Correspondence

A Compact Cooperative Recurrent Neural Network for Computing General Constrained L_1 Norm Estimators

Youshen Xia

***Abstract*—Recently, cooperative recurrent neural networks for solving three linearly constrained L_1 estimation problems were developed and applied to linear signal and image models under non-Gaussian noise environments. For wide applications, this paper proposes a compact cooperative recurrent neural network (CRNN) for calculating general constrained L_1 norm estimators. It is shown that the proposed CRNN converges globally to the constrained L_1 norm estimator without any condition. The proposed CRNN includes three existing CRNNs as its special cases. Unlike the three existing CRNNs, the proposed CRNN is easily applied and can deal with the nonlinear elliptical sphere constraint. Moreover, when computing the general constrained L_1 norm estimator, the proposed CRNN has a fast convergence speed due to low computational complexity. Simulation results confirm further the good performance of the proposed CRNN.**

***Index Terms*—Compact recurrent neural networks, constrained LAD estimation, elliptical sphere constraint, general linear constraints.**

I. Introduction

We are concerned with general constrained least absolute deviation (LAD) problems:

$$\begin{aligned} \min \quad & \|\mathbf{Dx}-\mathbf{d}\|_1 \\ \text{s.t.} \quad & \mathbf{Bx}=\mathbf{b}, \quad \mathbf{Ax}\le \mathbf{p} \\ & \mathbf{x}\in\Omega \end{aligned} \tag{1}$$

where $\|\mathbf{d}\|_1=\sum_{i=1}^m|d_i|$, constrained set Ω may be a box set $\{\mathbf{x}\in R^n\,|\,\mathbf{l}\le\mathbf{x}\le\mathbf{h}\}$ or an elliptical sphere set $\{\mathbf{x}\in R^n\,|\,\mathbf{x}^TQ\mathbf{x}\le\alpha\}$, $\mathbf{x}\in R^n, \mathbf{d}\in R^m, \mathbf{b}\in R^l$, and $\mathbf{p}\in R^r, \mathbf{D}\in R^{m\times n}, \mathbf{B}\in R^{l\times n}, \mathbf{A}\in R^{r\times n}$, and $Q\in R^{n\times n}$ is a positive semidefinite matrix. The optimal solution of constrained LAD problem (1) is called the constrained L_1 norm estimator. The considered constrained LAD problem is of major importance in the fields of linear signal and image processing, regression estimation, and system identification [1]–[4]. This is because the LAD estimation method is more robust and efficient than the least squares estimation method in the presence of outlier non-Gaussian noise errors [5]. The considered constrained LAD problem is general since it includes the unconstrained LAD problem and several constrained LAD problems discussed as its special cases. When $\Omega=R^n$, (1) becomes a linearly constrained LAD problem studied in [11]. When there are only box constraints, (1) is one constrained LAD problem studied in [9]

$$\begin{aligned} \min \quad & \|\mathbf{Dx}-\mathbf{d}\|_1 \\ \text{s.t.} \quad & \mathbf{l}\le\mathbf{x}\le\mathbf{h}. \end{aligned} \tag{2}$$

When $\mathbf{D}=[\mathbf{H},-I]\in R^{m\times(n+m)}, B=[1,\ldots,1]\in R^{1\times n}, \mathbf{b}=1, A=O, \mathbf{p}=0$, and $\Omega=R^n\times\{\mathbf{z}\in R^m\,|\,\hat{\mathbf{l}}\le\mathbf{z}\le\hat{\mathbf{h}}\}$ where I is an unit matrix, (1) reduces to another constrained LAD problem studied in [10]

$$\begin{aligned} \min \quad & \|\mathbf{Hx}-\mathbf{z}-\mathbf{d}\|_1 \\ \text{s.t.} \quad & \mathbf{Bx}=1, \quad \hat{\mathbf{l}}\le\mathbf{z}\le\hat{\mathbf{h}}. \end{aligned} \tag{3}$$

Conventional numerical methods for solving linear L1 problems include the linear programming (LP) method [6], the approximate method based on the Huber M-estimator [7], and the disciplined convex programming algorithm [8]. Because of the inherent nature of parallel and distributed information processing in neural networks, recurrent neural networks are promising computational models for real-time applications [12]. Recurrent neural networks for constrained optimization problems have been widely explored for control, prediction, and signal and image processing [9]–[11] and [13]. They are theoretically analyzed to be globally convergent under various conditions. In order to deal with degeneracy problems arisen in the constrained LAD estimation methods, recently three cooperative recurrent neural networks (CRNNs) [9]–[11] were developed and they can be used respectively to solve the linearly constrained LAD problem without constrained set Ω and the constrained LAD problems in (2) and (3). Although all the three existing CRNNs are shown to have good performance in several different applications, they cannot solve (1), specially in the case of the elliptical sphere constraint which is nonlinear. Moreover, the existing CRNN [11] cannot be used directly to solve (1) with linear box constraints. When applied to (1), the CRNN [11] must add its size greatly and thus result in implementation cost expensive and a slow convergence rate. However, in many applications such as signal and image processing, considering box constraints and sphere constraints in estimation problems is very desirable for better solutions [9]–[16].

In order to overcome disadvantages of the existing CRNNs, we propose a compact CRNN for solving the general constrained LAD problem (1). It is shown that the proposed CRNN converge globally to the constrained L_1 norm estimator of (1) without any condition. The proposed CRNN is compact and easily applied. More importantly, the proposed CRNN has a good performance that its computational complexity does not depend on both the linear box constraints and the nonlinear sphere constraint. In contrast, when applied to (1) with the box constraints, the existing CRNN [11] has to transform the liner box constraints such that its size increases greatly and thus will result in high computational complexity and slow convergence. Therefore, the proposed CRNN can significantly extend the existing CRNNs [9]–[11].

II. A Compact Cooperative Recurrent Neural Network

In this section, we drive a compact cooperative recurrent neural network model and proves its global convergence.

A. *Cooperative Recurrent Neural Network Model*

First, the following proposition shows that (1) can be converted into a system of four equations.

Proposition 1: $\mathbf{x}^*\in R^n$ is an optimal solution to (1) if and only if there exist $(\mathbf{y}^*,\mathbf{z}_I^*,\mathbf{z}_{II}^*)\in R^m\times R^l\times R^r$ such that $(\mathbf{x}^*,\mathbf{y}^*,\mathbf{z}_I^*,\mathbf{z}_{II}^*)$ satisfies the following four equations

$$\begin{cases} \mathbf{Bx}=\mathbf{b}, \\ \mathbf{y}=g_{X_1}(\mathbf{y}+\mathbf{Dx}-\mathbf{d}) \\ \mathbf{z}_{II}=g_{X_2}(\mathbf{z}_{II}+\mathbf{Ax}-\mathbf{p}) \\ \mathbf{x}=g_\Omega(\mathbf{x}-\mathbf{D}^T\mathbf{y}-\mathbf{B}^T\mathbf{z}_I-\mathbf{A}^T\mathbf{z}_{II}). \end{cases} \tag{4}$$

Manuscript received January 07, 2009; accepted March 31, 2009. First published April 21, 2009; current version published August 12, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Danilo P. Mandic. This work was supported by the National Natural Science Foundation of China under Grant 60875085 and the National Natural Science Foundation of Fujian Province of China under Grant 2008J0019.

The author is with College of Mathematics and Computer Science, Fuzhou University, 350108, China (e-mail: ysxia2001@yahoo.com; ysxia@fzu.edu.cn).

Digital Object Identifier 10.1109/TSP.2009.2021499

where $\mathbf{x} \in R^n, \mathbf{y} \in R^m, \mathbf{z}_I \in R^l$, $\mathbf{z}_{II} \in R^r, X_2 = \{\mathbf{z}_{II} \in R^r \mid \mathbf{z}_{II} \geq 0\}, g_{X_2}(\mathbf{z}) = [g_{X_2}(z_1), \ldots, g_{X_2}(z_r)]^T, g_{X_2}(z_i) = \max\{0, z_i\}, X_1 = \{\mathbf{y} \in R^m \mid \max_i |y_i| \leq 1\}, g_{X_1}(\mathbf{y}) = [g_{X_1}(y_1), \ldots, g_{X_1}(y_m)]^T, g_{X_1}(y_i) = \min\{\max\{y_i, -1\}, 1\}$ and when Ω is an elliptical sphere set,

$$g_\Omega(x) = \begin{cases} x & x^TQx \leq \alpha \\ \sqrt{\frac{\alpha}{x^TQx}}x & x^TQx > \alpha \end{cases}; \tag{5}$$

when Ω is a box set, $g_\Omega(\mathbf{x}) = [g_\Omega(x_1), \ldots, g_\Omega(x_n)]^T$ and

$$g_\Omega(x_i) = \begin{cases} l_i & x_i < l_i \\ x_i & l_i \leq x_i \leq h_i \\ h_i & x_i > h_i. \end{cases} \tag{6}$$

Proof: Consider the Lagrange function of (1) as

$$L(\mathbf{x}, \mathbf{y}, \mathbf{z}_I, \mathbf{z}_{II}) = \mathbf{y}^T(\mathbf{D}\mathbf{x} - \mathbf{d}) + \mathbf{z}_I^T(\mathbf{B}\mathbf{x} - \mathbf{b}) + \mathbf{z}_{II}^T(\mathbf{A}\mathbf{x} - \mathbf{p}) \tag{7}$$

where $(\mathbf{x}, \mathbf{y}, \mathbf{z}_I, \mathbf{z}_{II}) \in \Omega \times X_1 \times R^l \times X_2$, $(\mathbf{y}, \mathbf{z}_I, \mathbf{z}_{II})$ is the Lagrange multiplier vector. Similar to the discussion in paper [11], by the well-known saddle theorem, we know that $\mathbf{x}^*$ is an optimal solution of (1) if and only if there exists a Lagrange multiplier vector $(\mathbf{y}^*, \mathbf{z}_I^*, \mathbf{z}_{II}^*)$ such that $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}_I^*, \mathbf{z}_{II}^*)$ is a saddle point satisfying for any $\mathbf{x} \in \Omega, \mathbf{y} \in X_1, \mathbf{z}_I \in R^l, \mathbf{z}_{II} \in X_2$

$$L(\mathbf{x}^*, \mathbf{y}, \mathbf{z}_I, \mathbf{z}_{II}) \leq L(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}_I^*, \mathbf{z}_{II}^*) \leq L(\mathbf{x}, \mathbf{y}^*, \mathbf{z}_I^*, \mathbf{z}_{II}^*). \tag{8}$$

From the right inequality of (8) we have for any $\mathbf{x} \in \Omega$

$$\begin{aligned} &(\mathbf{y}^*)^T(\mathbf{D}\mathbf{x}^* - \mathbf{d}) + (\mathbf{z}_I^*)^T(\mathbf{B}\mathbf{x}^* - \mathbf{b}) \\ &\quad + (\mathbf{z}_{II}^*)^T(\mathbf{A}\mathbf{x}^* - \mathbf{p}) \\ &\quad \leq (\mathbf{y}^*)^T(\mathbf{D}\mathbf{x} - \mathbf{d}) + (\mathbf{z}_I^*)^T(\mathbf{B}\mathbf{x} - \mathbf{b}) \\ &\quad + (\mathbf{z}_{II}^*)^T(\mathbf{A}\mathbf{x} - \mathbf{p}). \end{aligned}$$

Note that $B\mathbf{x}^* = \mathbf{b}$ and $(\mathbf{z}_I^*)^T(\mathbf{B}\mathbf{x} - \mathbf{b}) = (\mathbf{z}_I^*)^T\mathbf{B}(\mathbf{x} - \mathbf{x}^*)$. We have

$$(\mathbf{x} - \mathbf{x}^*)^T(\mathbf{D}^T\mathbf{y}^* + \mathbf{B}^T\mathbf{z}_I^* + \mathbf{A}^T\mathbf{z}_{II}^*) \geq 0, \qquad \forall \mathbf{x} \in \Omega.$$

From the projection Theorem [18] it follows that

$$\mathbf{x}^* = g_\Omega(\mathbf{x}^* - (\mathbf{D}^T\mathbf{y}^* + \mathbf{B}^T\mathbf{z}_I^* + \mathbf{A}^T\mathbf{z}_{II}^*)).$$

Equation (4) is thus obtained.

By combining the four equations of (4) together we now propose a compact CRNN for (1) as follows:

State equation

$$\begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= \lambda\{E(t) - \mathbf{D}^TF_2(t) - \mathbf{B}^T(\mathbf{B}\mathbf{x}(t) - \mathbf{b}) \\ &\quad - \mathbf{A}^TF_3(t)\} \end{aligned} \tag{9a}$$

$$\frac{d\mathbf{y}(t)}{dt} = \lambda\{\mathbf{D}E(t) + F_2(t)\} \tag{9b}$$

$$\frac{d\mathbf{z}_I(t)}{dt} = \lambda\{\mathbf{B}E(t) + \mathbf{B}\mathbf{x}(t) - \mathbf{b}\} \tag{9c}$$

$$\frac{d\mathbf{z}_{II}(t)}{dt} = \lambda\{\mathbf{A}E(t) + F_3(t)\} \tag{9d}$$

Output equation

$$\mathbf{w}(t) = g_\Omega(\mathbf{x}(t))$$

where $\lambda > 0$ is a designing constant, $E(t) = g_\Omega(\mathbf{x}(t) - \mathbf{D}^T\mathbf{y}(t) - \mathbf{B}^T\mathbf{z}_I(t) - \mathbf{A}^T\mathbf{z}_{II}(t)) - \mathbf{x}(t), F_2(t) = g_{X_1}(\mathbf{y}(t) + \mathbf{D}\mathbf{x}(t) - \mathbf{d}) - \mathbf{y}(t)$, and $F_3(t) = g_{X_2}(\mathbf{z}_{II}(t) + \mathbf{A}\mathbf{x}(t) - \mathbf{p}) - \mathbf{z}_{II}(t)$ are three modular functions, $(\mathbf{x}(t), \mathbf{y}(t), \mathbf{z}_I(t), \mathbf{z}_{II}(t)) \in R^n \times R^m \times R^l \times R^r$ is the state vector trajectory of four neural state models (9a)–(9d) and $\mathbf{w}(t)$ is the output vector of the proposed CRNN. It can be seen that the proposed CRNN requires $4mn + 2n^2 + 4ln + 2rn$ multiplication operations per iteration. Moreover, the network size of the proposed CRNN is $m + n + l + r$.

B. Comparison

There are three CRNNs for solving constrained LAD problems in recent literatures. The first CRNN [9] is used to solve the constrained LAD problem (2). The second CRNN [10] solve the constrained LAD problem (3). The third CRNN [11] solve the linearly constrained LAD problem without the constrained set Ω. Note that linear and inequality constraints vanish and Ω is a box set, the proposed CRNN become the CRNN [9] and the CRNN [10], respectively. And note that the proposed CRNN will reduce to the third CRNN [11], where $E(t) = -\mathbf{D}^T\mathbf{y}(t) - \mathbf{B}^T\mathbf{z}_I(t) - \mathbf{A}^T\mathbf{z}_{II}(t)$. Thus, the proposed CRNN includes the three CRNNs as its special cases. On the other side, both the first and second CRNNs cannot be used to solve (1). The third CRNN cannot be directly used to solve (1) since it needs transforming the box constraints into inequality constraints. Moreover, when solving (1) by the third CRNN, $2n$ inequality constraints have to be added. More exactly, the third CRNN state model is then given by

$$\begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= \lambda\Big\{F_1(t) - \mathbf{D}^TF_2(t) - \mathbf{B}^T(\mathbf{B}\mathbf{x}(t) - \mathbf{b}) \\ &\quad -\mathbf{A}^TF_3^{(1)}(t) + F_3^{(2)}(t) - F_3^{(3)}(t)\Big\} \end{aligned} \tag{10a}$$

$$\frac{d\mathbf{y}(t)}{dt} = \lambda\{\mathbf{D}F_1(t) + F_2(t)\} \tag{10b}$$

$$\frac{d\mathbf{z}_I(t)}{dt} = \lambda\{\mathbf{B}F_1(t) + \mathbf{B}\mathbf{x}(t) - \mathbf{b}\} \tag{10c}$$

$$\frac{d\mathbf{z}_{II}^{(1)}(t)}{dt} = \lambda\Big\{\mathbf{A}F_1(t) + F_3^{(1)}(t)\Big\} \tag{10d}$$

$$\frac{d\mathbf{z}_{II}^{(2)}(t)}{dt} = \lambda\Big\{F_1(t) + F_3^{(2)}(t)\Big\} \tag{10e}$$

$$\frac{d\mathbf{z}_{II}^{(3)}(t)}{dt} = \lambda\Big\{-F_1(t) + F_3^{(3)}(t)\Big\} \tag{10f}$$

where $(\mathbf{x}(t), \mathbf{y}(t), \mathbf{z}_I(t), \mathbf{z}_{II}^{(1)}(t), \mathbf{z}_{II}^{(2)}(t), \mathbf{z}_{II}^{(3)}(t)) \in R^n \times R^m \times R^l \times R^r \times R^n \times R^n$ is the state vector trajectory of six models (10a)–(10f), $F_1(t) = -\mathbf{D}^T\mathbf{y}(t) - \mathbf{B}^T\mathbf{z}_I(t) - \mathbf{A}^T\mathbf{z}_{II}^{(1)}(t) + \mathbf{z}_{II}^{(2)}(t) - \mathbf{z}_{II}^{(3)}(t), F_2(t)$ is defined in (9), $F_3^{(1)}(t) = g_{X_2}(\mathbf{z}_{II}^{(1)}(t) + \mathbf{A}\mathbf{x}(t) - \mathbf{p}) - \mathbf{z}_{II}^{(1)}(t), F_3^{(2)}(t) = g_{\hat{X}_2}(\mathbf{z}_{II}^{(2)}(t) + \mathbf{x}(t) - \mathbf{h}) - \mathbf{z}_{II}^{(2)}(t), F_3^{(3)}(t) = g_{\hat{X}_2}(\mathbf{z}_{II}^{(3)}(t) - \mathbf{x}(t) + \mathbf{l}) - \mathbf{z}_{II}^{(3)}(t)$, and $\hat{X}_2 = \{\mathbf{z}_{II} \in R^n \mid \mathbf{z}_{II} \geq 0\}$. It can be seen that the third CRNN model requires $4mn + 6n^2 + 4ln + 2rn$ multiplication operations per iteration and size of its state variables is $m + 3n + l + r$. Thus, the third CRNN has very higher computational complexity than the proposed CRNN when $n > \max\{m, l, r\}$. Moreover, all the three CRNNs cannot be used to deal with the nonlinear constraint. Therefore, the proposed CRNN has good performance in dealing with the nonlinear constraint and speeding convergence.

TABLE I
SIMULATION RESULTS OF TWO ALGORITHMS WITH THREE SAME INITIAL POINTS

Methods	Initial point	CPU	Norm error
Existing CRNN (10)	$\mathbf{0} \in R^{29}$	1.032	0.0357
Proposed CRNN (9)	$\mathbf{0} \in R^{11}$	0.875	0.0356
Existing CRNN (10)	$-50\mathbf{e} \in R^{29}$	1.016	0.0357
Proposed CRNN(9)	$-50\mathbf{e} \in R^{11}$	0.797	0.0356
Existing CRNN (10)	$50\mathbf{e} \in R^{29}$	1.25	0.0356
Proposed CRNN (9)	$50\mathbf{e} \in R^{11}$	0.766	0.0356

C. Global Convergence

Lemma 1 [18]: Let $\Omega \subset R^n$ be a closed and convex set. For any $\mathbf{v} \in R^n, \mathbf{v}' \in \Omega$

$$(\mathbf{v} - g_\Omega(\mathbf{v}))^T (g_\Omega(\mathbf{v}) - \mathbf{v}') \geq 0.$$

Lemma 2: Assume that $(\mathbf{x}(t), \mathbf{y}(t), \mathbf{z}_I(t), \mathbf{z}_{II}(t))$ is the state trajectory of (9) with any given initial point. Then

$$\begin{aligned}
&(\mathbf{x}(t) - \mathbf{x}^*)^T E(t) + (\mathbf{y}(t) - \mathbf{y}^*)^T \mathbf{D} E(t)\\
&\quad + (\mathbf{z}_I(t) - \mathbf{z}_I^*)^T \mathbf{B} E(t)(\mathbf{z}_{II}(t) - \mathbf{z}_{II}^*)^T \mathbf{A} E(t)\\
&\quad \leq -\|E(t)\|_2^2 - (\mathbf{x}(\mathbf{t}) - \mathbf{x}^*)^T \mathbf{D}^T (\mathbf{y}(t) - \mathbf{y}^*)\\
&\qquad - (\mathbf{x}(\mathbf{t}) - \mathbf{x}^*)^T \mathbf{B}^T (\mathbf{z}_I(t) - \mathbf{z}_I^*)\\
&\qquad - (\mathbf{x}(\mathbf{t}) - \mathbf{x}^*)^T \mathbf{A}^T (\mathbf{z}_{II}(t) - \mathbf{z}_{II}^*) \qquad (11)\\
&(\mathbf{x}(t) - \mathbf{x}^*)^T \mathbf{D}^T F_2(t) + (\mathbf{y}(t) - \mathbf{y}^*)^T (-F_2(t))\\
&\quad \geq \|F_2(t)\|_2^2 - (\mathbf{x}(t) - \mathbf{x}^*)^T \mathbf{D}^T (\mathbf{y}(t) - \mathbf{y}^*) \qquad (12)\\
&(\mathbf{x}(t) - \mathbf{x}^*)^T \mathbf{A}^T F_3(t) + (\mathbf{z}_{II}(t) - \mathbf{z}_{II}^*)^T (-F_3(t))\\
&\quad \geq \|F_3(t)\|_2^2 - (\mathbf{x}(t) - \mathbf{x}^*)^T \mathbf{A}^T (\mathbf{z}_{II}(t) - \mathbf{z}_{II}^*) \qquad (13)
\end{aligned}$$

where $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}_I^*, \mathbf{z}_{II}^*)$ is the solution of (4) and $\|\cdot\|_2$ is l_2 norm.

Proof: Because both inequality (12) and inequality (13) can be obtained directly from analysis of the paper [11], we prove inequality (11) only. Let $\mathbf{v} = \mathbf{x}(t) - (\mathbf{D}^T \mathbf{y}(\mathbf{t}) + \mathbf{B}^T \mathbf{z}_I(t) + \mathbf{A}^T \mathbf{z}_{II}(t))$ and $\mathbf{v}' = \mathbf{x}^*$ in the first projection inequality of Lemma 1, we have

$$\begin{aligned}
&[\mathbf{x}(t) - (\mathbf{D}^T \mathbf{y}(t) + \mathbf{B}^T \mathbf{z}_I(t) + \mathbf{A}^T \mathbf{z}_{II}(t)) - g_\Omega(\mathbf{x}(t)\\
&\quad - (\mathbf{D}^T \mathbf{y}(t) + \mathbf{B}^T \mathbf{z}_I(t) + \mathbf{A}^T \mathbf{z}_{II}(t))]^T\\
&\quad \times [g_\Omega(\mathbf{x}(t) - (\mathbf{D}^T \mathbf{y}(t) + \mathbf{B}^T \mathbf{z}_I(t) + \mathbf{A}^T \mathbf{z}_{II}(t)) - \mathbf{x}^*] \geq 0. \qquad (14)
\end{aligned}$$

On the other side, since $g_\Omega(\mathbf{x}^* - (\mathbf{D}^T \mathbf{y}^* + \mathbf{B}^T \mathbf{z}_I^* + \mathbf{A}^T \mathbf{z}_{II}^*)) = \mathbf{x}^*$, from the projection theorem [18] it follows that

$$\begin{aligned}
&(g_\Omega(\mathbf{x}(t) - (\mathbf{D}^T \mathbf{y}(t) + \mathbf{B}^T \mathbf{z}_I(t) + \mathbf{A}^T \mathbf{z}_{II}(t)) - \mathbf{x}^*)^T\\
&\qquad \times (\mathbf{D}^T \mathbf{y}^* + \mathbf{B}^T \mathbf{z}_I^* + \mathbf{A}^T \mathbf{z}_{II}^*) \geq 0. \qquad (15)
\end{aligned}$$

Adding (14) and (15), we get

$$\begin{aligned}
&[E(t) + \mathbf{x}(t) - \mathbf{x}^*]^T [-E(t) - (\mathbf{D}^T (\mathbf{y}(t) - \mathbf{y}^*)\\
&\qquad + \mathbf{B}^T (\mathbf{z}_I(t) - \mathbf{z}_I^*) + \mathbf{A}^T (\mathbf{z}_{II}(t) - \mathbf{z}_{II}^*))] \geq 0
\end{aligned}$$

where $E(t) = g_\Omega(\mathbf{x}(t) - (\mathbf{D}^T \mathbf{y}(t) + \mathbf{B}^T \mathbf{z}_I(t) + \mathbf{A}^T \mathbf{z}_{II}(t)) - \mathbf{x}(t)$. Then

$$\begin{aligned}
&-\|E(t)\|^2 + E(t)^T ((\mathbf{D}^T (\mathbf{y}(t) - \mathbf{y}^*)\\
&\quad + \mathbf{B}^T (\mathbf{z}_I(t) - \mathbf{z}_I^*)\\
&\quad + \mathbf{A}^T (\mathbf{z}_{II}(t) - \mathbf{z}_{II}^*))) \geq (\mathbf{x}(t) - \mathbf{x}^*)^T E(t)\\
&\quad + [\mathbf{x}(t) - \mathbf{x}^*]^T [\mathbf{D}^T (\mathbf{y}(t) - \mathbf{y}^*)\\
&\quad + \mathbf{B}^T (\mathbf{z}_I(t) - \mathbf{z}_I^*) + \mathbf{A}^T (\mathbf{z}_{II}(t) - \mathbf{z}_{II}^*)].
\end{aligned}$$

It follows (11).

Theorem 1: The proposed CRNN defined in (9) is globally convergent to an optimal solution of the general constrained LAD problem (1).

Proof: First, let $\mathbf{u}(t) = (\mathbf{x}(t), \mathbf{y}(t), \mathbf{z}_I(t), \mathbf{z}_{II}(t))$ be the state trajectory of (9) with any given initial point. Consider the standard Lyapunov function

$$V(\mathbf{u}) = \frac{1}{2}\|\mathbf{u} - \mathbf{u}^*\|_2^2 \qquad (16)$$

where $\mathbf{u} = (\mathbf{x}, \mathbf{y}, \mathbf{z}_I, \mathbf{z}_{II})$ and $\mathbf{u}^* = (\mathbf{x}^*, \mathbf{y}^*, \mathbf{z}_I^*, \mathbf{z}_{II}^*)$ is a solution of (4). Then

$$\begin{aligned}
&\frac{dV(\mathbf{u}(t))}{dt}\\
&= (\mathbf{u}(t) - \mathbf{u}^*)^T \frac{d\mathbf{u}}{dt}\\
&= \lambda\{(\mathbf{x}(t) - \mathbf{x}^*)^T (E(t) - \mathbf{B}^T (\mathbf{B}\mathbf{x}(t) - \mathbf{b}))\\
&\quad + (\mathbf{x}(t) - \mathbf{x}^*)^T (-\mathbf{D}^T F_2(t)\\
&\quad - \mathbf{A}^T F_3(t)) + (\mathbf{y}(t) - \mathbf{y}^*)^T (\mathbf{D} E(t) + F_2(t))\\
&\quad + (\mathbf{z}_I(t) - \mathbf{z}_I^*)^T (\mathbf{B} E(t) + \mathbf{B}\mathbf{x}(t) - \mathbf{b})\\
&\quad + (\mathbf{z}_{II}(t) - \mathbf{z}_{II}^*)^T (\mathbf{A} E(t) + F_3(t))\}\\
&= (\mathbf{x}(t) - \mathbf{x}^*)^T E(t) + (\mathbf{z}_{II}(t) - \mathbf{z}_{II}^*)^T F_3(t)\\
&\quad - (\mathbf{x}(t) - \mathbf{x}^*)^T \mathbf{B}^T (\mathbf{B}\mathbf{x}(t) - \mathbf{b}))\\
&\quad + (\mathbf{y}(t) - \mathbf{y}^*)^T \mathbf{D} E(t) + (\mathbf{z}_I(t) - \mathbf{z}_I^*)^T \mathbf{B} E(t)\\
&\quad + (\mathbf{z}_{II}(t) - \mathbf{z}_{II}^*)^T \mathbf{A} E(t) + (\mathbf{y}(t) - \mathbf{y}^*)^T F_2(t)\\
&\quad + (\mathbf{x}(t) - \mathbf{x}^*)^T (-\mathbf{A}^T F_3(t) - \mathbf{D}^T F_2(t))\\
&\quad - (\mathbf{z}_I(t) - \mathbf{z}_I^*)^T \mathbf{B}\mathbf{x}(t) - \mathbf{b}).
\end{aligned}$$

Using $\mathbf{B}\mathbf{x}^* = \mathbf{b}$ and adding (11)–(13) in Proposition 2, we have

$$\begin{aligned}
\frac{dV(\mathbf{u}(t))}{dt} &\leq \lambda \left\{-\|E(t)\|_2^2 - \|F_2(t)\|_2^2 - \|F_3(t)\|_2^2\right.\\
&\qquad \left. -\|\mathbf{B}\mathbf{x}(t) - \mathbf{b}\|_2^2\right\} \leq 0.
\end{aligned}$$

It follows that $dV/dt = 0$ at $\hat{\mathbf{u}}$ if and only if $\hat{\mathbf{u}}$ satisfies $\|E(t)\|_2^2 = \|F_2(t)\|_2^2 = \|F_3(t)\|_2 = 0$ and $\|\mathbf{B}\mathbf{x}(t) - \mathbf{b}\|_2^2 = 0$. Thus, $d\mathbf{u}/dt = 0$ at $\hat{\mathbf{u}}$ which is an equilibrium point of (9). By the given analysis of the correspondence [13], [17], we know that the state trajectory of (9) will converge to $\hat{\mathbf{u}}$. Because the equilibrium point of (9) equals to the solution of (3), from Proposition 1, we obtain that $\mathbf{w}(t)$ converges globally to an optimal solution of (1).

III. ILLUSTRATIVE EXAMPLE

In this section, we give examples to demonstrate the effectiveness of the proposed compact CRNN. Since the existing CRNN (Xia and Kamel, 2008) has been shown to have a better performance in dealing with the LAD problems with degeneracy than the linear programming

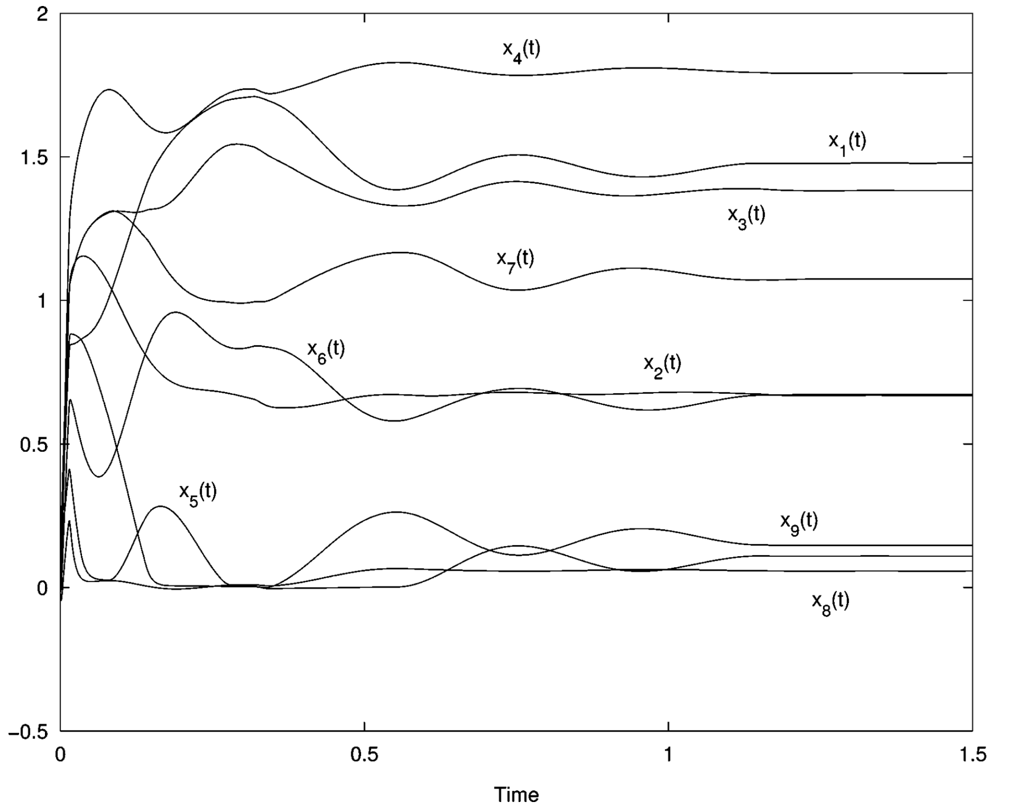

Fig. 1. Convergence behavior of the state trajectory $x(t)$ based on the proposed compact CRNN in Example 1.

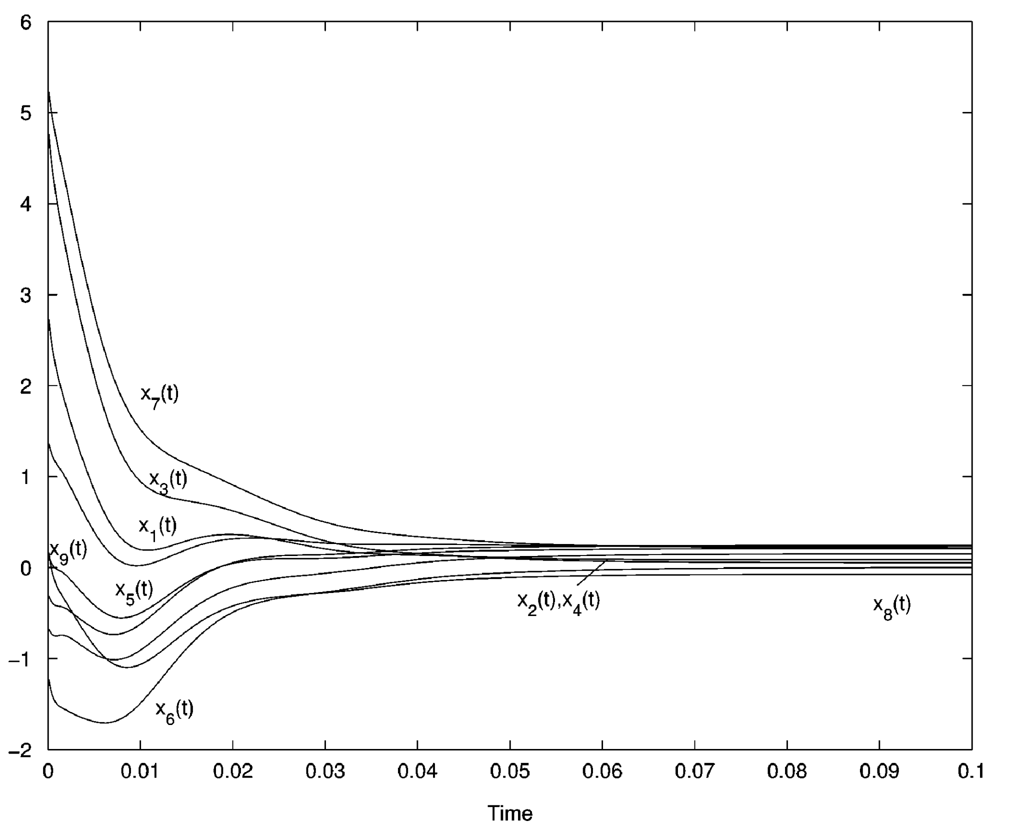

Fig. 2. Convergence behavior of the state trajectory $x(t)$ based on the proposed compact CRNN in Example 2.

method (Ruzinsky and Olsen, 1989) and the disciplined convex programming algorithm (Grant *et al.*, 2005), we here compare the proposed CRNN with the existing CRNN in simulation. The simulation is conducted in MATLAB.

Example 1: Consider the general constrained LAD problem (1), where $\Omega = \{\mathbf{x} \in R^9 \,|\, 0 \le x_i \le 2 (i = 1, \ldots, 9)\}$ and

$$\mathbf{D} = \begin{pmatrix} 5 & 9 & 6 & 9 & 3 & 8 & 1 & 3 & 0 \\ 3 & 7 & 6 & 9 & 0 & 1 & 9 & 1 & 9 \\ 4 & 3 & 0 & 7 & 1 & 8 & 8 & 1 & 3 \\ 12 & 19 & 12 & 25 & 4 & 17 & 18 & 5 & 12 \\ 4 & 13 & 12 & 11 & 2 & 1 & 2 & 3 & 6 \end{pmatrix}$$

$$\mathbf{d} = \begin{pmatrix} 45 \\ 46 \\ 36 \\ 125 \\ 55 \end{pmatrix}.$$

$\mathbf{p} = 7.5, \mathbf{b} = 1, \mathbf{A} = [1, -1, 1, -1, 1, -1, 1, -1, 1]$, and $\mathbf{B} = [1, 1, 1, 1, 1, 1, 1, 1, 1]$. Since the matrix $\mathbf{D}$ is rank deficient, the constrained LAD problem has nonunique optimal solutions. For a comparison, we consider an equivalent objective function $\|\hat{D}\mathbf{x} - \hat{\mathbf{d}}\|_1$ with the optimal objective value being 0.0356, where $\hat{D} = D/\|D\|_2$ and $\hat{\mathbf{d}} = \mathbf{d}/\|D\|_2$. The proposed compact CRNN defined in (9) has 11 number of neurons and always converges to an optimal solution $\mathbf{x}^*$ with the optimal value. On the other hand, we use the existing CRNN defined in (10) to solve the above problem. The existing CRNN not only requires 29 number of neurons but also has a slower convergence rate, as shown in Table I, where $\mathbf{e}$ is a vector with elements being 1, $\lambda = 100$, and the norm error is defined by $\|\hat{D}\mathbf{x}^* - \hat{\mathbf{d}}\|_1$. Finally, Fig. 1 displays the convergence behavior of the proposed compact CRNN with the zero initial point. From computed results we can see that the proposed CRNN defined in (9) has low computational complexity and has a faster convergence time than the existing CRNN defined in (10).

Example 2: Consider the general constrained LAD problem (1), where $\Omega = \{x \in R^9 \,|\, \mathbf{x}^T\mathbf{x} \le 0.5\}, \mathbf{D}, \mathbf{A}, \mathbf{B}, \mathbf{p}, \mathbf{d}$ are defined in Example 1.

The proposed compact CRNN defined in (9) always converges to the unique optimal solution. Let $\lambda = 100$, Fig. 2 displays the convergence behavior of the proposed compact CRNN with an random initial point. On the other hand, existing CRNNs [9]–[11], including (10), could not be used to solve the above problem.

IV. Conclusion

This correspondence develops a compact cooperative recurrent neural network (CRNN) for calculating general constrained L_1 norm estimators. The proposed CRNN significantly generalizes existing CRNNs as its special cases. Unlike the existing CRNNs, the proposed CRNN is compact and easily applied. Moreover, it can deal with the nonlinear constraint and a fast convergence rate due to low computational complexity. Simulation results confirm the good performance of the proposed CRNN.

Acknowledgment

The authors thank the associate editor and reviewers for their encouragement and valued comments.

References

[1] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1991.
[2] S. S. Kuo and J. J. Mammone, "Image-restoration by convex projections using adaptive constraints and L_1 norm estimation," *IEEE Trans. Signal Process.*, vol. 40, no. 1, pp. 159–168, Jan. 1992.
[3] J. M. Chen and B. S. Chen, "System parameter estimation with input/output noisy data and missing measurements," *IEEE Trans. Signal Process.*, vol. 48, no. 6, pp. 1548–1558, Jun. 2000.
[4] J. A. Cadzow, "Minimum l_1, l_2, and l_∞ norm approximate solutions an overdetermined system of linear equations," *Digital Signal Process.*, vol. 12, pp. 524–560, 2002.
[5] Y. Dodge and J. Jana, *Adaptive Regression*. New York: Springer, 2000.
[6] S. A. Ruzinsky and E. T. Olsen, "L_1 and L_∞ minimization via a variant of Karmarkar's algorithm," *IEEE Trans. Signal Process.*, vol. 37, no. 2, pp. 245–253, Feb. 1989.
[7] W. Li and J. J. Swetits, "The linear L_1 estimator and the Huber m-estimator," *SIAM J. Optimiz.*, vol. 8, pp. 457–475, 1998.
[8] M. Grant, S. Boyd, and Y. Y. Ye, "Disciplined convex programming," in *Global Optimization: From Theory to Implementation*, L. Liberti and N. Maculan, Eds. Dordrecht, Germany: Kluwer, 2005 [Online]. Available: http://www.stanford.edu/boyd/cvx
[9] Y. S. Xia and M. S. Kamel, "Cooperative recurrent neural networks for the constrained L1 norm estimator," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3192–3205, Jul. 2007.
[10] Y. S. Xia and M. S. Kamel, "Novel cooperative neural fusion algorithms for image restoration and image fusion," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 367–381, Feb. 2007.
[11] Y. S. Xia and M. S. Kamel, "A cooperative recurrent neural network for solving L1 estimation problems with linear constraints," *Neural Comput.*, vol. 20, no. 3, pp. 844–872, 2008.
[12] A. Cichocki and R. Unbehauen, *Neural Networks for Optimization and Signal Processing*. New York: Wiley, 1993.
[13] D. P. Mandic and J. A. Chambers, "Recurrent neural networks for prediction: Learning algorithms, architectures and stability," in *Adaptive and Learning Systems for Signal Processing, Communications and Control*. New York: Wiley, 2001.
[14] Z. Tian, K. L. Bell, and H. L. Van Trees, "A recursive least squares implementation for LCMP beamforming under quadratic constraint," *IEEE Trans. Signal Process.*, vol. 49, no. 6, pp. 1138–1145, Jun. 2001.
[15] Y. Tsaig and D. L. Donoho, "Breakdown of equivalence between the minimal L1-norm solution and the sparsest solution," *Signal Process.*, vol. 86, pp. 533–548, 2006.
[16] D. L. Donoho, "For most large underdetermined systems of equations, the minimal $L1$-norm near-solution approximates the sparsest near-solution," *Commun. Pure Appl. Math.*, vol. 59, pp. 907–934, 2006.
[17] Y. S. Xia and J. Wang, "A general projection neural network for solving monotone variational inequality and related optimization problems," *IEEE Trans. Neural Netw.*, vol. 15, no. 2, pp. 318–328, Mar. 2004.
[18] D. Kinderlehrer and G. Stampcchia, *An Introduction to Variational Inequalities and Their Applications*. New York: Academic, 1980.