Kybernetika

Alaeddin Malek; Najmeh Hosseinipour-Mahani

Solving a class of non-convex quadratic problems based on generalized KKT conditions and neurodynamic optimization technique

SOLVING A CLASS OF NON-CONVEX QUADRATIC PROBLEMS BASED ON GENERALIZED KKT CONDITIONS AND NEURODYNAMIC OPTIMIZATION TECHNIQUE

Alaeddin Malek and Najmeh Hosseinipour-Mahani

In this paper, based on a generalized Karush–Kuhn–Tucker (KKT) method a modified recurrent neural network model for a class of non-convex quadratic programming problems involving a so-called Z-matrix is proposed. The basic idea is to express the optimality condition as a mixed nonlinear complementarity problem. Then one may specify conditions for guaranteeing the global solutions of the original problem by using results from the S-lemma. This process is proved by building up a dynamic system from the optimality condition whose equilibrium point is exactly the solution of the mixed nonlinear complementarity problem. By the study of the resulting dynamic system it is shown that under given assumptions, steady states of the dynamic system are stable. Numerical simulations and comparisons with the other methods are presented to illustrate the efficiency of the practical technique that is proposed in this paper.

1. INTRODUCTION

Although there are many Recurrent Neural Network (RNN) to solve the constraint/ unconstraint convex programming problems [12] – [17],[24, 25, 27] there are few non-convex problem solvers. In addition, a very broad class of difficult combinatorial optimization problems such as integer programming, Celis–Dennis–Tapia (CDT) problem, quadratic assignment and the maximum clique problem can be formulated as non-convex quadratic programming problems. However, the neural network models for convex optimization problems are not successful in solving non-convex optimization problems, hence, the study of neural network models for non-convex optimization problems, is of quite importance. In last decade, some RNN models are used to solve non-convex optimization problems. Two neural network models for unconstrained non-convex optimization problems are presented by Beyer and Ogier [2] and Sun and Feng [20]. Two neural network models for non-convex quadratic objective function subject to a set of affine constraints and a box set are presented in [5] and [21], respectively. Xia et al. [24] and Hu [7] studied two different neural network models for solving differentiable and non-convex

optimization problems. In the year 2009, a neural network model based on differential inclusion for nonsmooth and non-convex optimization problems is introduced [3]. Yan et al. [26] presented a collective neurodynamic optimization method for solving non-convex optimization problems with bound constraints. Their approach is based on the KKT conditions that is necessary condition for the optimal point in the non-convex problems. The practical technique in this paper uses both necessary and sufficient conditions to solve the specified class of non-convex optimization problems.

In this paper, we consider a non-convex quadratic optimization problem with finitely many quadratic inequality constraints involving Z-matrices, which are matrices with non-positive off diagonal elements. Then, by designing a neurodynamic system (RNN model), we establish a practical way of finding the global optimal solution while we obey the corresponding sufficient conditions. First, we establish a relationship between optimality condition of the problem and mixed nonlinear complementarity problem. Then we propose a modified RNN model to transfer the non-convex quadratic problem into a specific dynamical system of the first order differential equations along with a modification parameter which leads to the stability of its equilibrium points. In the following, everything is boiled down to the study of the resulting dynamical system.

The outline of the paper is as follows. In Section 2, we introduce some preliminary definitions, lemmas and theorems for global optimality conditions for a class of non-convex constrained quadratic optimization problems. In Section 3, a RNN model is designed based on necessary and sufficient conditions for non-convex quadratic objective function subject to quadratic constraints. Moreover, the global convergence of the proposed neural network is analyzed. In Section 4, numerical examples are given to illustrate the efficiency of the proposed model. Comparisons with two other existing models are made. As an application, we apply the proposed neural network model for solving the CDT subproblem, which arises in some trust region algorithms for nonlinear optimization problems. Finally, some concluding remarks will be drawn in Section 5.

2. PROBLEM FORMULATION

In this section, we present basic results on global optimality and Lagrange multipliers for non-convex constrained quadratic optimization problems. In what follows, $\|.\|$ denotes l_2-norm on $\mathbb{R}^n$ ($\|\boldsymbol{x}\| = (\sum_{i=1}^{n} x_i^2)^{1/2}$) and e_i denotes the column vector with a 1 in the ith coordinate and 0's elsewhere. The space of all $n \times n$ symmetric matrices is denoted by S^n. For $g : \mathbb{R}^n \to \mathbb{R}$, $\bigtriangledown g(\boldsymbol{x}) \in \mathbb{R}^n$ and $\bigtriangledown^2 g(\boldsymbol{x}) \in \mathbb{R}^{n \times n}$ stand for gradient and the Hessian of g at $\boldsymbol{x}$. For vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $\boldsymbol{x} \geq \boldsymbol{y}$ means that $x_i \geq y_i$, for $i = 1, \ldots, n$. The notation $\mathbf{A} \succeq 0$ ($\mathbf{A} \preceq 0$) shows that the matrix $\mathbf{A}$ is positive (negative) semi-definite. Furthermore, if there exists a nonzero vector $\boldsymbol{x} \in \mathbb{R}^n$ such that $\boldsymbol{x}^{\mathrm{T}} \mathbf{A} \boldsymbol{x} < 0$ then we write $\mathbf{A} \not\succeq 0$. Consider the following smooth non-convex quadratic optimization problem:

$$\min f(\boldsymbol{x}) \qquad \text{s.t.} \qquad g_i(\boldsymbol{x}) \leq 0, \qquad i = 1, \ldots, m, \tag{1}$$

where $f, g_i : \mathbb{R}^n \to \mathbb{R}$ are defined by $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathbf{A}_f\boldsymbol{x} + \mathbf{b}_f^{\mathrm{T}}\boldsymbol{x} + c_f$ and $g_i(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathbf{A}_{g_i}\boldsymbol{x} + \mathbf{b}_{g_i}^{\mathrm{T}}\boldsymbol{x} + c_{g_i}$, $\mathbf{A}_f \not\succeq 0$ and $S_0 = \{\boldsymbol{x} \in \mathbb{R}^n | g_i(\boldsymbol{x}) \leq 0\}$ is the feasible set. Define $\mathbf{H}_f, \mathbf{H}_{g_i}$ for $i = 1, \ldots, m$ by

$$\mathbf{H}_f = \left(\begin{array}{cc} \mathbf{A}_f & \mathbf{b}_f \\ \mathbf{b}_f^{\mathrm{T}} & 2c_f \end{array} \right), \quad \mathbf{H}_{g_i} = \left(\begin{array}{cc} \mathbf{A}_{g_i} & \mathbf{b}_{g_i} \\ \mathbf{b}_{g_i}^{\mathrm{T}} & 2c_{g_i} \end{array} \right). \tag{2}$$

Definition 2.1. (Malek et al. [14])   A mixed nonlinear complementarity problem (MNCP) is finding a point $\boldsymbol{x} \in \mathbb{R}^n$ such that

$$\text{MNCP}(F) \quad \begin{cases} x_i F_i(\boldsymbol{x}) = 0, \ F_i(\boldsymbol{x}) \geq 0, \ x_i \geq 0, & \forall i \in I, \\ F_i(\boldsymbol{x}) = 0, & \forall i \in N \backslash I, \end{cases} \tag{3}$$

where F is a continuously differentiable mapping from $X = \{\boldsymbol{x} \in \mathbb{R}^n | x_i \geq 0, i \in I\}$ into $\mathbb{R}^n$, $N = \{1, 2, \ldots, n\}$ and $I \subseteq N$.

Remark 2.2. A matrix $\mathbf{A} \in S^n$ is called a Z-matrix if $a_{ij} \leq 0$ for all $i \neq j$. Therefore any diagonal matrix is a Z-matrix.

The fundamental lemma which has played a key role in many areas of optimization and control theory is given as follows.

Lemma 2.3. (S-Lemma) (Polik and Terlaky [19])   Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be quadratic functions, defined by

$$f(\boldsymbol{x}) = \frac{1}{2} \boldsymbol{x}^{\mathrm{T}} \mathbf{A}_f \boldsymbol{x} + \mathbf{b}_f^{\mathrm{T}} \boldsymbol{x} + c_f, \ g(\boldsymbol{x}) = \frac{1}{2} \boldsymbol{x}^{\mathrm{T}} \mathbf{A}_g \boldsymbol{x} + \mathbf{b}_g^{\mathrm{T}} \boldsymbol{x} + c_g$$

where $\mathbf{A}_f, \mathbf{A}_g \in S^n, \mathbf{b}_f, \mathbf{b}_g \in \mathbb{R}^n$ and $c_f, c_g \in \mathbb{R}$. Suppose that there exists $\boldsymbol{x}_0 \in \mathbb{R}^n$ such that $g(\boldsymbol{x}_0) < 0$. Then the following statements are equivalent

(i) $g(\boldsymbol{x}) \leq 0 \Rightarrow f(\boldsymbol{x}) \geq 0$.

(ii) $\exists \ \lambda \geq 0, \forall \boldsymbol{x} \in \mathbb{R}^n, \ f(\boldsymbol{x}) + \lambda g(\boldsymbol{x}) \geq 0$.

Theorem 2.4. (Jeyakumar et al. [9])   Let $f_i = \frac{1}{2} \boldsymbol{x}^{\mathrm{T}} \mathbf{A}_i \boldsymbol{x} + \mathbf{b}_i^{\mathrm{T}} \boldsymbol{x} + c_i, \ i = 1, \ldots, m$. Suppose that $\mathbf{H}_{f_i}, \ i = 1 \ldots, m$ are all Z-matrices. Then, exactly one of the following two statements holds.

(i) $(\exists \ \boldsymbol{x} \in \mathbb{R}^n) \quad f_i(\boldsymbol{x}) < 0, \ i = 1, \ldots, m$.

(ii) $(\exists \ \lambda \in \mathbb{R}_+^m \backslash \{\mathbf{0}\}) \ (\forall \boldsymbol{x} \in \mathbb{R}^n) \ \sum_{i=1}^m \lambda_i f_i(\boldsymbol{x}) \geq 0$.

Proposition 2.5. (Jeyakumar et al. [8])   For general non-convex quadratic programming problem (1), let $\boldsymbol{x}^* \in S_0$. If there exists $\lambda = (\lambda_1, \ldots, \lambda_m)^{\mathrm{T}} \in \mathbb{R}_+^m \backslash \{\mathbf{0}\}$ such that the following conditions

$$\begin{cases} (a) & \mathbf{A}_f + \sum_{i=1}^m \lambda_i \mathbf{A}_{g_i} \succeq 0; \\ (b) & (\mathbf{A}_f \boldsymbol{x}^* + \mathbf{b}_f) + \sum_{i=1}^m (\lambda_i \mathbf{A}_{g_i} \boldsymbol{x}^* + \lambda_i \mathbf{b}_{g_i}) = \mathbf{0}; \\ (c) & \sum_{i=1}^m \lambda_i g_i(\boldsymbol{x}^*) = 0; \end{cases} \tag{4}$$

hold, then $\boldsymbol{x}^*$ is a global minimizer of (1).

Remark 2.6.   For the non-convex optimization problem (1) when $m = 1$ and the strict feasibility condition holds, then conditions (4) are necessary and sufficient conditions [10].

A relevant counterexample for $m > 1$: Consider the following non-convex problem [8]:

$$\min \quad f(\boldsymbol{x}) = (x_1 - 1)^2 + x_2^2 - 10x_3^2 - 6x_2$$
$$\text{s.t.} \quad g_1(\boldsymbol{x}) = x_1^2 + x_2^2 + x_3^2 - 2 \le 0,$$
$$g_2(\boldsymbol{x}) = (x_1 - 2)^2 + x_2^2 + x_3^2 - 2 \le 0.$$

For the feasible point $\boldsymbol{x}^* = (1, 1, 0)^{\mathrm{T}}$ and $\lambda^* = (1, 1)^{\mathrm{T}} > 0$, the following KKT necessary conditions hold at $\boldsymbol{x}^*$:

$$\nabla f(\boldsymbol{x}^*) + \lambda_1^* \nabla g_1(\boldsymbol{x}^*) + \lambda_2^* \nabla g_2(\boldsymbol{x}^*) = \begin{pmatrix} 0 \\ -4 \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \end{pmatrix} + \begin{pmatrix} -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

$\lambda_1^* g_1(\boldsymbol{x}^*) = 0$ and $\lambda_2^* g_2(\boldsymbol{x}^*) = 0$. Indeed $\mathbf{A}_f + \sum_{i=1}^{2} \lambda_i^* \mathbf{A}_{g_i} = \mathrm{diag}(2, 2, -20) + \mathrm{diag}(2, 2, 2) + \mathrm{diag}(2, 2, 2) = \mathrm{diag}(6, 6, -16) \not\succeq 0$. Thus the sufficient optimality conditions (4) do not hold at this point (see also Example 4.4). Hence for $m > 1$ the condition (a) of (4) is just a sufficient (not necessary) global optimality condition.

Theorem 2.7. (Jeyakumar et al. [9]) For the non-convex quadratic problem (1), suppose that $\mathbf{H}_f$ and $\mathbf{H}_{g_i}, i = 1, \ldots, m$ are Z-matrices and the Slater condition holds, i. e., there exists $\boldsymbol{x}_0 \in \mathbb{R}^n$ such that $g_i(\boldsymbol{x}_0) < 0, \ i = 1, \ldots, m$. Then a feasible point $\boldsymbol{x}^*$ is a global optimal solution if and only if (4) holds.

Theorem 2.8. For the non-convex quadratic problem (1), suppose that $\mathbf{H}_{g_i}, i = 1, \ldots, m$ are Z-matrices. In the feasible set S_0, assume that there exists a Z-matrix $\mathbf{Q} \in S^n$ such that $\mathbf{A}_f - \mathbf{Q} \preceq 0$. Also there exists $\boldsymbol{x}_0 \in \mathbb{R}^n$ such that $g_i(\boldsymbol{x}_0) < 0, \ i = 1, \ldots, m$. If $\boldsymbol{x}^*$ is a global optimal solution of (1) and $\nabla f(\boldsymbol{x}^*) - \mathbf{Q}\boldsymbol{x}^* \le \mathbf{0}$, then there exists $(\lambda_1, \ldots, \lambda_m) \in \mathbb{R}_+^m \setminus \{\mathbf{0}\}$ such that

$$\mathbf{Q} + \sum_{i=1}^{m} \lambda_i \mathbf{A}_{g_i} \succeq 0, \quad \nabla \left(f + \sum_{i=1}^{m} \lambda_i g_i \right) (\boldsymbol{x}^*) = \mathbf{0} \quad \text{and} \quad \lambda_i g_i(\boldsymbol{x}^*) = 0, \ i = 1, \ldots, m. \tag{5}$$

P r o o f. See Appendix. □

3. DYNAMICAL SYSTEM

In this section, first we reformulate the optimality conditions of the problem (1) as the mixed nonlinear complementarity problem. Then we design a neurodynamic system along with a modification parameter to guarantee the stability of the equilibrium points. Hereafter, for the non-convex quadratic problem (1) suppose that $\mathbf{H}_f$ and $\mathbf{H}_{g_i}, i = 1, \ldots, m$ are Z-matrices. Let $\boldsymbol{x}^*$ be a global optimal solution of (1) and the Slater condition holds. According to the Theorem 2.7, there exists $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)^{\mathrm{T}} \in \mathbb{R}_+^m \setminus \{\mathbf{0}\}$ such that $(\boldsymbol{x}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})^{\mathrm{T}}$ satisfies the following generalized KKT conditions for $i = 1, \ldots, m$:

$$
\begin{cases}
\mathbf{A}_f + \sum_{i=1}^m \lambda_i^* \mathbf{A}_{g_i} \succeq 0, \\
\nabla f(\boldsymbol{x}^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}, \\
\lambda_i^* \left(\frac{1}{2} \boldsymbol{x}^{*\mathrm{T}} \mathbf{A}_{g_i} \boldsymbol{x}^* + \mathbf{b}_{g_i}^{\mathrm{T}} \boldsymbol{x}^* + c_{g_i} \right) = 0, \\
-\left(\frac{1}{2} \boldsymbol{x}^{*\mathrm{T}} \mathbf{A}_{g_i} \boldsymbol{x}^* + \mathbf{b}_{g_i}^{\mathrm{T}} \boldsymbol{x}^* + c_{g_i} \right) \geq 0, \\
\lambda_i^* \geq \boldsymbol{0}.
\end{cases} \tag{6}
$$

Consider the continuously differentiable function $G : Y \to \mathbb{R}^{n+m}$ as follows:

$$
G(\boldsymbol{y}) = \begin{pmatrix} \nabla (f + \sum_{i=1}^m \lambda_i g_i)(\boldsymbol{x}) \\ -g_1(\boldsymbol{x}) \\ \vdots \\ -g_m(\boldsymbol{x}) \end{pmatrix} = \begin{pmatrix} (\mathbf{A}_f \boldsymbol{x} + \mathbf{b}_f) + \sum_{i=1}^m \lambda_i (\mathbf{A}_{g_i} \boldsymbol{x} + \mathbf{b}_{g_i}) \\ -\left(\frac{1}{2} \boldsymbol{x}^{\mathrm{T}} \mathbf{A}_{g_1} \boldsymbol{x} + \mathbf{b}_{g_1}^{\mathrm{T}} \boldsymbol{x} + c_{g_1} \right) \\ \vdots \\ -\left(\frac{1}{2} \boldsymbol{x}^{\mathrm{T}} \mathbf{A}_{g_m} \boldsymbol{x} + \mathbf{b}_{g_m}^{\mathrm{T}} \boldsymbol{x} + c_{g_m} \right) \end{pmatrix}, \tag{7}
$$

for $Y = \{ \boldsymbol{y} = (\boldsymbol{x}^{\mathrm{T}}, \lambda^{\mathrm{T}})^{\mathrm{T}} | \boldsymbol{x} \in \mathbb{R}^n, \lambda \in \mathbb{R}_+^m \}$. We show $\boldsymbol{y}^* = (\boldsymbol{x}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})^{\mathrm{T}}$ is a solution of MNCP(G). Let $N = \{1, 2, \dots, n+m\}$ and $I = \{n+1, \dots, n+m\}$, it is obvious that $y_{n+i} = \lambda_i \geq 0$, $G_{n+i}(\boldsymbol{y}) = -g_i(\boldsymbol{x}) \geq 0$, $y_{n+i} G_{n+i}(\boldsymbol{y}) = 0$, $i = 1, \dots, m$ and $G_i(\boldsymbol{y}) = \nabla (f + \sum_{i=1}^m \lambda_i g_i)(\boldsymbol{x}) = 0$, $i = 1, \dots, n$ for $\boldsymbol{y} = \boldsymbol{y}^*$. Therefore $\boldsymbol{y}^* = (\boldsymbol{x}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})^{\mathrm{T}}$ is a solution of MNCP(G).

Conversely, let $\boldsymbol{y}^* = (\boldsymbol{x}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})$ be a solution of MNCP(G) such that $\mathbf{A}_f + \sum_{i=1}^m \lambda_i^* \mathbf{A}_{g_i} \succeq 0$. From the definition of MNCP(G), we conclude that $\boldsymbol{y}^*$ satisfies the generalized KKT conditions (6). Thus $\boldsymbol{x}^*$ is a global optimal solution for problem (1) by Theorem 2.7. This proves the following proposition.

Proposition 3.1. For the non-convex quadratic problem (1), suppose that $\mathbf{H}_f$ and $\mathbf{H}_{g_i}, i = 1, \dots, m$ are Z-matrices. Then $\boldsymbol{x}^* \in S_0$ is a global solution of the problem (1), if and only if there exists $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)^{\mathrm{T}} \in \mathbb{R}_+^m \backslash \{\boldsymbol{0}\}$, such that $\mathbf{A}_f + \sum_{i=1}^m \lambda_i^* \mathbf{A}_{g_i} \succeq 0$ and $\boldsymbol{y}^* = (\boldsymbol{x}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})$ is a solution of MNCP(G), where G is defined as in (7).

3.1. Neurodynamic model (RNN model)

Now, let $\boldsymbol{x}(.)$, $\lambda(.)$ and $\boldsymbol{y}(.)$ be some time dependent variables. The aim is to construct a continuous-time dynamical system that will settle down to the global optimal solution of the problem (1). We propose a modified recurrent neural network model for solving (1), whose dynamical system for initial point $(\boldsymbol{x}_0^{\mathrm{T}}, \lambda_0^{\mathrm{T}})^{\mathrm{T}}$ is defined as follows:

$$
\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = -(\mathbf{A}_f \boldsymbol{x} + \mathbf{b}_f) - \sum_{i=1}^m \frac{(\lambda_i + (\lambda_i + \frac{1}{2} \boldsymbol{x}^{\mathrm{T}} \mathbf{A}_{g_i} \boldsymbol{x} + \mathbf{b}_{g_i}^{\mathrm{T}} \boldsymbol{x} + c_{g_i})^+)}{2} (\mathbf{A}_{g_i} \boldsymbol{x} + \mathbf{b}_{g_i}), \tag{8}
$$

$$
\frac{\mathrm{d}\lambda_i}{\mathrm{d}t} = (\lambda_i + \frac{1}{2} \boldsymbol{x}^{\mathrm{T}} \mathbf{A}_{g_i} \boldsymbol{x} + \mathbf{b}_{g_i}^{\mathrm{T}} \boldsymbol{x} + c_{g_i})^+ - \lambda_i, \qquad i = 1, \dots, m. \tag{9}
$$

Define

$$
H(\boldsymbol{y}) = \begin{pmatrix} -\nabla f(\boldsymbol{x}) - \nabla g(\boldsymbol{x})^{\mathrm{T}} \left(\frac{\lambda + (\lambda + g(\boldsymbol{x}))^+}{2} \right) \\ (\lambda + g(\boldsymbol{x}))^+ - \lambda \end{pmatrix}, \tag{10}
$$

where $g(\boldsymbol{x}) = [g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), \dots, g_m(\boldsymbol{x})]^{\mathrm{T}}$, $(\boldsymbol{z})^+ = [(z_1)^+, \dots, (z_n)^+]^{\mathrm{T}}$, $(z_i)^+ = max\{0, z_i\}$. We propose the following neurodynamic model (RNN model):

$$
\begin{cases} \frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = H(\boldsymbol{y}), \\ \boldsymbol{x}(t) = (\mathbf{I}_n, \boldsymbol{0}) \boldsymbol{y}(t) \end{cases} \tag{11}
$$

where $\boldsymbol{y}(t) = (\boldsymbol{x}(t)^{\mathrm{T}}, \lambda(t)^{\mathrm{T}})^{\mathrm{T}}$ is the state vector, $\boldsymbol{x}(t)$ is the output vector and $\left(\frac{\lambda}{2} + \frac{(\lambda+g(\boldsymbol{x}))^+}{2}\right)$ is modification parameter. Model (11) is a modification of the neural network model in Ref [6] with $(\lambda + g(\boldsymbol{x}))^+$ replaced by $\left(\frac{\lambda+(\lambda+g(\boldsymbol{x}))^+}{2}\right)$ to solve non-convex optimization problem (1). This kind of modification have a serious role in making the RNN model (11) stable.

3.2. Connection between equilibrium point and global solution

Proposition 3.2. Let Ω^* be the set of equilibrium points of the model (11) in $\mathbb{R}^{n+m}$. Then $\boldsymbol{y}^* \in \Omega^*$ if and only if $\boldsymbol{y}^* = (\boldsymbol{x}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})^{\mathrm{T}}$ is the solution for MNCP(G).

P r o o f. Let $\boldsymbol{y}^* \in \Omega^*$, since $\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t}|_{\boldsymbol{y}=\boldsymbol{y}^*} = 0$ then $H(\boldsymbol{y}^*) = 0$. It follows that

$$(\mathbf{A}_f \boldsymbol{x}^* + \mathbf{b}_f) + \sum_{i=1}^{m} \left(\frac{\lambda_i^*}{2} + \frac{(\lambda_i^* + \frac{1}{2}\boldsymbol{x}^{*\mathrm{T}}\mathbf{A}_{g_i}\boldsymbol{x}^* + \mathbf{b}_{g_i}^{\mathrm{T}}\boldsymbol{x}^* + c_{g_i})^+}{2} \right) (\mathbf{A}_{g_i}\boldsymbol{x}^* + \mathbf{b}_{g_i}) = 0, \quad (12)$$

$$(\lambda_i^* + \frac{1}{2}\boldsymbol{x}^{*\mathrm{T}}\mathbf{A}_{g_i}\boldsymbol{x}^* + \mathbf{b}_{g_i}^{\mathrm{T}}\boldsymbol{x}^* + c_{g_i})^+ - \lambda_i^* = 0, \qquad i = 1, \dots, m. \quad (13)$$

Now by substituting (13) into (12) we have

$$(\mathbf{A}_f \boldsymbol{x}^* + \mathbf{b}_f) + \sum_{i=1}^{m} \lambda_i^*(\mathbf{A}_{g_i}\boldsymbol{x} + \mathbf{b}_{g_i}) = 0 \tag{14}$$

Moreover, it is clear that $(\lambda_i^* + \frac{1}{2}\boldsymbol{x}^{*\mathrm{T}}\mathbf{A}_{g_i}\boldsymbol{x}^* + \mathbf{b}_{g_i}^{\mathrm{T}}\boldsymbol{x}^* + c_{g_i})^+ = \lambda_i^*$ for $i = 1, \dots, m$ if and only if

$$\lambda_i^* \left(\frac{1}{2}\boldsymbol{x}^{*\mathrm{T}}\mathbf{A}_{g_i}\boldsymbol{x}^* + \mathbf{b}_{g_i}^{\mathrm{T}}\boldsymbol{x}^* + c_{g_i} \right) = 0, \quad \lambda_i^* \geq 0, \quad -\left(\frac{1}{2}\boldsymbol{x}^{*\mathrm{T}}\mathbf{A}_{g_i}\boldsymbol{x}^* + \mathbf{b}_{g_i}^{\mathrm{T}}\boldsymbol{x}^* + c_{g_i} \right) \geq 0.$$
$$(15)$$

From (14) and (15) we conclude that $\boldsymbol{y}^* = (\boldsymbol{x}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})^{\mathrm{T}}$ is a solution of MNCP(G) when $N = \{1, 2, \dots, n+m\}$ and $I = \{n+1, \dots, n+m\}$. The converse is immediate. $\qquad \square$

Theorem 3.3. Let $\boldsymbol{y}^* \in \Omega^*$ and $\mathbf{A}_f + \sum_{i=1}^{m} \lambda_i^* \mathbf{A}_{g_i} \succeq 0$ for $\lambda_i^* = \mathrm{e}_{n+i}^{\mathrm{T}}\boldsymbol{y}^*$, $i = 1, \dots, m$, then $\boldsymbol{x}^* = (\mathbf{I}_n, \mathbf{0}_{n \times m})\boldsymbol{y}^*$ is a global optimal solution for (1). On the other hand, if $\mathbf{H}_f$ and $\mathbf{H}_{g_i}, i = 1, \dots, m$ are Z-matrices and $\boldsymbol{x}^*$ is a global optimal solution of (1), then there exists $\lambda^* \in \mathbb{R}^m \setminus \{\mathbf{0}\}$ such that $(\boldsymbol{x}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})^{\mathrm{T}}$ is an equilibrium point of the RNN model (11).

P r o o f. Let $\boldsymbol{y}^*$ be an equilibrium point for the modified neural network model (11) and its associated λ^* satisfies $\mathbf{A}_f + \sum_{i=1}^{m} \lambda_i^* \mathbf{A}_{g_i} \succeq 0$. It follows from Proposition 3.2 that $\boldsymbol{y}^*$ is a solution of MNCP(G) and it satisfies in the conditions (4). Hence $\boldsymbol{x}^* = (\mathbf{I}_n, \mathbf{0}_{n \times m})\boldsymbol{y}^*$ is a global optimal solution for the problem (1) by Proposition 2.5. The converse is straightforward. $\qquad \square$

3.3. Stability and convergent analysis

Definition 3.4. (Miller and Michel [18]) A continuous-time neural network model is said to be globally convergent, if for any given initial point, the corresponding trajectory of the related dynamical system converges to an equilibrium point.

Definition 3.5. (Khalil [11]) The equilibrium point $\boldsymbol{y}^*$ is Lyapunov stable if, for each $\epsilon > 0$, there is $\delta > 0$ such that if $\|\boldsymbol{y}(t_0) - \boldsymbol{y}^*\| < \delta$, then $\|\boldsymbol{y}(t) - \boldsymbol{y}^*\| < \epsilon$, for $t \geq t_0$.

Theorem 3.6. (Khalil [11]) Let $\boldsymbol{y}^*$ be an equilibrium point for (11) and $D \subset \mathbb{R}^{n+m}$ be a domain containing $\boldsymbol{y}^*$. Let $E : D \to \mathbb{R}$ be a continuously differentiable function such that

$$\begin{cases} E(\boldsymbol{y}) \geq 0, & \forall \boldsymbol{y} \in D \setminus \{\boldsymbol{y}^*\}, \\ E(\boldsymbol{y}) = 0 & \text{iff } \boldsymbol{y} = \boldsymbol{y}^*, \\ \frac{\mathrm{d}E(\boldsymbol{y}(t))}{\mathrm{d}t} \leq 0, & \forall \boldsymbol{y} \in D, \end{cases} \tag{16}$$

then $\boldsymbol{y}^*$ is stable.

A continuously differentiable function $E(\boldsymbol{y})$ satisfying (16) is called a Lyapunov function.

Lemma 3.7. For any initial point $\mathbf{y}(t_0)$ there exists a unique continuous solution $\boldsymbol{y}(t)$ for model (11).

P r o o f . It is easily to verify that the right-hand term of (11) are locally Lipschitz continuous. According to the local existence uniqueness theorem of ODEs [4], there exists a unique local continuous solution for model (11) in the interval $(t_0, t_0 + \tau)$, where $\tau > 0$. $\qquad\square$

Theorem 3.8. Let there exists $\aleph \subseteq \mathbb{R}^{n+m}$ such that for any $\boldsymbol{y} = (\boldsymbol{x}^{\mathrm{T}}, \lambda^{\mathrm{T}})^{\mathrm{T}} \in \aleph$ we have $\mathbf{A}_f + \sum_{i=1}^{m} \frac{\lambda_i}{2} \mathbf{A}_{g_i} \succeq 0$. Then the Jacobian matrix $\nabla H(\boldsymbol{y})$ of the mapping H defined in (10) is a negative semi-definite matrix for $\boldsymbol{y} \in \aleph$.

P r o o f . See Appendix. $\qquad\square$

Theorem 3.9. Let $\Omega^* \subseteq \aleph \subseteq Y$, the set $\aleph$ being as in Theorem 3.8. Then

(i) equilibrium point of the model (11) is stable in the sense of Lyapunov,

(ii) the model (11) is globally convergent to the stationary point $\boldsymbol{y}^* = (\boldsymbol{x}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})^{\mathrm{T}}$ of (11).

P r o o f . See Appendix. $\qquad\square$

4. NUMERICAL EXAMPLES

In order to demonstrate the efficiency and performance of the RNN model (11) in solving a class of non-convex quadratic optimization problems, we give several illustrative examples in this section. Codes are written in MATLAB and have done on a Core i3-380M PC (2.53GHz) Processor.

Example 4.1. (Zheng et al. [29]) (General non-convex problem that its solution satisfies the sufficient global optimality condition)
Consider

$$
\begin{aligned}
\min \ & 8x_1x_2 + 3x_2^2 + 14x_1 + 12x_2 \\
\text{s.t.} \ & 18x_1^2 + 8x_2^2 + 2x_1 - 1 \le 0, \\
& 13x_1^2 - 4x_1x_2 + 8x_2^2 + 4x_2 - 1 \le 0, \\
& 5x_1^2 - 10x_1x_2 + 5x_2^2 + 16x_1 + 18x_2 - 1 \le 0.
\end{aligned}
$$

We use the proposed model (11) for solving this problem, it is seen that the state trajectory converges to $(\boldsymbol{x}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})$ where $\boldsymbol{x}^* = (-0.2190, -0.2680)^{\mathrm{T}}$ and $\lambda^* = (2.0148, 0, 0)^{\mathrm{T}}$. We have

$$
\mathbf{A}_f + \Sigma_{i=1}^3 \lambda_i^* \mathbf{A}_{g_i} = \begin{pmatrix} 36.2664 & 4 \\ 4 & 19.1184 \end{pmatrix} \succ 0.
$$

Therefore $\boldsymbol{x}^*$ is a global optimal solution by Theorem 3.3. In Figure 1 (a) the trajectories of the RNN model (11) with ten random initial vector points are illustrated. Figure 1 (b) displays the transient behavior of $\boldsymbol{x}(t)$ with five random initial points starting from outside of the feasible region S_0.

Note that if we use an adjusted parameter κ in RNN model (11), we will have

$$
\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \kappa H(\boldsymbol{y}), \tag{17}
$$

then a sufficiently large κ could accelerate the neurodynamic. This means that, it is possible for the RNN model to converge in millisecond scale, or even micro second scale. For $\kappa = 2000$, the state trajectories of the RNN model (18) with ten random initial points are converged in 6 milisecond (see Figure 1 (c)). Comparison between Figure 1 (a) and Figure 1 (c) shows that increasing the adjustable parameter κ leads to an appropriate gain.
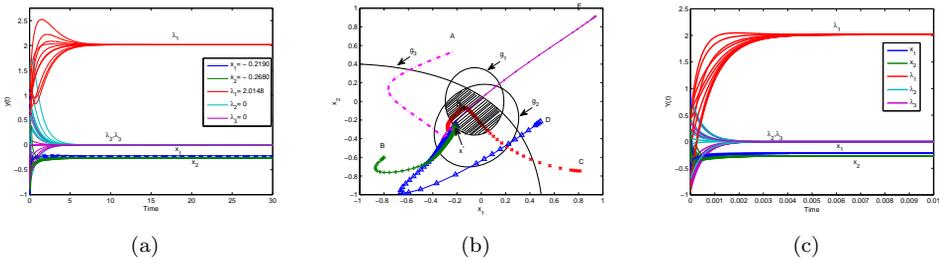

(a)　　　　(b)　　　　(c)

Fig. 1. (a) The transient behavior of the model (11) with ten different initial points, (b) The feasible region S_0 and the trajectories of the $(x_1(t), x_2(t))$ started from five different points, (c) The transient behavior of the mode (18) with $\kappa = 2000$ started from ten initial points for Example 4.1.

Example 4.2. (Tian and Lu [22]) (The non-convex problem that its solution does not satisfy the sufficient global optimality condition)

Consider the following subclass of mixed integer quadratic programming problem:

$$\min \ \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathbf{A}_f\boldsymbol{x} + \boldsymbol{b}_f^{\mathrm{T}}\boldsymbol{x}, \qquad \text{s.t.} \ \ \boldsymbol{x}_i \in [0,1], \ \ i = 1,\ldots,n \qquad (18)$$

with

$$\mathbf{A}_f = \begin{pmatrix} 263 & -97 & 62 & 217 & 52 & 621 & 935 & 258 & -61 & -10 \\ -97 & 299 & -17 & 9 & -4 & -123 & -17 & -40 & -3 & 37 \\ 62 & -17 & 178 & 71 & -118 & -83 & -110 & 9 & -56 & 42 \\ 217 & 9 & 71 & 143 & -5 & 842 & 228 & 42 & 58 & -41 \\ 52 & -4 & -118 & -5 & 177 & 102 & -15 & 120 & 13 & -52 \\ 621 & -123 & -83 & 842 & 102 & 219 & 574 & 22 & 73 & -53 \\ 935 & -17 & -110 & 228 & -15 & 574 & 457 & 154 & -25 & 84 \\ 258 & -40 & 9 & 42 & 120 & 22 & 154 & 473 & 18 & -29 \\ -61 & -3 & -56 & 58 & 13 & 73 & -25 & 18 & -4 & -79 \\ -10 & 37 & 42 & -41 & -52 & -53 & 84 & -29 & -79 & 224 \end{pmatrix},$$

$\boldsymbol{b}_f = (-20, -314, 46, -83.45, -128.7, 41.3, 43.85, -341.8, -34.05, -34.6)^{\mathrm{T}}$. In order to solve this mixed integer programming problem by the theories of this paper, we reformulate (18) by

$$\min \ \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathbf{A}_f\boldsymbol{x} + \boldsymbol{b}_f^{\mathrm{T}}\boldsymbol{x},$$
$$\text{s.t.} \ \ \boldsymbol{x}_i^2 - \boldsymbol{x}_i \leq 0, \ \ i = 1,\ldots,n.$$

This problem is solved by the RNN model (11) and the global minimizer $\boldsymbol{x}^*$ is calculated successfully. The output trajectories are converged to the global minimizer $\boldsymbol{x}^* = (0, 1.00, 0.50, 0, 0.75, 0, 0, 0.60, 1.00, 0.49)^{\mathrm{T}}$. Trajectories of the neural network model (11) with six random initial vector points are illustrated in Figure 2 (a). Figure 2 (b) displays the convergence behavior of the l_2-norm error $\|\boldsymbol{x}(t) - \boldsymbol{x}^*\|$ based on the model (11) with fifteen initial points.

Now, it is obvious that there exists $\lambda^* = (41.17, 35.15, 0, 20.05, 0, 13.14, 70.15, 0, 88.03, 0)^{\mathrm{T}}$ such that $(\boldsymbol{x}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})$ satisfies the KKT necessary condition, while $(\boldsymbol{x}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})$ does not satisfy the global optimality sufficient condition (4) (a). However, if $\mathbf{H}_f$ could be established such that satisfies in Z-matrix definition, then one must have the global optimality sufficient condition by Theorem 2.7.

In the following it is shown that non-convex problems involving Z-matrices (see Examples 4.3 and 4.4) satisfy both necessary and sufficient global optimality conditions (6).

Example 4.3. (Global solution for CDT problem) Consider the following problem

$$\min \ f(\boldsymbol{d}) = \frac{1}{2}\boldsymbol{d}^{\mathrm{T}}\mathbf{B}\boldsymbol{d} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{d},$$
$$\text{s.t.} \ \ \|\mathbf{A}^{\mathrm{T}}\boldsymbol{d} + \boldsymbol{a}\| \leq \theta, \qquad (19)$$
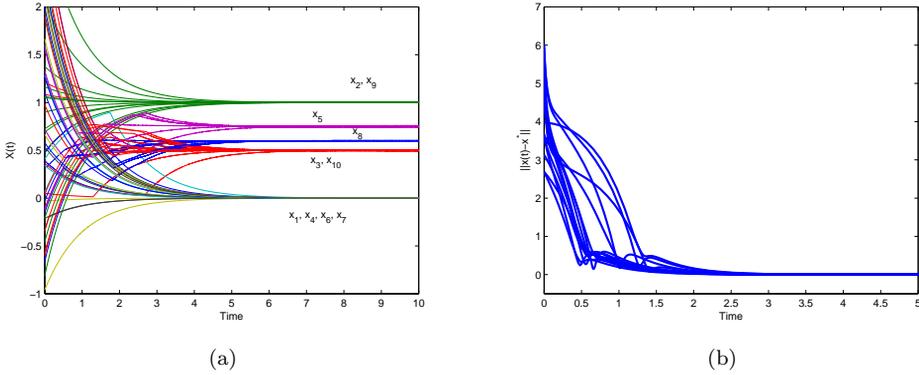$$\|\boldsymbol{d}\| \leq \delta,$$

Fig. 2. The transient behavior of the RNN model (11) with six random initial vector points, (b) Convergence behavior of the output trajectory norm error based on (11) with fifteen random initial points in Example 4.2.

where $\mathbf{B} \in S^n$, $\mathbf{A} \in \mathbb{R}^{n \times m} (m \leq n)$, $\boldsymbol{b} \in \mathbb{R}^n$, $\boldsymbol{a} \in \mathbb{R}^m$, $\theta > 0$ and $\delta > 0$. In this paper, it is assumed that $f(\boldsymbol{d})$ is non-convex quadratic function, namely, $\mathbf{B}$ may be indefinite.

The problem (19) comes from applying the successive quadratic programming (SQP) method and a trust-region technique to minimize a general function $q(\boldsymbol{x})$ subject to $h(\boldsymbol{x}) = 0$. At the kth iteration, one must calculate the correction step d_k to the current x_k by minimizing a quadratic function $\boldsymbol{b}^\mathrm{T} \boldsymbol{d} + \frac{1}{2} \boldsymbol{d}^\mathrm{T} \mathbf{B} \boldsymbol{d}$ subject to the norm of the linearized constraints $\mathbf{A}^\mathrm{T} \boldsymbol{d} + \boldsymbol{a}$, in a given tolerance where $\boldsymbol{b} = \nabla q(x_k)$, $\mathbf{B}$ is the Hessian or an approximate Hessian of the Lagrangian function with respect to $\boldsymbol{x}$, $\mathbf{A} = \nabla h(x_k)$ and $\boldsymbol{a} = h(x_k)$. Meanwhile, the trust-region restriction $\|\boldsymbol{d}\| \leq \delta$ must be impose [28].

In order to solve the CDT problem (19) by the theories of this paper, we replace (19) by

$$\min \ f(\boldsymbol{d}) = \frac{1}{2} \boldsymbol{d}^\mathrm{T} \mathbf{B} \boldsymbol{d} + \boldsymbol{b}^\mathrm{T} \boldsymbol{d},$$
$$\text{s.t.} \ \ g_1(\boldsymbol{d}) = \|\mathbf{A}^\mathrm{T} \boldsymbol{d} + \boldsymbol{a}\|^2 - \theta^2 \leq 0,$$
$$g_2(\boldsymbol{d}) = \ \|\boldsymbol{d}\|^2 - \delta^2 \leq 0.$$

For $n = m = 2$, $\mathbf{B} = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}$ and $\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\boldsymbol{a} = (0, -6)^\mathrm{T}$, $\boldsymbol{b} = (0, -6)^\mathrm{T}$, $\delta = 5$, $\theta = 5$, it is clear that $\mathbf{H}_f = \begin{pmatrix} \mathbf{B} & \boldsymbol{b} \\ \boldsymbol{b} & 0 \end{pmatrix}$, $\mathbf{H}_{g_1} = 2 \begin{pmatrix} \mathbf{A}\mathbf{A}^\mathrm{T} & \mathbf{A}\boldsymbol{a} \\ (\mathbf{A}\boldsymbol{a})^\mathrm{T} & \|\boldsymbol{a}\|^2 - \theta^2 \end{pmatrix}$ and $\mathbf{H}_{g_2} = 2 \begin{pmatrix} \mathbf{I}_n & 0 \\ 0 & -\delta^2 \end{pmatrix}$ are Z-matrices and the Slater condition holds. Let $\boldsymbol{y}^* = (\boldsymbol{d}^{*\mathrm{T}}, \lambda^{*\mathrm{T}})^\mathrm{T}$ and $\boldsymbol{y}^{**} = (\boldsymbol{d}^{**\mathrm{T}}, \lambda^{**\mathrm{T}})^\mathrm{T}$ where $\boldsymbol{d}^* = (4, 3)^\mathrm{T}$, $\boldsymbol{d}^{**} = (-4, 3)^\mathrm{T}$ and $\lambda^* = \lambda^{**} = (1, 1)^\mathrm{T}$. Simulation results show that the state trajectories of the proposed model (11) converge to $\boldsymbol{y}^*$ and $\boldsymbol{y}^{**}$ in which $\boldsymbol{y}^*$ and $\boldsymbol{y}^{**}$ satisfy the global optimality conditions in Theorem 2.7. Hence $\boldsymbol{d}^* = (4, 3)^\mathrm{T}$ and $\boldsymbol{d}^{**} = (-4, 3)^\mathrm{T}$ are two different global optimal solutions for

Model	Initial point	Solution	CPU time (sec.)	l_2-norm error
PNN	(-5,4,-3,2)	(-3.9024,3.0000)	0.33	0.09
ALN	(-5,4,-3,2)	(-3.9998,3.0000)	0.17	2.11×10^{-4}
Current	(-5,4,-3,2)	(-3.9999,3.0000)	0.05	6.28×10^{-5}
PNN	(2,2,2,2)	(3.8656,3.0000)	0.37	0.1344
ALN	(2,2,2,2)	(3.9992,3.0000)	0.14	7.92×10^{-4}
Current	(2,2,2,2)	(3.9999,3.0000)	0.04	6.2998×10^{-5}
PNN	(-5,-5,-5,-5)	(-4.3653,3.0000)	0.34	0.36
ALN	(-5,-5,-5,-5)	(-4.0009,3.0000)	0.15	8.51×10^{-4}
Current	(-5,-5,-5,5)	(-3.9999,3.0000)	0.04	9.92×10^{-5}

Tab. 1. Comparison of the results for three RNN models with
different initial points for Example 4.3.

Example 4.3. Figure 3 (a) shows that the trajectories of the neural network model (11)
with fifteen random initial vector points converge to the global optimal solutions of
this problem. Figure 3 (b) displays the transient behavior of $\boldsymbol{d}(t)$ when one may start
from points inside, outside and on the boundary of the feasible region. Furthermore, we
compare the modified neural network with the existing projection neural network (PNN)
defined in Ref. [24] and the augmented Lagrange network (ALN) algorithm defined in
Ref. [7], where adjusted parameters in PNN and ALN algorithms are assumed to be 1.
Table 1 gives their computational results under different initial points. It shows that
the modified neural network has a better performance in convergence time and solution
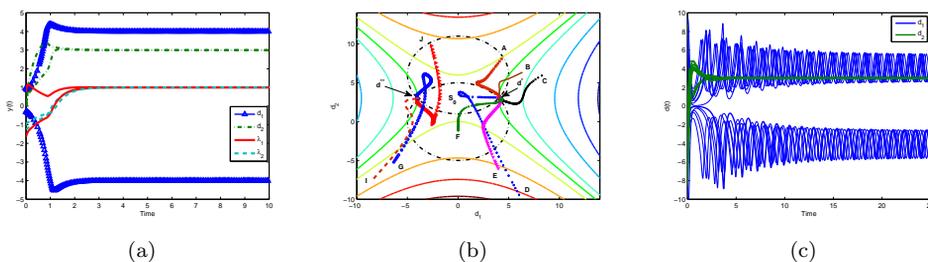accuracy than the PNN and ALN algorithms.

(a)   (b)   (c)

Fig. 3. (a) The transient behavior of the RNN model (11) with two
initial points, (b) The objective function contours, the feasible region
and the output trajectories started from different points, (c) The
transient behavior of the PNN model with fifteen initial points for
Example 4.3.

Model	Initial point	Solution	CPU time (sec.)	l_2-norm error
PNN	(2,2,2,2,2)	(5.142,0.000,0.000,0.933,0.049)	0.11	1.538
ALN	(2,2,2,2,2)	(3.608,0.000,0.000,0.990,0.054)	1.36	0.011
Current	(2,2,2,2,2)	(3.605,0.000,0.000,0.999,0.055)	0.27	2.193×10^{-4}
PNN	(-1,-1,-1,-1,-1)	(-5.811,0.000,0.946,0.048)	0.12	2.206
ALN	(-1,-1,-1,-1,-1)	(3.664,0.000,0.000,0.977,0.053)	0.28	0.063
Current	(-1,-1,-1,-1,-1)	(-3.6051,-0.000,-0.000,1.000,0.055)	0.29	1.222×10^{-4}
PNN	(2,-1,3,-2,4)	(2.458,0.000,0.000,1.028,0.056)	0.10	1.147
ALN	(2,-1,3,-2,4)	(3.602,0.000,0.000,0.985,0.054)	1.95	0.015
Current	(2,-1,3,-2,4)	(3.605,0.000,0.000,0.999,0.055)	0.28	5.547×10^{-4}

Tab. 2. Comparison of the results for three different RNN models
with different initial points for Example 4.4.

Example 4.4. Consider the following non-convex programming problem:

$$\min \quad -3x_1^2 + x_2^2 + \frac{3}{2}x_3^2 + 2x_4^2 + 3x_5^2$$

$$\text{s.t.} \quad \frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 - 14) \le 0,$$

$$\frac{1}{4}(x_1^2 + x_2^2 + x_3^2 + (x_4 - 3)^2 + x_5^2 - 17) \le 0,$$

$$-x_2 x_3 - 0.5x_3^2 - 1.5x_4 + x_5^2 - 2.5 \le 0,$$

$$-2x_2 x_3 + 0.5x_4^2 - 9x_5 \le 0, \quad -x_2 x_3 - 9x_5 \le 0.$$

This problem is solved using the proposed model (11). Simulation results show that the state trajectories of the model (11) converge to $\boldsymbol{x}^* = (3.6051, 0.0000, 0.0000, 1.0000, 0.0556)^{\mathrm{T}}$ and $\boldsymbol{x}^{**} = (-3.6051, 0.0000, 0.0000, 1.0000, 0.0556)^{\mathrm{T}}$ with $\lambda^* = \lambda^{**} = (5.2840, 6.7160, 0.0000, 0.0741, 0.0000)^{\mathrm{T}}$. Then for λ^* we have $\mathbf{A}_f + \Sigma_{i=1}^5 \lambda_i^* \mathbf{A}_{g_i} \succeq 0$ and since $\mathbf{H}_f$ and $\mathbf{H}_{g_i}, i = 1, \ldots, m$ are Z-matrices, thus we can conclude by Theorem 3.3 that $\boldsymbol{x}^*$ and $\boldsymbol{x}^{**}$ are two different global solutions of this problem. Trajectories of the neural network model (11) with ten random initial vector points, are shown in Figure 4 (a). The RNN model (11), the PNN and ALN algorithms are compared in Table 2. Simulation results show that the trajectory of the PNN algorithm does not converge to the global solutions, whereas the RNN (11) calculates converge to $\boldsymbol{x}^*$ and $\boldsymbol{x}^{**}$ as global solutions.

Note: The ALN algorithm converges while the PNN algorithm does not converge, when dealing with the non-convex quadratic problem in Examples 4.3 and 4.4. The Xia approach (PNN algorithm) provides unstable solution (see Figures 3 (c) and 4 (b)), while the Hu approach (ALN algorithm) provides the equilibrium points which satisfy the global optimality conditions in Theorem 2.7, hence the ALN algorithm converges to the global optimal solutions of Examples 4.3 and 4.4.
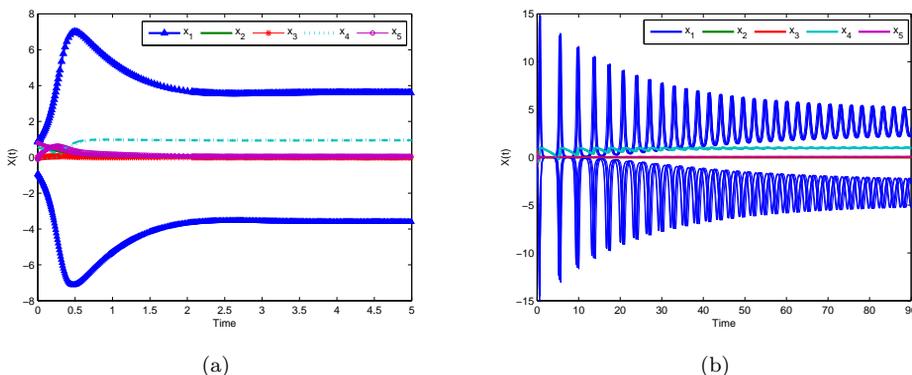
Fig. 4. (a) The transient behavior of the model (11) with two
random initial points, (b) The transient behavior of the PNN model
with ten various initial points for Example 4.4.

Practical neurodynamic optimization technique: Theories that is proposed in this paper may be applied to the PNN algorithm, in order to make this algorithm stable. This may be done by

Step 1: Replacing the given non-convex quadratic problem by a problem of the type:

$$\min \quad \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathbf{Q}\boldsymbol{x}$$
$$\text{s.t.} \quad \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathbf{A}_{g_i}\boldsymbol{x} + \mathbf{b}_{g_i}^{\mathrm{T}}\boldsymbol{x} + c_{g_i} \leq 0, \quad i = 1, \ldots, m, \tag{20}$$

where $\mathbf{Q}$ is a Z-matrix and $\mathbf{A}_f - \mathbf{Q} \preceq 0$.

Step 2: Solve (20) by the PNN algorithm and find $\boldsymbol{x}^*$ as a global optimal solution.

Step 3: If $\nabla f(\boldsymbol{x}^*) - \mathbf{Q}\boldsymbol{x}^* \leq \mathbf{0}$ and there exists $\lambda \geq \mathbf{0}$ such that $(\boldsymbol{x}^{*\mathrm{T}}, \lambda^{\mathrm{T}})^{\mathrm{T}}$ satisfies in condition (5), Go to Step 4. Otherwise, Go to Step 1.

Step 4: If $\boldsymbol{x}^*$ satisfies (4) (a), then the necessary and sufficient condition holds, Stop. Otherwise, Go to Step 1.

This practical technique is explained in the following example.

Example 4.5. Consider the CDT problem in Example 4.3. For Z-matrix $\mathbf{Q} = \mathrm{diag}(-2, 2)$, solve (20) using the PNN algorithm. This algorithm calculates $\boldsymbol{d}_1 = (4, 3)^{\mathrm{T}}$ and $\boldsymbol{d}_2 = (-4, 3)^{\mathrm{T}}$. It is easy to check that $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ satisfy (5) with this choice of $\mathbf{Q}$ and $\lambda = (1, 1)^{\mathrm{T}}$. Moreover $\nabla f(\boldsymbol{d}) - \mathbf{Q}\boldsymbol{d} = (0, -6)^{\mathrm{T}} \leq \mathbf{0}$ for $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$. Thus by Theorem 2.7, $\boldsymbol{d}_1$ and $\boldsymbol{d}_2$ are global optimal solutions. This gives a practical way of finding global minimizers for Example 4.3 when one uses the PNN algorithm.

Example 4.6. In the similar way we propose $\mathbf{Q} = \mathrm{diag}(-5, 2, 3, 8, 6)$ and obtain $\lambda = (5.2840, 6.7160, 0.0000, 0.0741, 0.0000)^{\mathrm{T}}$ for the PNN algorithm in Example 4.4, in order to find stable global optimal solutions as those are given by current method.

5. CONCLUSION

In this paper, we have established a connection between the global optimality conditions of a class of non-convex quadratic optimization problems and mixed nonlinear complementarity problems. Based on this connection, we have proposed a neurodynamic model with a modification parameter which leads to the stability condition for RNN equilibrium points. It is proved that steady states of the dynamic system is exactly the solution of MNCP problem. Then by using results from S-lemma, we have shown certain conditions for guaranteeing the global solutions. Furthermore, we obtained necessary condition for global minimizers of a class of non-convex quadratic programming problems involving Z-matrices. We have given a practical neurodynamic optimization technique to determine global solutions for some algorithms that fail to solve some non-convex problems.

FUTURE WORK

The proposed RNN model in this paper is based on the KKT conditions, which is generally not the sufficient for global optimality. The proposed RNN model can converge to a KKT point from any initial state, but this solution may not be the global optimal solution. Idea here to compute the global optimal solution from a KKT point is based on the sufficient condition that is related to the matrices $\mathbf{A}_f + \sum_{i=1}^{m} \lambda_i^* \mathbf{A}_{g_i}$, $\mathbf{H}_f$ and $\mathbf{H}_{g_i}$ for $i = 1, \ldots, m$. If matrix $\mathbf{A}_f + \sum_{i=1}^{m} \lambda_i^* \mathbf{A}_{g_i}$ is a positive semi-definite matrix or $\mathbf{H}_f$ and $\mathbf{H}_{g_i}$ for $i = 1, \ldots, m$ are Z-matrices, then the candidate KKT point is a global optimal solution. This idea is highlighted in Propositions 3.1 and 3.2 and Theorem 3.3. Any idea to further research for the global optimal solution from a KKT point which $\mathbf{A}_f + \sum_{i=1}^{m} \lambda_i^* \mathbf{A}_{g_i}$ is not positive semi-definite, may be considered by using underestimators and/or overestimators for the objective function. This is the plan for the next investigation of the authors.

APPENDIX

P r o o f. o f T h e o r e m 2.8. Suppose that $\boldsymbol{x}^*$ is a global minimizer of (1). Let $h(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathbf{Q}\boldsymbol{x} + (\nabla f(\boldsymbol{x}^*) - \mathbf{Q}\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{x}$. Define $\phi(\boldsymbol{x}) = f(\boldsymbol{x}) - h(\boldsymbol{x})$, $\boldsymbol{x} \in S_0$. Then $\nabla^2\phi(\boldsymbol{x}) = \nabla^2 f(\boldsymbol{x}) - \nabla^2 h(\boldsymbol{x}) = \mathbf{A}_f - \mathbf{Q} \preceq 0$, $\forall \boldsymbol{x} \in S_0$. So ϕ is a concave function over S_0. Moreover $\nabla\phi(\boldsymbol{x}^*) = \nabla f(\boldsymbol{x}^*) - \nabla h(\boldsymbol{x}^*) = \mathbf{0}$. Hence, by concavity, $\phi(\boldsymbol{x}) \leq \phi(\boldsymbol{x}^*)$, $\forall \boldsymbol{x} \in S_0$. Thus $f(\boldsymbol{x}) - f(\boldsymbol{x}^*) \leq h(\boldsymbol{x}) - h(\boldsymbol{x}^*)$, $\forall \boldsymbol{x} \in S_0$. Since $f(\boldsymbol{x}) - f(\boldsymbol{x}^*) \geq 0$ hence $h(\boldsymbol{x}) - h(\boldsymbol{x}^*) \geq 0$. Let $\tilde{h}(\boldsymbol{x}) = h(\boldsymbol{x}) - h(\boldsymbol{x}^*)$ and

$$\mathbf{H}_{\tilde{h}} = \begin{pmatrix} \mathbf{Q} & \nabla f(\boldsymbol{x}^*) - \mathbf{Q}\boldsymbol{x}^* \\ (\nabla f(\boldsymbol{x}^*) - \mathbf{Q}\boldsymbol{x}^*)^{\mathrm{T}} & \boldsymbol{x}^{*\mathrm{T}}\mathbf{Q}\boldsymbol{x}^* - 2\nabla f(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{x}^* \end{pmatrix}.$$

Note that $\mathbf{H}_{\tilde{h}}$ is a Z-matrix if and only if $\mathbf{Q}$ is a Z-matrix and $\nabla f(\boldsymbol{x}^*) - \mathbf{Q}\boldsymbol{x}^* \leq \mathbf{0}$. Then the system $\tilde{h}(\boldsymbol{x}) < 0$, $g_i(\boldsymbol{x}) < 0$, $i = 1, \ldots, m$, has no solution. Thus, from Theorem 2.4, there exists $(\mu_0, \ldots, \mu_m) \in \mathbb{R}_+^{m+1} \setminus \{\mathbf{0}\}$ such that for all $\boldsymbol{x} \in S_0$

$$\mu_0 \tilde{h}(\boldsymbol{x}) + \sum_{i=1}^{m} \mu_i g_i(\boldsymbol{x}) = \mu_0(h(\boldsymbol{x}) - h(\boldsymbol{x}^*)) + \sum_{i=1}^{m} \mu_i g_i(\boldsymbol{x}) \geq 0.$$

In particular, one has $\mu_i g_i(\boldsymbol{x}^*) = 0$, $i = 1, \ldots, m$. Thus, $\mu_0 h(\boldsymbol{x}) + \sum_{i=1}^m \mu_i g_i(\boldsymbol{x})$ attains its minimum at $\boldsymbol{x}^*$ over S_0. Thus $\mu_0 \nabla f(\boldsymbol{x}^*) + \sum_{i=1}^m \mu_i \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0}$ and $\mu_0 \mathbf{Q} + \sum_{i=1}^m \mu_i \mathbf{A}_{g_i} \succeq 0$. We next, show that $\mu_0 > 0$. Suppose, contrary to our claim, that $\mu_0 = 0$, then $\sum_{i=1}^m \mu_i g_i(\boldsymbol{x}) \geq 0$ for all $x \in S_0$. Note that $g_i(\boldsymbol{x}_0) < 0$, $i = 1, \ldots, m$ under the hypothesis that the Slater condition holds. It follows that $\mu_i = 0$, $i = 1, \ldots, m$. This contradicts the fact that $(\mu_0, \ldots, \mu_m) \neq (0, \ldots, 0)$. Hence $h(\boldsymbol{x}) + \sum_{i=1}^m \lambda_i g_i(\boldsymbol{x}) \geq h(\boldsymbol{x}^*)$, where $\lambda_i = \mu_i/\mu_0$. This implies that $\lambda_i g_i(\boldsymbol{x}^*) = 0$ for $i = 1, \ldots, m$. Therefore, $\boldsymbol{x}^*$ is a global minimizer of $h(\boldsymbol{x}) + \sum_{i=1}^m \lambda_i g_i(\boldsymbol{x})$ over S_0. This gives us that

$$\nabla f(\boldsymbol{x}^*) + \sum_{i=1}^m \lambda_i \nabla g_i(\boldsymbol{x}^*) = \boldsymbol{0} \quad \text{and} \quad \mathbf{Q} + \sum_{i=1}^m \lambda_i \mathbf{A}_{g_i} \succeq 0. \tag{21}$$

$\square$

Proof. of Theorem 3.8. For $\boldsymbol{y} \in \aleph$ consider the following three cases.

Case 1. Without loss of generality, assume that there exists $0 < p < m$ such that

$$(\lambda + g(\boldsymbol{x}))^+ = (\lambda_1 + g_1(\boldsymbol{x}), \ldots, \lambda_p + g_p(\boldsymbol{x}), \underbrace{0, 0, \ldots, 0}_{m-p})^{\mathrm{T}} \geq \boldsymbol{0}.$$

An easy computation shows that the Jacobian matrix of H is

$$\begin{pmatrix} -\xi & -\nabla g_1(\boldsymbol{x}) & \cdots & -\nabla g_p(\boldsymbol{x}) & -\frac{1}{2}\nabla g_{p+1}(\boldsymbol{x}) & \cdots & -\frac{1}{2}\nabla g_m(\boldsymbol{x}) \\ \nabla g_1(\boldsymbol{x})^{\mathrm{T}} & 0 & \ldots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots \\ \nabla g_p(\boldsymbol{x})^{\mathrm{T}} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & -1 \end{pmatrix},$$

where

$$\xi = \mathbf{A}_f + \sum_{i=1}^p \left(\frac{\lambda_i + (\lambda_i + g_i(\boldsymbol{x}))}{2} \right) \mathbf{A}_{g_i} + \frac{1}{2} \sum_{i=1}^p \nabla g_i(\boldsymbol{x}) \nabla g_i(\boldsymbol{x})^{\mathrm{T}} + \sum_{j=p+1}^m \frac{\lambda_j}{2} \mathbf{A}_{g_j}.$$

It is clear that $\boldsymbol{x}^{\mathrm{T}}(\nabla g_i(\boldsymbol{x}) \nabla g_i(\boldsymbol{x})^{\mathrm{T}})\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}(\mathbf{A}_{g_i}\boldsymbol{x} + \mathbf{b}_{g_i})(\mathbf{A}_{g_i}\boldsymbol{x} + \mathbf{b}_{g_i})^{\mathrm{T}}\boldsymbol{x} = (\boldsymbol{x}^{\mathrm{T}}\mathbf{A}_{g_i}\boldsymbol{x} + \mathbf{b}_{g_i}^{\mathrm{T}}\boldsymbol{x})^2 \geq 0$, thus $\nabla g_i(\boldsymbol{x}) \nabla g_i(\boldsymbol{x})^{\mathrm{T}}$ is a positive semi-definite matrix. Moreover, since $\mathbf{A}_f + \sum_{i=1}^m \frac{\lambda_i}{2} \mathbf{A}_{g_i}$ and $\mathbf{A}_{g_i}$, $i = 1, \ldots, m$ are positive semi-definite matrices and $\lambda_i + g_i(\boldsymbol{x}) \geq 0$ for $i = 1, \ldots, p$, we see that

$$\mathbf{A}_f + \sum_{i=1}^p \left(\frac{\lambda_i + (\lambda_i + g_i(\boldsymbol{x}))}{2} \right) \mathbf{A}_{g_i} + \sum_{j=p+1}^m \frac{\lambda_j}{2} \mathbf{A}_{g_j}$$

$$= (\mathbf{A}_f + \sum_{i=1}^m \frac{\lambda_i}{2} \mathbf{A}_{g_i}) + \sum_{i=1}^p \frac{\lambda_i + g_i(\boldsymbol{x})}{2} \mathbf{A}_{g_i} \succeq 0.$$

Therefore $-\xi \preceq 0$ and we conclude that $\nabla H(\boldsymbol{y})$ is negative semi-definite.

Case 2. If $p = m$, i.e. $(\lambda + g(\boldsymbol{x}))^{+} = (\lambda_1 + g_1(\boldsymbol{x}), \lambda_2 + g_2(\boldsymbol{x}), \ldots, \lambda_m + g_m(\boldsymbol{x}))^{\mathrm{T}}$, then we obtain $\nabla H(\boldsymbol{y})$ is

$$
\begin{pmatrix}
-\mathbf{A}_f - \sum_{i=1}^{m} \alpha_i \mathbf{A}_{g_i} - \frac{1}{2}\sum_{i=1}^{m} \nabla g_i(\boldsymbol{x})\nabla g_i(\boldsymbol{x})^{\mathrm{T}} & -\nabla g_1(\boldsymbol{x}) & \cdots & -\nabla g_m(\boldsymbol{x}) \\
\nabla g_1(\boldsymbol{x})^{\mathrm{T}} & 0 & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots \\
\nabla g_m(\boldsymbol{x})^{\mathrm{T}} & 0 & \cdots & 0
\end{pmatrix},
$$

where $\alpha_i = \frac{\lambda_i + (\lambda_i + g_i(\boldsymbol{x}))}{2}$. Similar to the previous case, it is easily proved that $\nabla H(\boldsymbol{y})$ is a negative semi-definite.

Case 3. Finally if $p = 0$, i.e. $(\lambda + g(\boldsymbol{x}))^{+} = (\underbrace{0, 0, \ldots, 0}_{m})$, then

$$
\nabla H(\boldsymbol{y}) =
\begin{pmatrix}
-\mathbf{A}_f - \sum_{i=1}^{m} \frac{\lambda_i}{2} \mathbf{A}_{g_i} & -\frac{1}{2}(\mathbf{A}_{g_1}\boldsymbol{x} + \mathbf{b}_{g_1}) & \cdots & -\frac{1}{2}(\mathbf{A}_{g_m}\boldsymbol{x} + \mathbf{b}_{g_m}) \\
\mathbf{0}_{1 \times n} & -1 & \cdots & 0 \\
\vdots & \vdots & \cdots & \vdots \\
\mathbf{0}_{1 \times n} & 0 & \cdots & -1
\end{pmatrix}.
$$

In this case, the Jacobian matrix $\nabla H(\boldsymbol{y})$ is a negative semi-definite since $\mathbf{A}_f + \sum_{i=1}^{m} \frac{\lambda_i}{2} \mathbf{A}_{g_i}$ is positive semi-definite for $\boldsymbol{y} \in \aleph$. $\square$

Proof. of Theorem 3.9. (i) We first prove that the equilibrium point of the model (11) is stable in the sense of Lyapunov. Let $\boldsymbol{y}(t)$ be the trajectory of the model (11) with arbitrary initial point $\boldsymbol{y}(t_0)$ and $\tilde{\boldsymbol{y}}$ as an equilibrium point of (11). Consider the Lyapunov function $E : \aleph \to \mathbb{R}$ as

$$
E(\boldsymbol{y}) = \|H(\boldsymbol{y})\|^2 + \beta\|\boldsymbol{y} - \tilde{\boldsymbol{y}}\|^2, \tag{22}
$$

where β is a positive constant. By the definition of $H(\boldsymbol{y})$ it is seen that

$$
\frac{\mathrm{d}H(\boldsymbol{y})}{\mathrm{d}t} = \frac{\partial H(\boldsymbol{y})}{\partial \boldsymbol{y}}\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \nabla H(\boldsymbol{y})H(\boldsymbol{y}). \tag{23}
$$

Calculating the derivative of $E(\boldsymbol{y}(t))$ along the trajectory of the proposed neural network (11), we have

$$
\frac{\mathrm{d}E(\boldsymbol{y}(t))}{\mathrm{d}t} = \left(\frac{\mathrm{d}H(\boldsymbol{y})}{\mathrm{d}t}\right)^{\mathrm{T}} H(\boldsymbol{y}) + H(\boldsymbol{y})^{\mathrm{T}}\left(\frac{\mathrm{d}H(\boldsymbol{y})}{\mathrm{d}t}\right) + 2\beta(\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathrm{T}}\frac{\mathrm{d}\boldsymbol{y}(t)}{\mathrm{d}t}
$$
$$
= H(\boldsymbol{y})^{\mathrm{T}}(\nabla H(\boldsymbol{y})^{\mathrm{T}} + \nabla H(\boldsymbol{y}))H(\boldsymbol{y}) + 2\beta(\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathrm{T}}H(\boldsymbol{y}). \tag{24}
$$

Theorem 3.8 now shows that $\nabla H(\boldsymbol{y}) \preceq 0$ for $\boldsymbol{y} \in \aleph$. Thus, we have

$$
\frac{\mathrm{d}E(\boldsymbol{y}(t))}{\mathrm{d}t} \le 2\beta(\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathrm{T}}H(\boldsymbol{y}). \tag{25}
$$

Now, according to the mean value theorem, there exists $\hat{\boldsymbol{y}} = \gamma \boldsymbol{y} + (1 - \gamma)\tilde{\boldsymbol{y}}$, where $0 < \gamma < 1$, such that

$$H(\boldsymbol{y}) - H(\tilde{\boldsymbol{y}}) = \nabla H(\hat{\boldsymbol{y}})(\boldsymbol{y} - \tilde{\boldsymbol{y}}).$$

Since $\boldsymbol{y}, \tilde{\boldsymbol{y}} \in \aleph$ then $\hat{\boldsymbol{y}} \in \aleph$ and consequently $\nabla H(\hat{\boldsymbol{y}}) \preceq 0$. By multiplying both sides of the above equation by $(\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathrm{T}}$, we have

$$(\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathrm{T}}(H(\boldsymbol{y}) - H(\tilde{\boldsymbol{y}})) = (\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathrm{T}}\nabla H(\hat{\boldsymbol{y}})(\boldsymbol{y} - \tilde{\boldsymbol{y}}) \leq 0. \tag{26}$$

Hence

$$(\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathrm{T}}(H(\boldsymbol{y}) - H(\tilde{\boldsymbol{y}})) = (\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathrm{T}}H(\boldsymbol{y}) \leq 0. \tag{27}$$

We conclude from (25) and (27) that

$$\frac{\mathrm{d}E(\boldsymbol{y}(t))}{\mathrm{d}t} \leq 0. \tag{28}$$

It follows from Theorem 3.6 that $\tilde{\boldsymbol{y}}$ is stable in the sense of Lyapunov.

(ii) It remains to prove that the proposed model (11) is globally convergent. Since $\beta\|\boldsymbol{y} - \tilde{\boldsymbol{y}}\|^2 \leq E(\boldsymbol{y})$, then $\boldsymbol{y}$ is bounded for $t \geq t_0$. It follows that for any initial point $\boldsymbol{y}(t_0) \in Y$, there exists a convergent subsequence $\{\boldsymbol{y}(t_k)\}_{k=1}^{+\infty}$ such that, $\lim_{k \to +\infty} \boldsymbol{y}(t_k) = \boldsymbol{y}^*$. Thus $\boldsymbol{y}^*$ is a positive limit point of $\boldsymbol{y}(t)$. Let L^+ denote the set of all positive limit point of $\boldsymbol{y}(t)$. Since L^+ is a compact and invariant set [11], we have $\frac{\mathrm{d}E(\boldsymbol{y})}{\mathrm{d}t} = 0$ on L^+ [11]. But $\frac{\mathrm{d}E(\boldsymbol{y})}{\mathrm{d}t} = 0$ only at equilibrium points. This means that for $\boldsymbol{y} \in Y$

$$\frac{\mathrm{d}E(\boldsymbol{y})}{\mathrm{d}t} = 0 \text{ if and only if } \frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = 0. \tag{29}$$

Clearly, if $\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = 0$, then $\frac{\mathrm{d}E(\boldsymbol{y}(t))}{\mathrm{d}t} = \left(\frac{\mathrm{d}E(\boldsymbol{y})}{\mathrm{d}\boldsymbol{y}}\right)^{\mathrm{T}}\left(\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t}\right) = 0$. Conversely, let $\frac{\mathrm{d}E(\boldsymbol{y}(t))}{\mathrm{d}t} = 0$. We conclude from (25) that $(\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathrm{T}}H(\boldsymbol{y}) \geq 0$ and from (27) that $(\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathrm{T}}H(\boldsymbol{y}) \leq 0$, thus $(\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathrm{T}}H(\boldsymbol{y}) = 0$ which leads to $\boldsymbol{y} = \tilde{\boldsymbol{y}}$ or $H(\boldsymbol{y}) = 0$. In either case, it is clear that $\boldsymbol{y}$ is an equilibrium point of the (11) and hence $\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = 0$. Therefore from (29), $L^+ \subseteq \Omega^*$ and consequently $\boldsymbol{y}^* \in \Omega^*$.

Now we are in the stage to define the following Lyapunov function that works for the new task.

$$\bar{E}(\boldsymbol{y}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{y}^*\|^2, \tag{30}$$

Then $\bar{E}(\boldsymbol{y})$ is continuously differentiable and $\bar{E}(\boldsymbol{y}^*) = 0$. Therefore we have

$$\lim_{k \to +\infty} \bar{E}(\boldsymbol{y}(t_k)) = \bar{E}(\boldsymbol{y}^*).$$

So, $\forall \epsilon \geq 0$ there exists $q \geq 0$ such that for all $t \geq t_q$, we have $\bar{E}(\boldsymbol{y}(t)) \leq \epsilon$. Similarly, we can obtain $\frac{\mathrm{d}\bar{E}(\boldsymbol{y}(t))}{\mathrm{d}t} \leq 0$. It follows that for $t \geq t_q$

$$\frac{1}{2}\|\boldsymbol{y}(t) - \boldsymbol{y}^*\|^2 \leq \bar{E}(\boldsymbol{y}(t)) \leq \epsilon.$$

Thus $\lim_{t \to +\infty} \|\boldsymbol{y}(t) - \boldsymbol{y}^*\| = 0$ and $\lim_{t \to +\infty} \boldsymbol{y}(t) = \boldsymbol{y}^*$. Therefore the proposed neural network model (11) is globally convergent to an equilibrium point $\boldsymbol{y}^* = (\boldsymbol{x}^{*\mathrm{T}}, \lambda^*)^{\mathrm{T}}$. $\qquad\square$

REFERENCES

[1] M. S. Bazaraa and C. M. Shetty: Nonlinear Programming Theory and Algorithms. Wiley and Sons, New York 1990.

[2] D. Beyer and R. Ogier: Tabu learning: A neural network search method for solving nonconvex optimization problems. IEEE Int. Joint Conf. Neural Networks *2* (2000), 953–961.

[3] W. Bian and X. Xue: Subgradient-based neural networks for nonsmooth nonconvex optimization problems. IEEE Trans. Neural Networks *20* (2009), 6, 1024–1038. DOI:10.1109/tnn.2009.2016340

[4] C. Chicone: Ordinary Differential Equations with Applications. Second edition. Springer–Verlag, New York 2006.

[5] M. Forti, P. Nistri, and M. Quincampoix: Convergence of neural networks for programming problems via a nonsmooth Lojasiewicz inequality. IEEE Trans. Neural Networka *17* (2006), 6, 1471–1486. DOI:10.1109/tnn.2006.879775

[6] X. B. Gao: A novel neural network for nonlinear convex programming problems. IEEE Trans. Neural Network *15* (2004), 613–621. DOI:10.1109/tnn.2004.824425

[7] X. Hu: Neurodynamic optimization: Towards nonconvexity. In: Recurrent Neural Networks (X. Hu and P. Balasubramaniam, ed.), IN-TECH, 2008, pp. 289–308. DOI:10.5772/5551

[8] V. Jeyakumar, A. M. Rubinov, and Z. Y. Wu: Non-convex quadratic minimization problems with quadratic constraints: global optimality conditions. Math. Program., Ser. A *110* (2007), 521–541. DOI:10.1007/s10107-006-0012-5

[9] V. Jeyakumar, G. M. Lee, and G. Y. Li: Alternative theorems for quadratic inequality systems and global quadratic optimization. SIAM J. Optim *20* (2009), 2, 983–1001. DOI:10.1137/080736090

[10] V. Jeyakumar and S. Srisatkunarajah: Lagrange multiplier necessary condition for global optimality for non-convex minimization over a quadratic constraint via S-lemma. Optim. Lett. *3* (2009), 23–33. DOI:10.1007/s11590-008-0088-3

[11] H. K. Khalil: Nonlinear Systems. Third edition. Prentice Hall, 2002.

[12] A. Malek: Application of recurrent neural networks to optimization problems. In: Recurrent Neural Networks (X. Hu and P. Balasubramaniam, eds.), IN-TECH, 2008, pp. 255–288. DOI:10.5772/5556

[13] A. Malek and M. Alipour: Numerical solution for linear and quadratic programming problems using a recurrent neural network. Appl. Math. Comput *192* (2007), 27–39. DOI:10.1016/j.amc.2007.02.149

[14] A. Malek, N. Hosseinipour-Mahani, and S. Ezazipour: Efficient recurrent neural network model for the solution of general nonlinear optimization problems. Optimization Methods and Software *25* (2010), 489–506. DOI:10.1080/10556780902856743

[15] A. Malek, S. Ezazipour, and N. Hosseinipour-Mahani: Double projection neural network for solving pseudomonotone variational inequalities. Fixed Point Theory *12* (2011), 2, 401–418.

[16] A. Malek, S. Ezazipour, and N. Hosseinipour-Mahani: Projected dynamical systems and optimization problems. Bull. Iranian Math. Soc. *37* (2011), 2, 81–96.

[17] A. Malek and A. Yari: Primal-dual solution for the linear programming problem using neural network. Appl. Math. Comput. *169* (2005), 198–211. DOI:10.1016/j.amc.2004.06.081

[18] R. K. Miller and A. N. Michel: Ordinary Differential Equations. Academic Press, 1982. DOI:10.1016/b978-0-12-497280-3.50008-6

[19] I. Polik and T. Terlaky: A survey of the S-Lemma. SIAM Rev. *49* (2007), 371–418. DOI:10.1137/s003614450444614x

[20] C. Y. Sun and C. B. Feng: Neural networks for nonconvex nonlinear programming problems: A switching control approach. In: Lecture Notes in Computer Science *3495*, Springer–Verlag, Berlin 2005, pp. 694–699. DOI:10.1007/11427391_111

[21] Q. Tao, X. Liu, and M. S. Xue: A dynamic genetic algorithm based on continuous neural networks for a kind of non-convex optimization problems. Appl. Math. Comput. *150* (2004), 811–820. DOI:10.1016/s0096-3003(03)00309-6

[22] Y. Tian and Ch. Lu: Nonconvex quadratic reformulations and solvable conditions for mixed integer quadratic programming problems J. Industr. Managment Optim. *7* (2011), 1027–1039. DOI:10.3934/jimo.2011.7.1027

[23] Y. Xia, G. Feng, and J. Wang: A recurrent neural network with exponential convergence for solving convex quadratic program and related linear piecewise equation. Neural Networks *17* (2004), 1003–1015. DOI:10.1016/j.neunet.2004.05.006

[24] Y. S. Xia, G. Feng, and J. Wang: A novel recurrent neural network for solving nonlinear optimization problems with inequality constraints. IEEE Trans. Neural Networks *19* (2008), 1340–1353. DOI:10.1109/tnn.2008.2000273

[25] X. Xue and W. Bian: A project neural network for solving degenerate convex quadratic program. Neurocomputing *70* (2007), 2449–2459. DOI:10.1016/j.neucom.2006.10.038

[26] Z. Yan, J. Wang, and G. Li: A collective neurodynamic optimization approach to bound-constrained nonconvex optimization. Neural networks *55* (2014), 20–29. DOI:10.1016/j.neunet.2014.03.006

[27] M. Yashtini and A. Malek: Solving complementarity and variational inequalities problems using neural networks. Appl. Math. Comput. *190* (2007), 216–230. DOI:10.1016/j.amc.2007.01.036

[28] Y. Zhang: Computing a Celis–Dennis–Tapia trust-region step for equality constrained optimization. Math. Programming *55* (1992), 109–124. DOI:10.1007/bf01581194

[29] X. J. Zheng, X. L. Sun, D. Li, and Y. F. Xu: On zero duality gap in nonconvex quadratic programming problems. Global Optim. *52* (2012), 229–242. DOI:10.1007/s10898-011-9660-y

Alaeddin Malek, Department of Applied Mathematics, Faculty of Mathematical Sciences, Tarbiat Modares University, P.O. Box 14115-134, Tehran. Iran.
e-mail: mala@modares.ac.ir

Najmeh Hosseinipour-Mahani, Department of Applied Mathematics, Faculty of Mathematical Sciences, Tarbiat Modares University, Tehran. Iran.
e-mail: n.mahani@modares.ac.ir