2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

A Single Layer Recurrent Neural Network For Pseudoconvex Optimization Subject to Quasiconvex Constraints

Jingjing Huang and Guocheng Li

***Abstract*—This paper presents a single layer recurrent network for solving optimization problems with pseudoconvex objectives subject to quasiconvex constraints. The penalty method using a finite penalty parameter is applied for the design and analysis of the neural network. The lower bounder of the penalty parameter is given in order to guarantee the exact penalty property. It is rigorously proved that the neural network is globally convergent to the global optimal solution of the corresponding optimization problem. Simulation results are included to illustrate the performances of the proposed neural network.**

I. Introduction

RECURRENT neural networks (RNNs) constitute one of the most successful computational intelligent models. They have achieved great successes in many engineering applications, such as kinematic control of redundant robot manipulators [1], nonlinear model predictive control [2], [3], hierarchical control of interconnected dynamic systems [4] , compressed sensing in adaptive signal processing [5], and so on. Particularly, since the pioneering work of Hopfield neural networks [6], [7], RNNs have shown promises for online optimization. Compared with traditional numerical optimization algorithms, RNNs offer a highly computationally efficient optimization paradigm due to the parallel and distributed information processing.

The past three decades witnessed remarkable progress in the area of online optimization using RNNs. For example, a deterministic annealing neural network was proposed for solving convex programming problems based on the simulated annealing algorithm [8], a Lagrangian network was developed for solving convex optimization problems with linear equality constraints based on the Lagrangian optimality conditions [9], the dual networks [10]–[12] were developed for solving convex optimization problems based on the Karush-Kuhn-Tucker optimality conditions, projection neural networks were developed for constrained optimization problems based on the projection method [13]–[16]. These neural networks models can globally converge to the unique optimal solutions of convex optimization problems.

In addition to convex optimization, RNNs have been successfully extended to optimization problems with general convex objectives. For example, [17] proposed a projection neural network for solving pseudomonotone variational inequalities and pseudoconvex optimization problems. [18] proposed a recurrent neural network for solving the differentiable pseudoconvex optimization problems with linear equality constraints. [19] proposed a finite-time convergent recurrent neural network for constrained optimization problems with piecewise-linear objective functions, [20] proposed a penalty-based recurrent neural network for solving a class of constrained optimization problems with generalized convex objective functions. [21] proposed a one-layer projection neural network without any design parameter for solving nonsmooth optimization problems with generalized convex objective functions.

The authors are with Department of Mathematics, Beijing Information Science and Technology University, Beijing, China (email: xyliguocheng@sohu.com).

Despite of the effectiveness and usefulness and the aforementioned RNNs, there are still some issues worth in-depth investigation. Specifically, previous studies mainly focused on extending the objectives to generalized convex functions, but they did not pay much attention on the constraints. In other words, most existing neural network models can only deal with convex constraints. However, many real world optimization problems involve nonconvex constraints. It is obviously interesting and necessary to investigate RNNs that allow the constraint functions not necessarily to be convex.

In this paper, a single layer recurrent neural network is proposed for optimization problems with pseudoconvex objectives and quasiconvex constraint functions. To design the neural network, the penalty method is applied where an exact penalty function with finite penalty parameter is constructed. It is proved that the neural network is globally convergent to its equilibrium point which corresponds to the global optimal solution of the underlying optimization problem.

The remainder of this paper is organized as follows. Section II introduces some definitions and preliminary results. Section III discusses an exact penalty function. Section IV presented a neural network model and analyzed its convergent properties. Section V provides simulation results. Finally, Section VI concludes this paper.

II. Preliminaries

Consider a constrained optimization in the following form

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_i(x) \le 0,\ i = 1, 2, \cdots, m, \end{array} \tag{1}$$

where $x \in R^n$ is the decision vector; f and $g_i, : R^n \to R$ ($i = 1, 2, \cdots, m$) are continuously differentiable functions, but not necessarily convex. f is assumed to be radially unbounded. The feasible region

$$\mathcal{F} = \{x \in R^n : g_i(x) \le 0,\ i = 1, 2, \cdots, m\}$$

is assumed to be a nonempty set. The global solutions of the problem (1) is defined as

$$\mathcal{G} = \{x^* \in \mathcal{F} : f(x) \ge f(x^*),\ \forall\ x \in \mathcal{F}\}.$$

Some definitions and propositions are presented which are needed to obtain the main results. We refer readers to [22]–[26] for a more thorough research on these topics.

Definition 1: A differentiable function f, defined on an open convex set $\mathcal{D} \subset R^n$, is called pseudoconvex if

$$x_1,\ x_2 \in \mathcal{D}, f(x_1) > f(x_2) \Rightarrow \nabla f(x_1)^T(x_2 - x_1) < 0.$$

Definition 2: A differentiable function f, defined on an open convex set $\mathcal{D} \subset R^n$, is called quasiconvex if

$$x_1,\ x_2 \in \mathcal{D}, f(x_1) \geq f(x_2) \Rightarrow \nabla f(x_1)^T(x_2 - x_1) \leq 0.$$

Definition 3: A function $f : R^n \to R$ satisfies the growth condition if

$$\lim_{||x|| \to +\infty} f(x) = +\infty.$$

Proposition 1: A function $f : R^n \to R$ satisfies the growth condition if and only if $\forall \alpha \in R$, level set $L(\alpha) = \{x \in R^n : f(x) \leq \alpha\}$ is bounded.

Proposition 2: Let $f : R^n \to R$ be continuously differentiable. Then $\max\{0, f(x)\}$ is a regular function, it's Clarke's generalized gradient as follows:

$$\partial \max\{0, f(x)\} = \begin{cases} \nabla f(x), & \text{for } f(x) > 0; \\ (0,1)\nabla f(x), & \text{for } f(x) = 0; \\ 0, & \text{for } f(x) < 0. \end{cases}$$

Proposition 3: If $f : R^n \to R$ is a regular at $x(t)$ and $x : R \to R^n$ is differentiable at t and Lipschitz near t, then

$$\frac{d}{dt} f(x(t)) = \langle \xi, \dot{x}(t) \rangle \quad \forall \xi \in \partial f(x(t)).$$

Let $\mathcal{A} \subset R^n$, we denote the closures of $\mathcal{A}$ by cl$\mathcal{A}$, the internal of $\mathcal{A}$ by int$\mathcal{A}$, the border of $\mathcal{A}$ by bd$\mathcal{A}$, the complement of $\mathcal{A}$ by $\mathcal{A}^c$ and the ball centered at the origin with radius 1 by $\mathcal{B}$.

Throughout this paper, the following assumptions hold.

Assumption 1: The objective function $f(x)$ satisfies the growth condition.

Assumption 2: The gradients of constraint functions $\nabla g_i(z), i = 1, \ldots, m$, are linearly independent, where $g_1(z) = g_2(z) = \ldots = g_m(z) = 0$.

Assumption 3: $\mathcal{F} = \text{cl}(\text{int}\mathcal{F})$.

III. Exact Penalty Function

In this section, a penalty function is defined and analyzed based on an appropriate neighborhood of the feasible region $\mathcal{F}$. Consider the following function:

$$V(x) = \sum_{i=1}^{m} \max\{0, g_i(x)\}.$$

$V(x)$ is a continuous function and satisfies the growth condition. By Proposition 1 the feasible region $\mathcal{F} = \{x \in R^n : V(x) \leq 0\}$ is a compact subset. For any $x \in R^n$, we define the index sets:

$$I_0(x) = \{i : g_i(x) = 0,\ i \in I\},$$

$$I_+(x) = \{i : g_i(x) > 0,\ i \in I\},$$

$$I_-(x) = \{i : g_i(x) < 0,\ i \in I\}.$$

The Clarke's generalized gradient of $V(x)$ as follows:

$$\partial V(x) = \sum_{i \in I_+(x)} \nabla g_i(x) + \sum_{i \in I_0(x)} [0,1] \nabla g_i(x). \tag{2}$$

Since $V(x)$ is a continuous function and satisfies the growth condition, then $\exists\ r > 0$ such that

$$\mathcal{D} := \{x \in R^n : V(x) < r\} \subseteq \mathcal{F} + R\mathcal{B}.$$

For the problem (1), a penalty function is commonly defined as follows:

$$E_\sigma(x) = f(x) + \frac{1}{\sigma} V(x),$$

where $\sigma > 0$ is penalty parameter. The $E_\sigma(x)$ is a continuous function and satisfies the growth condition.

Consider the following problem:

$$\min E_\sigma(x),\ x \in \mathcal{D}. \tag{3}$$

The Clarke's generalized gradient of $E_\sigma(x)$ as follows:

$$\partial E_\sigma(x) = \nabla f(x) + \frac{1}{\sigma} \partial V(x).$$

Since $\mathcal{D}$ is an open set, any local solution of problem (3), provided it exists, is unconstrained; thus problem (3) can be considered as an essentially unconstrained problem. The sets of global solutions of problem (3) is denoted by $\mathcal{G}(\sigma)$:

$$\mathcal{G}(\sigma) = \{x \in \mathcal{D} : E_\sigma(y) \geq E_\sigma(x),\ \forall y \in \mathcal{D}\}.$$

Definition 4: The function $E_\sigma(x)$ is an exact penalty function for problem (1) with respect to the set $\mathcal{D}$ if there exists an $\sigma^* > 0$ such that for all $\sigma \in (0, \sigma^*]$, $\mathcal{G}(\sigma) = \mathcal{G}$.

The following lemma plays an important role in convergence analysis.

Lemma 1: There exist $R > 0$ and $m_g > 0$ such that

$$\min_{x \in (\mathcal{F} + R\mathcal{B}) \setminus \text{int}\mathcal{F}} \text{dist}(0, \partial V(x)) \geq m_g > 0. \tag{4}$$

Proof: First, we prove that $\forall\ x \in \text{bd}\mathcal{F}$, $0 \notin \partial V(x)$. If not, then $\exists\ \tilde{x} \in \text{bd}\mathcal{F}$ such that, $\text{dist}(0, \partial V(\tilde{x})) = 0$. Since $x \to \text{dist}(0, \partial V(x))$ is a lower semi-continuous real-valued function and $\mathcal{F} = \text{cl}(\text{int}\mathcal{F})$, we can take $\{x_k\} \subset \mathcal{F}^c$, $x_k \to \tilde{x}$ and $\eta_k \in \partial V(x_k)$ such that $\lim_{k\to\infty} ||\eta_k|| = 0$. There exist $\alpha_{k_i} \in [0,1]$ $(i \in I_0(x_k))$ such that

$$\eta_k = \sum_{i \in I_+(x_k)} \nabla g_i(x_k) + \sum_{i \in I_0(x_k)} \alpha_{k_i} \nabla g_i(x_k). \tag{5}$$

Extracting a subsequence and re-indexing, we assume with out loss of generality that, for all natural numbers k, $I_+(x_k)$ = $I_+(x_1)$, $I_0(x_k) = I_0(x_1)$ and $\lim_{k\to\infty} \alpha_{k_i} = \alpha_i$ Taking $k \to \infty$ in (5), then

$$\sum_{i \in I_+(x_1)} \nabla g_i(\tilde{x}) + \sum_{i \in I_0(x_1)} \alpha_i \nabla g_i(\tilde{x}) = 0. \tag{6}$$

Note that $I_+(x_1) \neq \emptyset$ and $g_i(\tilde{x}) = 0$ ($i \in I_+(x_1) \cup I_0(x_1)$), (6) is a contradiction to Assumption 2, thus $\forall\ x \in \text{bd}\mathcal{F}$, $0 \notin \partial V(x)$.

Next, by Proposition 4.1 in [27], there exist $R > 0$ and $m_g > 0$ such that

$$\min_{x \in (\mathcal{F}+R\mathcal{B})\setminus \text{int}\mathcal{F}} \text{dist}(0, \partial V(x)) \geq m_g > 0.$$

■

The following theorem establishes a sufficient condition for $E_\sigma(x)$ to be an exact penalty function.

Theorem 1: $E_\sigma(x)$ is an exact penalty function for (1) with respect to the set $\mathcal{D}$.

Proof: By the compactness of cl$\mathcal{D}$ and the continuity of $E_\sigma(x)$, for all $\sigma > 0$, $E_\sigma(x)$ admits a global minimum point on cl$\mathcal{D}$. We show first that there exists an $\sigma_1^* > 0$ such that, for all $\sigma \in (0, \sigma_1^*]$ we have $\mathcal{G}(\sigma) \neq \emptyset$. Suppose that this assertion is false. Then, for any integer k there must exist an $\sigma_{1k} \leq 1/k$ and $y^{(k)} \in$ bd$\mathcal{D}$ such that

$$E_{\sigma_{1k}}(y^{(k)}) = \inf_{x \in \text{cl}\mathcal{D}} E_{\sigma_1 k}(x).$$

There exists a convergent subsequence, which we relabel $\{y^{(k)}\}$ such that, $\lim_{k\to\infty} y^{(k)} = \bar{y}$, and $\bar{y} \in$ bd$\mathcal{D}$. For each natural number k,

$$\begin{array}{rl} & \inf_{x\in\mathcal{F}} f(x) = \inf_{x\in\mathcal{F}} E_{\sigma_{1k}}(x) \geq \inf_{x\in \text{cl}\mathcal{D}} E_{\sigma_{1k}}(x) \\ = & E_{\sigma_{1k}}(y^{(k)}) = f(y^{(k)}) + \frac{1}{\sigma_{1k}} V(y^{(k)}) \\ \geq & \inf_{x\in \text{cl}\mathcal{D}} f(x) + \frac{1}{\sigma_{1k}} V(y^{(k)}). \end{array} \quad (7)$$

It follows from (7) that

$$\begin{array}{l} \max\{g_i(\bar{y}), 0\} = \lim_{k\to\infty} \max\{g_i(y^{(k)}), 0\} \\ \leq \limsup_{k\to\infty} \sigma_{1k}[\inf_{x\in\mathcal{F}} f(x) - \inf_{x\in \text{cl}\mathcal{D}} f(x)] = 0. \end{array}$$

Therefore,

$$g_i(\bar{y}) \leq 0, \quad \text{for all } i \in I. \quad (8)$$

From (8), we have $\bar{y} \in \mathcal{F} \subset \mathcal{D}$, which is a contradiction with $\bar{y} \in$ bd$\mathcal{D}$. Therefore, there exists an $\sigma_1^* > 0$ such that, for all $\sigma \in (0, \sigma_1^*]$ we have $\mathcal{G}(\sigma) \neq \emptyset$.

Next, we will prove that there exists an $\sigma^* > 0$ $(0 < \sigma^* \leq \sigma_1^*)$ such that, for all $\sigma \in (0, \sigma^*]$ we have $\mathcal{G}(\sigma) \subseteq \mathcal{G}$. Namely, there exists $\sigma^* > 0$, such that if $\sigma \in (0, \sigma^*]$, and $z \in \mathcal{D}$ satisfies

$$E_\sigma(z) = \inf_{x\in\mathcal{D}} E_\sigma(x),$$

then $z \in \mathcal{F}$, and $f(z) = \inf_{x\in\mathcal{F}} f(x)$.

Let us assume the converse. Then there exist a sequence $\{\sigma_k\}_{k=1}^\infty \subset (0, \sigma_1^*]$ and a sequence $\{z^{(k)}\}_{i=1}^\infty \subset \mathcal{D}$ such that for all natural numbers k,

$$\sigma_k \leq \frac{1}{k},\ E_{\sigma_k}(z^{(k)}) = \inf_{x\in\mathcal{D}} E_{\sigma_k}(x),\ \ z^{(k)} \notin \mathcal{F} \quad (9)$$

For each natural number k,

$$\begin{array}{rl} & \inf_{x\in\mathcal{F}} f(x) = \inf_{x\in\mathcal{F}} E_{\sigma_k}(x) \geq \inf_{x\in\mathcal{D}} E_{\sigma_k}(x) \\ = & E_{\sigma_k}(z^{(k)}) = f(z^{(k)}) + \frac{1}{\sigma_k} V(z^{(k)}) \\ \leq & \inf_{x\in\mathcal{D}} f(x) + \frac{1}{\sigma_k} V(z^{(k)}) \end{array} \quad (10)$$

Extracting a subsequence and re-indexing, we may assume without loss of generality that there exists

$$\lim_{k\to\infty} z^{(k)} = \bar{z}. \quad (11)$$

By (10) and (11), similar to proof of (8), we have $\bar{z} \in \mathcal{F}$.

Since $z^{(k)}$ is a global minimizer of $E_{\sigma_k}(x)$ on $\mathcal{D}$, then for each integer $k \geq 1$,

$$0 \in \partial E_\sigma(z^{(k)}) = \nabla f(z^{(k)}) + \frac{1}{\sigma_k} \partial V(z^{(k)}),$$

$$0 \in \nabla f(z^{(k)}) + \frac{1}{\sigma_k}[\sum_{i\in I_+(z^{(k)})} \nabla g_i(z^{(k)}) + \sum_{i\in I_0(z^{(k)})} [0,1] \nabla g_i(z^{(k)})]. \quad (12)$$

For each integer $k \geq 1$, $\exists\ \alpha_{ki} \in [0,1]$ such that

$$0 = \sigma_k \nabla f(z^{(k)}) + \sum_{i\in I_+(z^{(k)})} \nabla g_i(z^{(k)}) + \sum_{i\in I_0(z^{(k)})} \alpha_{ki} \nabla g_i(z^{(k)}). \quad (13)$$

By (9), for all integers $k \geq 1$,

$$I_+(z^{(k)}) \neq \emptyset. \quad (14)$$

Extracting a subsequence and re-indexing, we may assume without loss of generality that, for all natural numbers k, $I_+(z^{(1)}) = I_+(z^{(k)})$, $I_-(z^{(1)}) = I_-(z^{(k)})$, and for each $i \in I_0(z^{(1)})$,

$$\lim_{k\to\infty} \alpha_{ki} = \alpha_i. \quad (15)$$

Set $k \to \infty$ in (13), it follows from (9), (11) and (15) that

$$0 = \sum_{i\in I_+(z^{(1)})} \nabla g_i(\bar{z}) + \sum_{i\in I_0(z^{(1)})} \alpha_i \nabla g_i(\bar{z}). \quad (16)$$

By $\bar{z} \in \mathcal{F}$,

$$g_i(\bar{z}) = 0,\ \ i \in I_+(z^{(1)}) \cup I_0(z^{(1)}). \quad (17)$$

Therefore, (16) is a contradiction with Assumption 1. The contradiction proves that there exists $\sigma^* > 0$, such that if $\sigma \in (0, \sigma^*]$, then $\mathcal{G}(\sigma) \subseteq \mathcal{G}$.

Let $\tilde{x}$ be a global minimizer of problem (1) and $x_\sigma \in \mathcal{G}(\sigma)$ $(\sigma \in (0, \sigma^*])$, then

$$f(x_\sigma) = E_\sigma(x_\sigma), \quad f(\tilde{x}) = E_\sigma(\tilde{x}). \quad (18)$$

Therefore, as $f(x_\sigma) = f(\tilde{x})$, (18) implies that $E_\sigma(\tilde{x}) = E_\sigma(x_\sigma)$ and this proves that $\tilde{x}$ is a global solution to problem (3). ■

IV. Neural network model

To solve the the optimization problem (1), a neural network is presented based of the exact penalty property of $E_\sigma(x)$ as follows

$$\dot{x}(t) \in -\partial E_\sigma(x(t)), x_0 \in \mathcal{D}. \quad (19)$$

$\bar{x} \in \mathcal{D}$ is said to be an equilibrium point of system (19), if $0 \in -\partial E_\sigma(\bar{x})$. We denote by $\mathcal{E}(\sigma)$ the set of equilibrium point of (19).

Proposition 4 (see [28]): Let $\bar{x} \in \mathcal{F}$, $\bar{x} \in \mathcal{E}(\sigma)$, then $\bar{x} \in \mathcal{G}$ if f is a pseudoconvex function and $g_i, i = 1, \ldots, m$ are quasiconvex functions. Moreover, there exists a $\sigma^* > 0$, such that for all $\sigma \in (0, \sigma^*]$, if $x_\sigma \in \mathcal{E}(\sigma)$, then $x_\sigma \in \mathcal{G}$.

Proposition 5: Let $x^* \in \mathcal{G}$, then $x^* \in \mathcal{E}(\sigma)$ for all $\sigma > 0$ such that $\lambda_i^* \leq 1/\sigma$, $i \in I_0(x^*)$.

Since $\nabla f(x)$ is continuous and $\partial V(x)$ is upper semicontinuous with nonempty compact convex values, $\mathrm{cl}\mathcal{D} \setminus \mathrm{int}\mathcal{F}$ is a compact subset, there exist $L_f > 0$ and $L_V > 0$ such that, for $x \in \mathrm{cl}\mathcal{D} \setminus \mathrm{int}\mathcal{F}$, $||\nabla f(x)|| \leq L_f$ and $|\partial V(x)| = \max\{||\eta|| : \eta \in \partial V(x)\} \leq L_V$.

The following corollary shows that any equilibrium point of (19) corresponds to an optimal solution of (1) when the penalty parameter is sufficiently small.

Corollary 1: If $\sigma^* = m_g/2L_f$ and $\sigma \in (0, \sigma^*]$, then $\mathcal{E}(\sigma) \subseteq \mathcal{G}$.

Proof: Note that for all $x \in \mathcal{D} \setminus \mathcal{F}$ and $\sigma \in (0, \sigma^*]$, $\exists\ \nu_0 \in \partial V(x)$ such that

$$\begin{aligned} &|\partial E_\sigma(x)| := \min\{||\nabla f(x) + (1/\sigma)\nu|| : \nu \in \partial V(x)\} \\ &= ||\nabla f(x) + (1/\sigma)\nu_0|| \\ &\geq (1/\sigma)||\nu_0|| - ||\nabla f(x)|| \\ &\geq (1/\sigma)m_g - L_f > 0. \end{aligned}$$

Therefore, if $x \in \mathcal{E}(\sigma)$, then $x \in \mathcal{G}$. ∎

Theorem 2: Any state of (19) converges to an optimal solution of Problem (1) if $\sigma^* = m_g/2L_f, \sigma \in (0, \sigma^*]$ and $f(x)$ is a pseudoconvex function and $g_i(x),\ i \in I$ are quasiconvex functions.

Proof: By Corollary 1, $\mathcal{E}(\sigma) \subseteq \mathcal{F}$.

If $x(t) \in \mathcal{E}(\sigma)$, then $x(t)$ reaches $\mathcal{E}(\sigma)$ in finite time.

If $x(t) \notin \mathcal{E}(\sigma)$, $f(x(t)) > f(x^*)$. Since $f(x)$ is pseudoconvex, then

$$\langle \nabla f((t)), x(t) - x^* \rangle > 0. \tag{20}$$

If $x(t) \in \mathrm{bd}\mathcal{F}$, $\exists\ \alpha_i \in [0, 1],\ i \in I_0(x(t))$,

$$\dot{x}(t) = -\nabla f(x(t)) - (1/\sigma) \sum_{i \in I_0(x(t))} \alpha_i \nabla g_i(x(t)), \tag{21}$$

thus

$$\begin{aligned} (d/dt)E(t, x^*) &= -||\dot{x}(t)||^2 - \langle \nabla f(x(t)), x(t) - x^* \rangle \\ &\quad -(1/\sigma) \sum_{i \in I_0(x(t))} \alpha_i \langle \nabla g_i(x(t)), x(t) - x^* \rangle. \end{aligned} \tag{22}$$

Since $g_i(x)$ ($i \in I_0(x(t))$) are quasiconvex functions and $g_i(x(t)) \geq g_i(x^*)$, then

$$\langle \nabla g_i(x(t)), x(t) - x^* \rangle \geq 0, \quad i \in I_0(x(t)).$$

By (20 and (22), $(d/dt)E(t, x^*) \leq 0$.

If $x(t) \in \mathrm{int}\mathcal{F}$,

$$(d/dt)E(t, x^*) = -||\dot{x}(t)||^2 - \langle \nabla f(x(t)), x(t) - x^* \rangle. \tag{23}$$

By (20) and (23) $(d/dt)E(t, x^*) \leq 0$.

Therefore, the state of (19) either converges to an optimal solution of (1) in finite time or converges to an optimal solution of (1) asymptotically. ∎

Remark 1: Compared with existing results on recurrent neural networks for pseudoconvex optimization, such as [29], the contribution of the neural network model (19) lies in this applicability of deal with generalized convex constraint functions. In contrast, previous models and theoretic results are only valid for convex constraints.

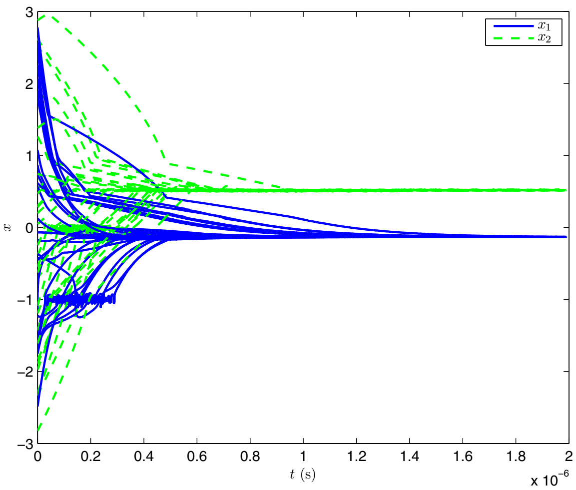

Fig. 1. Transient behaviors of the proposed neural network for Example 1

V. Simulation Results

In this section, simulation results on a numerical example are provided to illustrate the effectiveness and efficiency of the proposed recurrent neural network model (19).

Example 1: Consider an optimization problem as follows:

$$\begin{array}{ll} \min & f(x) = \frac{x_1^2 + x_1 + x_2}{x_1 + 1} \\ \text{subject to} & \ln(x_1^2 + x_2^2) \leq 0,\ \ \frac{1}{2} - x_2 + x_1^2 \leq 0. \end{array} \tag{24}$$

The objective function f is pseudoconvex, the constraint function $g_1 = \ln(x_1^2 + x_2^2)$ is quasiconvex, and the constraint function $g_2 = \frac{1}{2} - x_2 + x_1^2$ is convex. The generalized gradient ∂E_σ is computed as

$$\partial E_\sigma(x) = (\frac{x_1^2 + 2x_1 - x_2 + 1}{(x_1 + 1)^2}, \frac{1}{x_1 + 1})^T + (1/\sigma)\partial V(x),$$

where

$$\partial V(x) = \partial \max\{0, \ln(x_1^2 + x_2^2)\} + \partial \max\{0, \frac{1}{2} - x_2 + x_1^2\}.$$

The global optimal solution of (24) is $(-0.1321, 0.5175)^T$. Fig. 1 illustrates the transient behaviors of the proposed neural network from 30 random initial states. Fig. 2 shows the 2-dimensional phase plot from 100 random initial states. The simulation results show that the proposed neural network always converges to the global optimization problem (24). Moreover, the proposed neural network is capable of computing the optimal solution in micro-second scale, which is highly efficient.

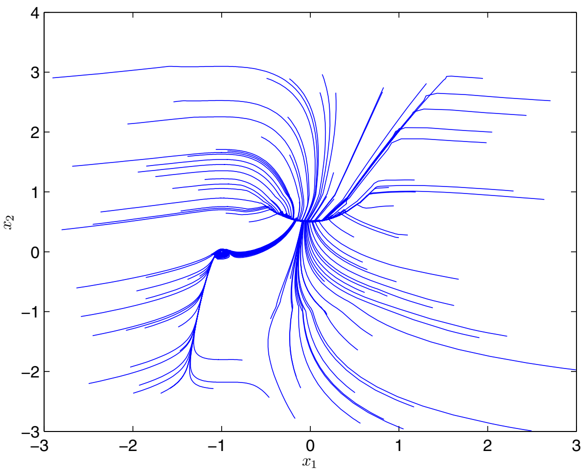

Fig. 2. Phase plot of the proposed neural network for Example 1

VI. Conclusion

This paper presents a single layer recurrent neural network for optimization problems with pseudoconvex objectives and quasiconvex inequality constraints based on an exact penalty design. The proposed neural network is proved to be convergent to the global optimal solution of the corresponding optimization problem. Simulation results are discussed to substantiate the characteristics and effectiveness of the proposed neural network. Future investigations are directed to general nonconvex optimization.

References

[1] Y. Xia and J. Wang, "A dual neural network for kinematic control of redundant robot manipulators," *IEEE Trans. Systems, Man and Cybernetics - Part B*, vol. 31, no. 1, pp. 147–154, 2001.

[2] Z. Yan and J. Wang, "Model predictive control of nonlinear systems with unmodeled dynamics based on feedforward and recurrent neural networks," *IEEE Trans. Ind. Informat.*, vol. 8, no. 4, pp. 746–756, 2012.

[3] ——, "Model predictive control of tracking of underactuated vessels based on recurrent neural networks," *IEEE J. Ocean. Eng.*, vol. 37, no. 4, pp. 717–726, 2012.

[4] Z.-G. Hou, M. M. Gupta, P. N. Nikiforuk, M. Tan, and L. Cheng, "A recurrent neural network for hierarchical control of interconnected dynamic systems," *IEEE Transactions on Neural Networks*, vol. 18, no. 2, pp. 466–481, 2007.

[5] A. Balavoine, J. Romberg, and C. J. Rozell, "Convergence and rate analysis of neural networks for sparse approximation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 9, pp. 1377–1389, 2012.

[6] J. J. Hopfield and D. W. Tank, "Neural computation of decisions in optimization problems," *Biological Cybernetics*, vol. 52, no. 3, pp. 141–152, 1985.

[7] 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 and Systems*, vol. 33, no. 5, pp. 533–541, 1986.

[8] J. Wang, "A deterministic annealing neural network for convex programming," *Neural Networks*, vol. 7, no. 4, pp. 629–641, 1994.

[9] Y. Xia, "Global convergence analysis of Lagrangian networks," *IEEE Trans. Circuits Syst.*, vol. 50, no. 6, pp. 818–822, 2003.

[10] ——, "A new neural network for solving linear and quadratic programming problems," *IEEE Transactions on Neural Networks*, vol. 7, no. 6, pp. 1544–1548, 1996.

[11] Y. Xia, G. Feng, and J. Wang, "A recurrent neural network with exponential convergence for solving convex quadratic programand related linear piecewise equations," *Neural Networks*, vol. 17, no. 7, pp. 1003–1005, 2004.

[12] S. Liu and J. Wang, "A simplified dual neural network for quadratic programming with its kwta application," *IEEE Transactions on Neural Networks*, vol. 17, no. 6, pp. 1500–1510, 2006.

[13] Y. Xia, H. Leung, and J. Wang, "A projection neural network and its application to constrained optimization problems," *IEEE Trans. Circuits and Systems - Part I*, vol. 49, no. 4, pp. 447–458, 2002.

[14] X. Gao, "A novel neural network for nonlinear convex programming," *IEEE Trans. Neural Networks*, vol. 15, no. 3, pp. 613–621, 2004.

[15] X. Hu and J. Wang, "Design of general projection neural network for solving monotone linear variational inequalities and linear and quadratic optimization problems," *IEEE Trans. Systems, Man and Cybernetics - Part B: Cybernetics*, vol. 37, no. 5, pp. 1414–1421, 2007.

[16] Q. Liu, J. Cao, and G. Chen, "A novel recurrent neural network with finite-time convergence for linear programming," *Neural Computation*, vol. 22, pp. 2962–2978, 2010.

[17] X. Hu and J. Wang, "Solving pseudomonotone variational inequalities and pseudoconvex optimization problems using the projection neural network," *IEEE Trans. Neural Networks*, vol. 17, no. 6, pp. 1487–1499, 2006.

[18] Z. Guo, Q. Liu, and J. Wang, "A one-layer recurrent neural network for pseudoconvex optimization subject to linear equality constraints," *IEEE Trans. Neural Networks*, vol. 22, no. 12, pp. 1892–1900, 2011.

[19] Q. Liu and J. Wang, "Finite-time convergent recurrent neural network with a hard-limiting activation function for constrained optimization with piecewise-linear objective functions," *IEEE Trans. Neural Networks*, vol. 22, no. 4, pp. 601–613, 2011.

[20] A. Hosseini, J. Wang, and S. M. Hosseini, "A recurrent neural network for solving a class of generalized convex optimization problems," *Neural Networks*, vol. 44, pp. 78–86, 2013.

[21] Q. Liu and J. Wang, "A one-layer projection neural network for nonsmooth optimization subject to linear equalities and bound constraints," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 5, pp. 812–824, 2013.

[22] F. Clarke, *Optimization and Non-Smooth Analysis*. New York: Wiley, 1969.

[23] J. Aubin and A. Cellina, *Differential Inclusions*. Verlin: Springer-Verlag, 1984.

[24] A. Filippov, *Differential Equations with Discontinuous Right-Hand Side*. Dordrecht: Kluwer Academic, 1988.

[25] P. Pardalos, *Nonconvex Optimization and Its Application*. Berlin Heidelberg, 2008.

[26] A. Cambini and L. Martein, *Generalized Convexity and Optimization :Theory and Applications*. Springer-Verlag Berlin Heidelberg, 2009.

[27] P. Nistri and M. Quincampoix, "On the dynamics of a differential inclusion built upon a nonconvex constrained minimization problem," *Journal of Optimization Ttheory and Applications*, vol. 124, no. 3, pp. 659–672, 2005.

[28] G. D. Pillo and L. Grippo, "Exact penalty function in constrained optimization," *SIAM J. Control and optimization*, vol. 27, no. 6, pp. 1333–1360, 1989.

[29] Q. Liu, Z. Guo, and J. Wang, "A one-layer recurrent neural network for constrained pseudoconvex optimization and its application for dynamic portfolio optimization," *Neural Networks*, vol. 26, pp. 99–109, 2012.