otherwise,} \end{cases}$$

where (ξ, η) is any vector satisfying $\|(\xi, \eta)\| \leq 1$ and

$$\partial c_+ = \begin{cases} 1 & \text{if } c > 0, \\ [0, 1] & \text{if } c = 0, \\ 0 & \text{otherwise.} \end{cases}$$

LEMMA 2.2 [5]. *If F and G are both continuously differentiable, then f is continuously differentiable, and its gradient at the point $z := (x, \lambda, \omega) \in R^{n+s+t}$ is given by $\nabla f(z) = V^{\mathrm{T}}\Theta(z)$, where $V \in \partial\Theta(z)$.*

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

By simple calculation, it is not difficult to show that the generalized Jacobian of Θ at any $(x, \lambda, \omega) \in R^{n+s+t}$ is

$$V = \begin{pmatrix} D_a A F'(x) & D_b & 0 \\ B F'(x) & 0 & 0 \\ G'(x) & -A^{\mathrm{T}} & -B^{\mathrm{T}} \end{pmatrix}, \tag{2.1}$$

where

$$D_a := \mathrm{diag}\{a_1(z), \ldots, a_s(z)\}, \quad D_b := \mathrm{diag}\{b_1(z), \ldots, b_s(z)\},$$

with

$$a_i(z) = \mu\left(1 - \frac{[AF(x)]_i}{\sqrt{[AF(x)]_i^2 + \lambda_i^2}}\right) + (1-\mu)(\lambda_i)_+ \partial([AF(x)]_i)_+,$$
$$b_i(z) = \mu\left(1 - \frac{\lambda_i}{\sqrt{[AF(x)]_i^2 + \lambda_i^2}}\right) + (1-\mu)([AF(x)]_i)_+ \partial(\lambda_i)_+,$$

if $[AF(x)]_i^2 + \lambda_i^2 > 0$, and

$$a_i(z) = \mu(1 - \xi_i), \quad b_i(z) = \mu(1 - \eta_i),$$

where $\xi_i^2 + \eta_i^2 \le 1$, if $[AF(x)]_i^2 + \lambda_i^2 = 0$.

From the above equations

$$a_i(z) \ge 0 \quad \text{and} \quad b_i(z) \ge 0 \quad \text{for all } i = 1, 2, \ldots, s,$$

which imply that D_a and D_b are positive semidefinite diagonal matrices.

To end this subsection, we propose a traditional way of obtaining the generalized Jacobian V at any point $z \in R^{n+s+t}$, which is similar to the algorithm in [2].

Algorithm 2.3 (The procedure to evaluate an element $V \in \partial\Theta(z)$).

Step 1. Let $z := (x, \lambda, \omega) \in R^{n+s+t}$ be given. Compute $F'(x)$ and $G'(x)$.

Step 2. Set

$$S_1 := \{i \mid \lambda_i = (AF(x))_i = 0\}$$

and

$$S_2 := \{i \mid \lambda_i > 0, (AF(x))_i > 0\}.$$

Step 3. Set $c \in R^s$ such that $c_i = 0$ for $i \notin S_1$ and $c_i = 1$ for $i \in S_1$.

Step 4. For $i \in S_1$, set

$$a_i = \mu\left(1 - \frac{c^{\mathrm{T}}(\nabla F(x) A^{\mathrm{T}})_i}{\|(c_i, c^{\mathrm{T}}(\nabla F(x) A^{\mathrm{T}})_i)\|}\right),$$
$$b_i = \mu\left(1 - \frac{c_i}{\|(c_i, c^{\mathrm{T}}(\nabla F(x) A^{\mathrm{T}})_i)\|}\right).$$

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

Step 5. For $i \in S_2$, set

$$a_i = \mu\left(1 - \frac{(AF(x))_i}{\|(\lambda_i, (AF(x))_i)\|}\right) + (1-\mu)\lambda_i,$$
$$b_i = \mu\left(1 - \frac{\lambda_i}{\|(\lambda_i, (AF(x))_i)\|}\right) + (1-\mu)(AF(x))_i.$$

Step 6. For $i \notin S_1 \cup S_2$, set

$$a_i = \mu\left(1 - \frac{(AF(x))_i}{\|(\lambda_i, (AF(x))_i)\|}\right),$$
$$b_i = \mu\left(1 - \frac{\lambda_i}{\|(\lambda_i, (AF(x))_i)\|}\right).$$

Step 7. According to (2.1), compute the generalized Jacobian V.

2.2. Stability in differential equations. Now we recall some facts about first order differential equations(ODE).

$$\dot{z}(t) = H(z(t)), \quad z(t_0) = z_0 \in R^n, \tag{2.2}$$

where $H : R^n \to R^n$ is a mapping. Next, we introduce three kinds of stability that will be discussed later (see [9]).

DEFINITION 2.4. A point $z^* = z(t^*)$ is called an equilibrium point or a steady state of the dynamic system (2.2) if $H(z^*) = 0$. If there is a neighborhood $\Omega^* \subseteq R^n$ of z^* such that $H(z^*) = 0$ and $H(z) \neq 0$ for $\forall z \in \Omega^* \backslash \{z^*\}$, then z^* is called an isolated equilibrium point.

LEMMA 2.5 [5]. *Assume that H is a continuous mapping from R^n to R^n. Then for arbitrary $t_0 \geq 0$ and $z_0 \in R^n$, there exists a local solution $z(t)$, $t \in [t_0, \tau)$ for some $\tau > t_0$. If, in addition, H is locally Lipschitz continuous at z_0, then the solution is unique; if H is Lipschitz continuous in R^n, then τ can be extended to ∞.*

If a local solution defined on $[t_0, \tau)$ cannot be extended to a local solution on a larger interval $[t_0, \tau_1)$, $\tau_1 > \tau$, then it is called a maximal solution, and the interval $[t_0, \tau)$ is the maximal interval of existence. Clearly, any local solution has an extension to a maximal one. We denote $[t_0, \tau(z_0))$ by the maximal interval of existence associated with z_0.

LEMMA 2.6 [5]. *Assume that $H : R^n \to R^n$ is continuous. If $z(t)$ with $t \in [t_0, \tau(z_0))$ is a maximal solution and $\tau(z_0) < \infty$, then*

$$\lim_{t \nearrow \tau(z_0)} \|z(t)\| = \infty.$$

DEFINITION 2.7 (Stability in the sense of Lyapunov). Let $z(t)$ be a solution of (2.2). An isolated equilibrium point z^* is Lyapunov stable if, for any $z_0 = z(t_0)$ and any scalar $\varepsilon > 0$, there exists a $\delta > 0$ such that if $\|z(t_0) - z^*\| < \delta$ then $\|z(t) - z^*\| < \varepsilon$ for $t \geq t_0$.

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

Definition 2.8 (Asymptotic stability). An isolated equilibrium point z^* is said to be asymptotically stable if, in addition to being Lyapunov stable, it has the property that $z(t) \to z^*$ as $t \to +\infty$ if $\|z(t_0) - z^*\| < \delta$.

Definition 2.9 (Lyapunov function). Let $\Omega \subseteq R^n$ be an open neighborhood of $\bar{z}$. A continuously differentiable function $W : R^n \to R^n$ is said to be a Lyapunov function at the state $\bar{z}$ over the set Ω for (2.2) if

$$\begin{cases} W(\bar{z}) = 0, \quad W(z) > 0, \quad \forall z \in \Omega \backslash \{\bar{z}\}, \\ \dfrac{dW(z(t))}{dt} = \nabla W_z(z(t))^{\mathrm{T}} H(z(t)) \le 0, \quad \forall z \in \Omega. \end{cases}$$

The next result addresses the relationship between stabilities and a Lyapunov function.

Lemma 2.10 [5].

(a) *An isolated equilibrium point z^* is Lyapunov stable if there exists a Lyapunov function over some neighborhood Ω^* of z^*.*

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

A stronger notion than the Lyapunov stability is the so-called exponential stability. Moreover, exponentially stable equilibria are asymptotically stable.

Definition 2.11 (Exponential stability). An isolated equilibrium point z^* is exponentially stable if there exist $\varrho > 0, \kappa > 0, \delta > 0$ such that an arbitrary solution $z(t)$ of (2.2) with the initial conditions $z(t_0) = z_0$ and $\|z(t_0) - z^*\| < \delta$ is defined on $[0, \infty)$ and satisfies

$$\|z(t) - z^*\| \le \kappa e^{-\varrho t} \|z(t_0) - z^*\|, \quad t \ge t_0.$$

To end this subsection, we provide a diagram to show how the first order ANN (1.4) is implemented on hardware (see Figure 1).

3. Convergence and stability of the trajectory

In this section, we focus on two aspects of the stability issues of the first order ANN (1.4). Firstly, we analyze the general behavior of the solution trajectory of (1.4), including various properties such as existence, uniqueness and convergence. Secondly, we address three kinds of stability for an isolated equilibrium.

Proposition 3.1. *For $z^* = (x^*, \lambda^*, \omega^*) \in R^{n+s+t}$, x^* is a solution to GNCP (F, G, K) and $\Theta(z^*) = 0$. Then z^* is an equilibrium point of the ANN (1.4).*

Proof. Since $\Theta(z^*) = 0$, according to Lemma 2.2, $\nabla f(z^*) = 0$. So z^* is the equilibrium point of (1.4). □

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

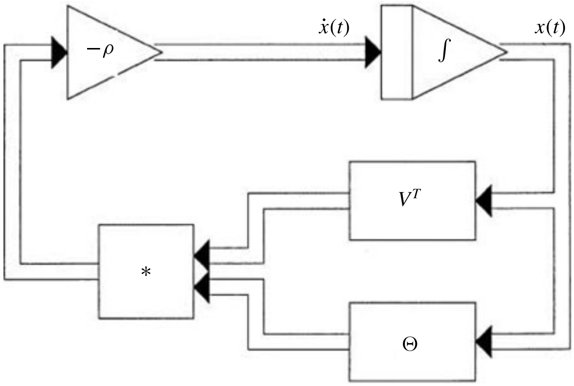

Figure 1. A simplified block diagram for the first order ANN (1.4).

The following proposition gives a suitable condition that guarantees that every stationary point of the ANN (1.4) solves the GNCP.

Proposition 3.2. *Suppose that $z^* = (x^*, \lambda^*, \omega^*)$ is an equilibrium point of the ANN (1.4), $F'(x^*)$ is nonsingular, and $G'(x^*)[F'(x^*)]^{-1}$ is positive definite in the null space of B. Then x^* is a solution of the GNCP (F, G, K).*

Proof. Since $z^* = (x^*, \lambda^*, \omega^*)$ is an equilibrium point of the ANN (1.4) (that is $\nabla f(z^*) = 0$), and D_a, D_b are positive semidefinite diagonal matrices. Using a method similar to the proof of Theorem 4.1 in [14], we can verify that $\Theta(z^*) = 0$: that is, x^* is the solution of the GNCP (F, G, K). □

Proposition 3.3. *Suppose $z^* = (x^*, \lambda^*, \omega^*)$ is an equilibrium point of the ANN (1.4), $F'(x^*)$ and $G'(x^*)$ are nonsingular, the matrix B has full row rank, and $AF'(x^*)[G'(x^*)]^{-1}A^{\mathrm{T}}$ is a P-matrix. Then x^* is a solution of the GNCP (F, G, K).*

Proof. Since $z^* = (x^*, \lambda^*, \omega^*)$ is an equilibrium point of the ANN (1.4) (that is $\nabla f(z^*) = 0$). Using a method similar to the proof of [14, Theorem 4.2], we can verify that any $V \in \partial\Theta(z^*)$ is nonsingular. Then we have $\Theta(z^*) = 0$: that is, x^* is the solution of the GNCP (F, G, K). □

We recall that z^* is said to be a regular solution to $f(z) = 0$ if every $V \in \partial\Theta(z^*)$ is nonsingular. According to Lemmas 1.1 and 2.2, it is easy to get the following result.

Proposition 3.4. *If $z^* = (x^*, \lambda^*, \omega^*) \in R^{n+s+t}$ is an equilibrium point of the ANN (1.4), and all $V \in \partial\Theta(z^*)$ are nonsingular. Then z^* is a regular solution to $f(z) = 0$.*

Proposition 3.5. *The function $f(z(t))$ is nonincreasing with respect to the variable t.*

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

PROOF. Since

$$\frac{df(z(t))}{dt} = \nabla f(z(t))^{\mathrm{T}} \frac{dz(t)}{dt} = \nabla f(z(t))^{\mathrm{T}}[-\rho \nabla f(z(t))] = -\rho \|\nabla f(z(t))\|^2 \le 0.$$

This concludes the proof. □

Let $\Omega(z_0)$ denote the level set associated with the initial state z_0 and be given by

$$\Omega(z_0) = \{z \in R^{n+s+t} \mid f(z) \le f(z_0)\}.$$

THEOREM 3.6. *For an arbitrary initial state* $z_0 := (x_0, \lambda_0, \omega_0) \in R^{n+s+t}$*, the following results hold:*

(a) *there exists a local solution* $z(t)$*,* $t \in [t_0, \tau(z_0))$ *for some* $\tau(z_0) > t_0$*; and*
(b) *if* $\Omega(z_0)$ *is bounded, then* $\tau(z_0) = +\infty$*.*

PROOF. (a) It is known that $\nabla f(z)$ is continuous. Hence Lemma 2.5 implies this result.

(b) If $\tau(z_0) < +\infty$, it follows from Lemma 2.6 that

$$\lim_{t \nearrow \tau(z_0)} \|z(t)\| = \infty.$$

Let

$$\tau_0 = \inf\{s \ge 0 \mid s < \tau(z_0), z(s) \in \Omega^c(z_0)\} < \infty,$$

where $\Omega^c(z_0)$ is the complement of the set $\Omega(z_0)$ in R^{n+s+t}. Moreover, $\Omega(z_0)$ is compact since it is bounded by assumption and it is also closed because of the continuity of f. Therefore, we have $z(\tau_0) \in \Omega(z_0)$ and $\tau_0 < \tau(z_0)$, implying that

$$f(z(s)) > f(z_0) \ge f(z(\tau_0)) \quad \text{for some } s \in (\tau_0, \tau(z_0)). \tag{3.1}$$

However, Proposition 3.5 says that $f(z(\cdot))$ is nonincreasing on $[t_0, \tau(z_0))$, which contradicts (3.1). This completes (b). □

COROLLARY 3.7. *For an arbitrary initial state* $z_0 := (x_0, \lambda_0, \omega_0) \in R^{n+s+t}$*, let* $z(t) : [t_0, \tau(z_0))$ *be the unique maximal solution to (1.4). If* $\tau(z_0) = +\infty$ *and* $\{z(t)\}$ *is bounded, then*

$$\lim_{t \to +\infty} \nabla f(z(t)) = 0. \tag{3.2}$$

Moreover, if z^* *is a accumulation point of the trajectory* $z(t)$ *and all* $V \in \partial\Theta(z^*)$ *are nonsingular, then* x^* *is a solution to the GNCP* (F, G, K)*.*

PROOF. It is proved in Proposition 3.5 that $f(z(t))$ is a nonincreasing function with respect to t. We also note that $f(z(t))$ is a nonnegative function over R^{n+s+t}: that is, $f(z(t))$ is bounded from below. Those arguments are exactly the same as those for [9, Corollary 4.3]. Thus we omit them. If z^* is a accumulation point of the trajectory $z(t)$ (that is, $\lim_{t \to +\infty} z(t) = z^*$), it follows, from (3.2), that $\nabla f(z^*) = 0$. Then applying Proposition 3.4 leads to the desired result. □

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

THEOREM 3.8. *If z^* is an isolated equilibrium point of (1.4), then z^* is asymptotically stable for (1.4).*

PROOF. Using a method similar to [9, Theorem 4.4], we can show that $f(x)$ is a Lyapunov function over the set Ω^* for (1.4). Furthermore, using Lemma 2.10(b), we have that z^* is asymptotically stable for (1.4). □

THEOREM 3.9. *If z^* is a regular solution to $f(z) = 0$, then z^* is exponentially stable for (1.4).*

PROOF. Since F and G are both continuously differentiable, then Θ is semismooth which is similar to the [14, Lemma 2.4]. Furthermore, with Lemma 2.2 and Theorem 3.8, the arguments are exactly the same as those for [9, Theorem 4.6]. Thus we omit them. □

4. Simulation results

In this section, we test some examples obtained by the neural network model (1.4). The numerical implementation is coded by Matlab 2011(b) and the ordinary differential equation solver adopted is ode23s.

EXAMPLE 4.1 [15, Example 2]. Find $x \in R^5$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad x^{\mathrm{T}} F(x) = 0,$$

where

$$F(x) = \begin{pmatrix} x_1 + x_2x_3x_4x_5/50 \\ x_2 + x_1x_3x_4x_5/50 - 3 \\ x_3 + x_1x_2x_4x_5/50 - 1 \\ x_4 + x_1x_2x_3x_5/50 - 0.5 \\ x_5 + x_1x_2x_3x_4/50 \end{pmatrix}.$$

It has only one solution,

$$x^* = (0, 3, 1, 0, 0)^{\mathrm{T}}, \quad F(x^*) = (0, 0, 0, 0.5, 0)^{\mathrm{T}}.$$

For Example 4.1, four starting vectors are used, namely,

$$x_0^{(1)} = (0.01, 1, 0.5, 0.01, 0.01)^{\mathrm{T}}, \quad x_0^{(2)} = (1, 1, 1, 1, 1)^{\mathrm{T}},$$
$$x_0^{(3)} = (5, 5, 5, 5, 5)^{\mathrm{T}}, \quad x_0^{(4)} = (10, 10, 10, 10, 10)^{\mathrm{T}}.$$

EXAMPLE 4.2 [7]. Find $x \in R^4$ such that

$$x \geq 0, \quad F(x) \geq 0, \quad x^{\mathrm{T}} F(x) = 0,$$

where

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

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

This is a nondegenerate GNCP and the solution is

$$x^* = \left(\frac{\sqrt{6}}{2}, 0, 0, 0.5\right)^{\mathrm{T}}, \quad F(x^*) = \left(0, 2 + \frac{\sqrt{6}}{2}, 5, 0\right)^{\mathrm{T}}.$$

For Example 4.2, four starting vectors are used, namely,

$$x_0^{(1)} = (2, 0.01, 0.01, 0.1)^{\mathrm{T}}, \quad x_0^{(2)} = (0, 0, 0, 0)^{\mathrm{T}},$$
$$x_0^{(3)} = (1, 1, 1, 1)^{\mathrm{T}}, \quad x_0^{(4)} = (10, 10, 10, 10)^{\mathrm{T}}.$$

Example 4.3. Find $x \in R^4$ such that

$$F(x) \geq 0, \quad G(x) = x - m(x) \geq 0, \quad F(x)^{\mathrm{T}} G(x) = 0,$$

where

$$F(x) = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

and $m(x) = \varphi(F(x)) : R^4 \to R^4$ is twice continuously differentiable. The function $\varphi(\cdot)$ that defines our test problems is

$$\varphi_i(t) = 0.5 - t_i, \quad i = 1, 2, 3, 4.$$

It has only one solution,

$$x^* = (-0.3, -0.4, -0.4, -0.3)^{\mathrm{T}},$$

and

$$F(x^*) = (0.8, 0.9, 0.9, 0.8)^{\mathrm{T}}, \quad G(x^*) = x^* - m(x^*) = (0, 0, 0, 0)^{\mathrm{T}}.$$

For Example 4.3, four starting vectors are used, namely,

$$x_0^{(1)} = (0, 0, 0, 0)^{\mathrm{T}}, \quad x_0^{(2)} = (-0.5, -0.5, -0.5, -0.5)^{\mathrm{T}},$$
$$x_0^{(3)} = (1, 1, 1, 1)^{\mathrm{T}}, \quad x_0^{(4)} = (10, 10, 10, 10)^{\mathrm{T}}.$$

Firstly, we test the influence of the parameters μ and ρ on the value $\|x(t) - x^*\|$. From Figures 2 and 3, we can see that it generates the fastest decrease of $\|x(t) - x^*\|$ when $\mu = 0.99$ and $\rho = 2$. We will emphasize that the convergence behavior of the error $\|x(t) - x^*\|$ is astable when $\rho \geq 3$. We set $\mu = 0.95$, $\rho = 2$ in the following computational experiments.

The transient behavior of $x(t)$ for Examples 4.1–4.3 is depicted in Figures 4, 6 and 8, respectively. We can see that they are both very close to the solution of the GNCP with suitable initial states.

Figures 5, 7 and 9 describe how $\|x(t) - x^*\|$ varies with different initial states. We emphasize that the initial state x_0 is not required to be close to the solution.

In summary, the neural network (1.4) is a better alternative to the network based on the penalized Fischer–Burmeister function if appropriate values of μ and ρ are chosen. From the numerical simulations above, we see that, to obtain a better convergence rate of the trajectory $x(t)$, the parameter μ cannot be set too small and the parameter ρ cannot be set too large. In addition, we should emphasize that the initial state $x(t_0)$ has little influence on the convergence behavior of $\|x(t) - x^*\|$.

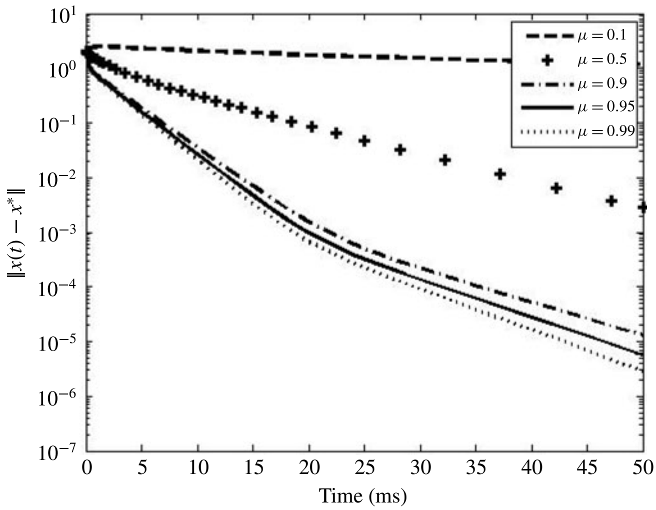

FIGURE 2. Convergence behavior of the error $\|x(t) - x^*\|$ in Example 4.1 with five different values of μ(0.1, 0.5, 0.9, 0.95 and 0.99) and initial point $x_0^{(1)}$.

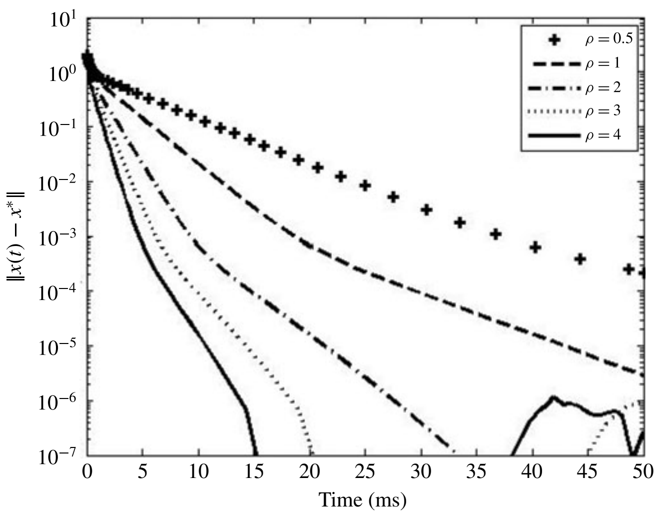

FIGURE 3. Convergence behavior of the error $\|x(t) - x^*\|$ in Example 4.1 with five different values of ρ(0.5, 1, 2, 3 and 4) and initial point $x_0^{(1)}$.

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

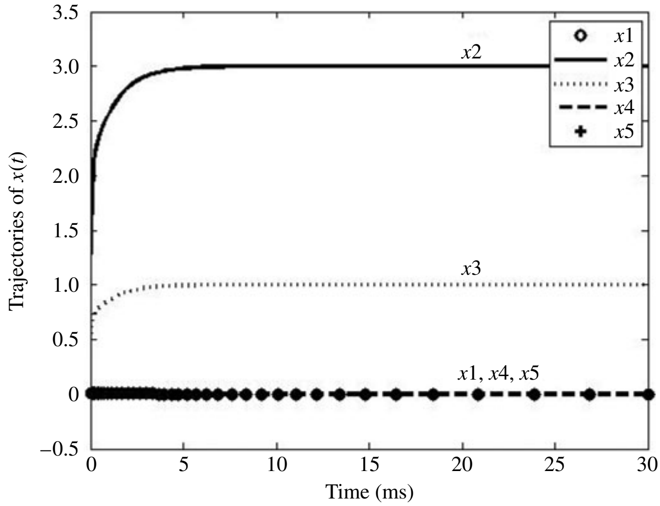

FIGURE 4. Transient behavior of $x(t)$ in Example 4.1 with initial point $x_0^{(1)}$.

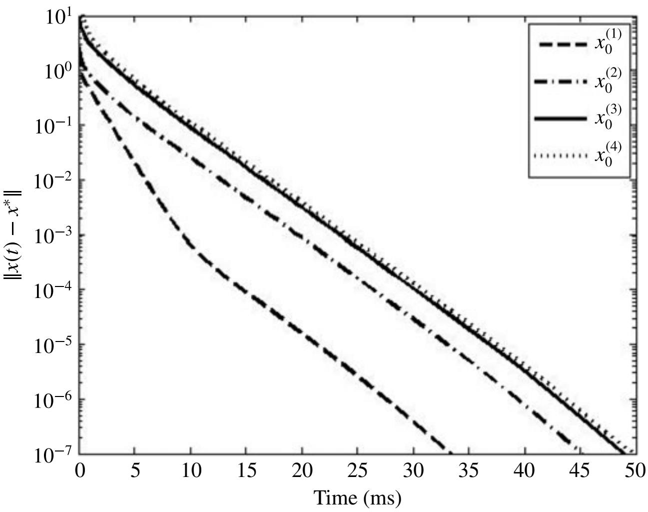

FIGURE 5. Convergence behavior of the error $\|x(t) - x^*\|$ in Example 4.1 with four different initial points $x_0^{(1)}, x_0^{(2)}, x_0^{(3)}$ and $x_0^{(4)}$.

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

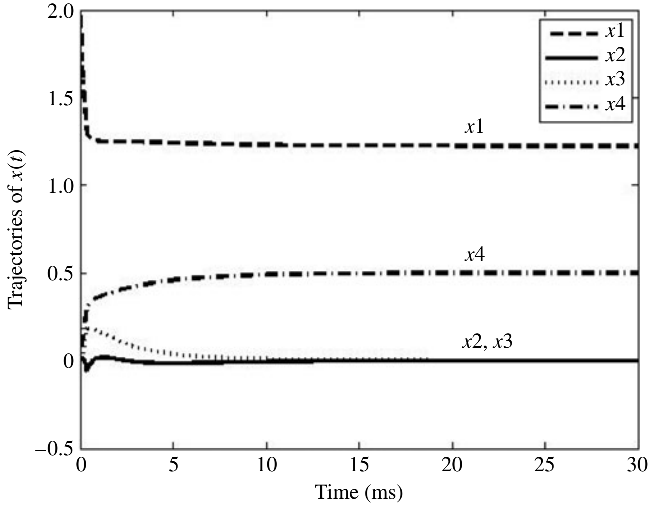

FIGURE 6. Transient behavior of $x(t)$ in Example 4.2 with initial point $x_0^{(1)}$.

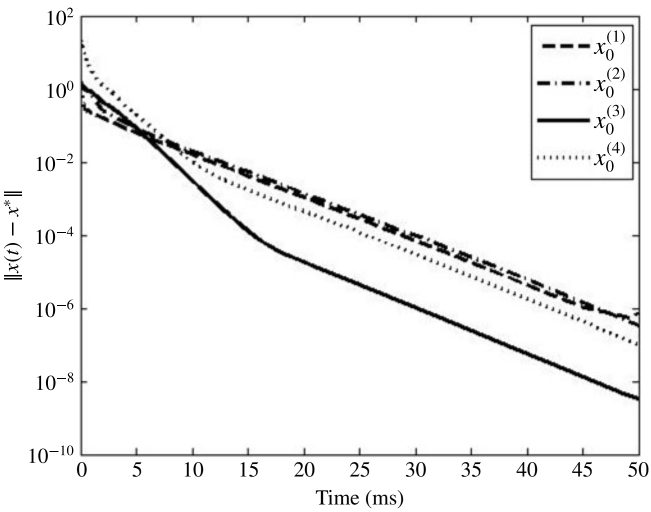

FIGURE 7. Convergence behavior of the error $\|x(t) - x^*\|$ in Example 4.2 with four different initial points $x_0^{(1)}, x_0^{(2)}, x_0^{(3)}$ and $x_0^{(4)}$.

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

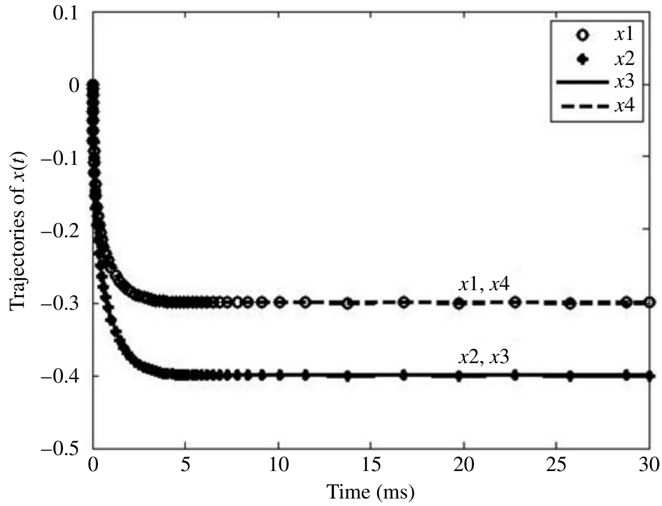

FIGURE 8. Transient behavior of $x(t)$ in Example 4.3 with initial point $x_0^{(1)}$.

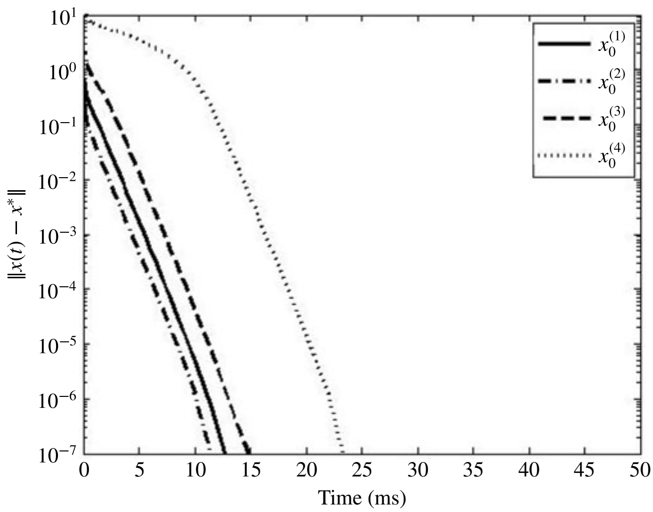

FIGURE 9. Convergence behavior of the error $\|x(t) - x^*\|$ in Example 4.3 with four different initial points $x_0^{(1)}, x_0^{(2)}, x_0^{(3)}$ and $x_0^{(4)}$.

https://doi.org/10.1017/S1446788715000300 Published online by Cambridge University Press

Acknowledgements

The authors are most grateful to the anonymous referees for their constructive comments and helpful suggestions, which greatly improved the original paper.

References

[1] A. Bouzerdoum and T. R. Pattison, 'Neural network for quadratic optimization with bound constraints', *IEEE Trans. Neural Netw.* **4** (1993), 293–304.
[2] B.-T. Chen, X.-J. Chen and C. Kanzow, 'A penalized Fischer–Burmeister NCP-function: theoretical investigation and numerical results', *Math. Program.* **88**(1) (1997), 211–216.
[3] X.-H Chen and C.-F. Ma, 'A regularization smoothing Newton method for solving nonlinear complementarity problem', *Nonlinear Anal. Real World Appl.* **10** (2009), 1702–1711.
[4] B.-L. Chen and C.-F. Ma, 'Superlinear/quadratic smoothing Broyden-like method for the generalized nonlinear complementarity problem', *Nonlinear Anal. Real World Appl.* **12** (2011), 1250–1263.
[5] J.-S. Chen, C. H. Ko and S. H. Pan, 'A neural network based on the generalized Fischer–Burmeister function for nonlinear complementarity problems', *Inform. Sci.* **180** (2010), 697–711.
[6] J. J. Hopfield and D. W. Tank, 'Neural computation of decision in optimization problems', *Biol. Cybernet.* **52** (1985), 141–152.
[7] M. Kojima and S. Shindo, 'Extensions of Newton and quasi-Newton methods to systems of PC^1 equations', *J. Oper. Res. Soc. Japan* **29** (1986), 352–374.
[8] L.-Z. Liao and H.-D. Qi, 'A neural network for the linear complementarity problem', *Math. Comput. Model.* **29** (1999), 9–18.
[9] L.-Z. Liao, H.-D. Qi and L. Qi, 'Solving nonlinear complementarity problems with neural networks: a reformulation method approach', *J. Comput. Appl. Math.* **131** (2001), 342–359.
[10] C.-F. Ma, 'A new smoothing and regularization Newton method for P_0-NCP', *J. Global Optim.* **48** (2010), 241–261.
[11] C.-F. Ma, L.-J. Chen and D.-S. Wang, 'A globally and superlinearly convergent smoothing Broyden-like method for solving nonlinear complementarity problem', *Appl. Math. Comput.* **198** (2008), 592–604.
[12] C.-F. Ma, L.-H. Jiang and D.-S. Wang, 'The convergence of a smoothing damped Gauss–Newton method for nonlinear complementarity problem', *Nonlinear Anal. Real World Appl.* **10** (2009), 2072–2087.
[13] D. W. Tank and J. J. Hopfield, 'Simple neural optimization networks: an A/D converter, signal decision circuit, and a linear programming circuit', *IEEE Trans. Circuits Syst. I. Regul. Pap.* **33** (1986), 533–541.
[14] Y.-J. Wang, F.-M. Ma and J.-Z. Zhang, 'A nonsmooth L-M method for solving the generalized nonlinear complementarity problem over a polyhedral cone', *Appl. Math. Optim.* **52**(1) (2005), 73–92.
[15] Y. Xia, H. Leung and J. Wang, 'A projection neural network and its application to constrained optimization problems', *IEEE Trans. Circuits Syst. I. Regul. Pap.* **49** (2002), 447–458.
[16] S. H. Zak, V. Upatising and S. Hui, 'Solving linear programming problems with neural networks: a comparative study', *IEEE Trans. Neural Netw.* **6** (1995), 94–104.

YIFEN KE, School of Mathematics and Computer Science,
Fujian Normal University, Fuzhou 350117, PR China
e-mail: yfke89@163.com

CHANGFENG MA, School of Mathematics and Computer Science,
Fujian Normal University, Fuzhou 350117, PR China
e-mail: macf@fjnu.edu.cn