

Design of Interval Type-2 Fuzzy Logic Systems Using Prior Knowledge via Optimization Algorithms

Tiechao Wang and Jianqiang Yi, *Senior Member, IEEE*

Institute of Automation
Chinese Academy of Sciences
Beijing 100190, China

Email: {tiechao.wang & jianqiang.yi}@ia.ac.cn

Tiechao Wang

College of Electrical Engineering
Liaoning University of Technology
Jinzhou 121001, China

Email: tiechao.wang@gmail.com

Abstract—The paper presents the methods of integrating prior knowledge with a first-order Single-Input Single-Output (SISO) Interval Type-2 Takagi-Sugeno-Kang (TSK) Fuzzy Logic System (IT2FLS) for function approximation under noisy circumstances. Firstly, sufficient conditions on the antecedent and the consequent parameters of the IT2FLS are given to ensure that three kinds of prior knowledge – monotonicity, symmetry and special points, can be embedded into the IT2FLS. And then, we use three optimization algorithms – constrained least squares algorithm, active-set algorithm and hybrid learning algorithm to design the IT2FLS, respectively. The effectiveness of the three algorithms and the comparisons of their performance are demonstrated by simulation examples.

Keywords – interval type-2 fuzzy logic system; optimization algorithm; prior knowledge

I. INTRODUCTION

Prior knowledge is a kind of important information sources and can play a significant role in many real problems, especially when input-output data are not informative enough or corrupted by noise. For a grey-box identification problem, it is not necessary to estimate what we already know, but instead we should take advantage of the prior knowledge of plants or systems. In the past several years, some work was done to incorporate some prior knowledge with neural network systems [1] or fuzzy logic systems [2] – [4]. In [1], Joerding studied methods for incorporating the monotonicity into feedforward networks. Lindskog [2] proposed a fuzzy model structure to ensure monotonic gain characteristics in identified models. Li [3, 4] integrated the prior knowledge of bounded range, symmetry and monotonicity into a zeroth-order IT2FLS and designed the fuzzy system by constrained least squares algorithm.

Type-2 Fuzzy Logic Systems (T2FLSs) can better handle different sources of uncertainties and debilitate noisy disturbance than traditional or type-1 fuzzy logic systems (T1FLSs) [5, 6]. A lot of excellent work focused on how to design T1FLSs and T2FLSs. Jang [7] presented the ANFIS architecture which is a type-1 fuzzy inference system implemented in the framework of adaptive network, and by using hybrid learning procedure, the ANFIS can yield remarkable results. In [8], Castro proposed an efficient mechanism for modeling

by combined hybrid learning algorithm with an interval type-2 fuzzy neural network. Méndez [9] presented a novel hybrid learning method to construct an interval type-2 TSK fuzzy model capable of approximating the behavior of steel strip temperature based on orthogonal least squares and back propagation methods. At present, to the authors' knowledge, there is no literatures about designing first-order SISO IT2FLSs with prior knowledge description via constrained hybrid method. In this paper, we first give the sufficient conditions which ensure the above-mentioned three kinds of prior knowledge can be combined with the IT2FLS. And then, we design the first-order SISO IT2FLS via constrained least squares algorithm, active-set algorithm and hybrid learning algorithm, respectively.

II. IT2FLSs AND PARAMETER CONDITIONS FOR PRIOR KNOWLEDGE

In this section, the basic concepts about first-order SISO IT2FLSs are introduced. And then, we give the sufficient conditions on the parameters of the IT2FLS which ensure that the prior knowledge of monotonicity, symmetry and special points can be integrated into the IT2FLS.

A. First-order SISO IT2FLSs

Consider a first-order SISO IT2FLS whose fuzzy rule base consists of M fuzzy rules. The i th rule R^i is denoted as

$$R^i : \text{IF } x \text{ is } \tilde{A}^i, \text{ THEN } y^i = w_0^i + w_1^i x,$$

where $i = 1, 2, \dots, M$; x is the input variable, $\tilde{A}^i$ s are triangular IT2FSs which can be completely depicted by the lower and the upper membership functions – $\underline{\mu}_{\tilde{A}^i}(x)$ and $\overline{\mu}_{\tilde{A}^i}(x)$ which are expressed by (1) and (2), as illustrated in Fig. 1.

$$\underline{\mu}_{\tilde{A}^i}(x) = \begin{cases} 0, & x < \underline{a}^i; \\ \frac{\underline{b}^i(x - \underline{a}^i)}{\underline{b}^i - \underline{a}^i}, & \underline{a}^i \leq x < \underline{b}^i; \\ \frac{\underline{b}^i(\underline{c}^i - x)}{\underline{c}^i - \underline{b}^i}, & \underline{b}^i \leq x < \underline{c}^i; \\ 0, & x \geq \underline{c}^i. \end{cases} \quad (1)$$

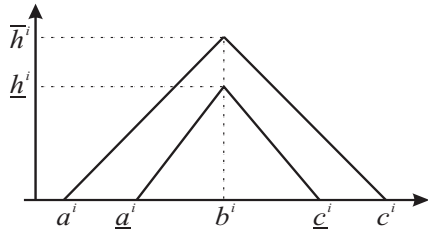


Fig. 1. Interval type-2 triangular membership function.

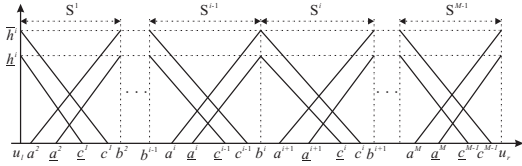


Fig. 2. Interval type-2 fuzzy partition.

$$\bar{\mu}_{\tilde{A}^i}(x) = \begin{cases} 0, & x < a^i; \\ \frac{\bar{h}^i(x-a^i)}{b^i-a^i}, & a^i \leq x < b^i; \\ \frac{\bar{h}^i(c^i-x)}{c^i-b^i}, & b^i \leq x < c^i; \\ 0, & x \geq c^i. \end{cases} \quad (2)$$

where $a^i \leq \underline{a}^i \leq b^i \leq \underline{c}^i \leq c^i$.

The firing set F^i of rule R^i is an interval type-1 set, i.e., $F^i = [\underline{\mu}_{\tilde{A}^i}(x), \bar{\mu}_{\tilde{A}^i}(x)]$. w_0^i and w_1^i are real constants of the consequent part, so the consequent output y^i of each rule is a crisp value.

Then, using the center-of-sets (COS) type-reducer and the center average defuzzifier [5], the defuzzification output $\hat{y}(x)$ of the IT2FLS is inferred as

$$\hat{y}(x) = \frac{1}{2}(y_l(x) + y_r(x)). \quad (3)$$

where $y_l(x)$ and $y_r(x)$ are the left and the right end points of a type-reduced interval set and can be expressed as

$$y_l(x) = \frac{\sum_{i=1}^L \bar{\mu}_{\tilde{A}^i}(x)y^i + \sum_{i=L+1}^M \underline{\mu}_{\tilde{A}^i}(x)y^i}{\sum_{i=1}^L \bar{\mu}_{\tilde{A}^i}(x) + \sum_{i=L+1}^M \underline{\mu}_{\tilde{A}^i}(x)}, \quad (4)$$

$$y_r(x) = \frac{\sum_{i=1}^R \underline{\mu}_{\tilde{A}^i}(x)y^i + \sum_{i=R+1}^M \bar{\mu}_{\tilde{A}^i}(x)y^i}{\sum_{i=1}^R \underline{\mu}_{\tilde{A}^i}(x) + \sum_{i=R+1}^M \bar{\mu}_{\tilde{A}^i}(x)}, \quad (5)$$

in which L and R can be computed by the iterative Karnik-Mendel algorithm [5].

B. Prior Knowledge of Monotonicity

We shall study the prior knowledge of monotonicity which includes the monotonically increasing and the monotonically decreasing properties. Here, for conciseness, we give the sufficient conditions to ensure that IT2FLSs monotonic properties without proof.

Theorem 2.1: Assume that an IT2FLS is single-input single-output, and that the input domain of the IT2FLS is $U = [u_l, u_r]$. The IT2FLS increases monotonically with respect to x , if the following conditions are satisfied:

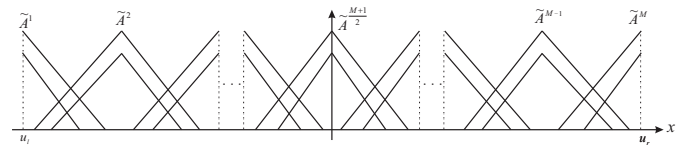


Fig. 3. Interval type-2 fuzzy symmetric partition with M triangular membership function.

- 1) The input domain U is partitioned by IT2FSs $\tilde{A}^1, \tilde{A}^2, \dots, \tilde{A}^M$ shown in Fig. 2, that is, $a^1 = b^1 = \underline{a}^1 = \underline{b}^1 = u_l$, $b^M = c^M = \underline{b}^M = \underline{c}^M = u_r$, $a^{i+1} \geq b^i$, $b^{i+1} \geq c^i$, $\underline{c}^i \geq \underline{a}^{i+1}$, for $i = 1, 2, \dots, M-1$;
- 2) $w_1^i \geq 0$, for $i = 1, 2, \dots, M$;
- 3) $w_0^{i+1} + w_1^{i+1}x \geq w_0^i + w_1^i x$, for all $x \in S^i$, and $i = 1, 2, \dots, M-1$.

Theorem 2.2: Assume that an IT2FLS is single-input single-output, and that the input domain of the IT2FLS is $U = [u_l, u_r]$. The IT2FLS decreases monotonically with respect to x , if the following conditions are satisfied:

- 1) The input domain U is partitioned by M IT2FSs $\tilde{A}^1, \tilde{A}^2, \dots, \tilde{A}^M$ shown in Fig. 2, that is, $a^1 = b^1 = \underline{a}^1 = \underline{b}^1 = u_l$, $b^M = c^M = \underline{b}^M = \underline{c}^M = u_r$, $a^{i+1} \geq b^i$, $b^{i+1} \geq c^i$, $\underline{c}^i \geq \underline{a}^{i+1}$, for $i = 1, 2, \dots, M-1$;
- 2) $w_1^i \leq 0$, for $i = 1, 2, \dots, M$;
- 3) $w_0^{i+1} + w_1^{i+1}x \leq w_0^i + w_1^i x$, if $u_l \geq 0$ or $u_r \leq 0$ for all $x \in S^i$ and $i = 1, 2, \dots, M-1$; if $U \in \mathbb{R}$ for all $x \in U$ and $i = 1, 2, \dots, M-1$.

C. Prior Knowledge of Symmetry

In this subsection, we give the sufficient conditions about the prior knowledge of symmetry which includes even symmetry and odd symmetry without proof.

Theorem 2.3: An IT2FLS is even symmetric, that is, $\hat{y}(x) = \hat{y}(-x)$, if the following conditions are satisfied:

- 1) The input domain of x is partitioned symmetrically around 0 by triangular IT2FSs $\tilde{A}^1, \tilde{A}^2, \dots, \tilde{A}^M$ shown in Fig. 3, where M is an odd number.
- 2) The consequent parameters of the rules R^i and R^{M-i+1} satisfy that $w_0^i = w_0^{M-i+1}$, $w_1^i = -w_1^{M-i+1}$ ($i \neq \frac{M+1}{2}$), and $w_1^{\frac{M+1}{2}} = 0$.

Theorem 2.4: An IT2FLS is odd symmetric, that is, $\hat{y}(x) = -\hat{y}(-x)$, if the following conditions are satisfied:

- 1) The input domain of x is partitioned symmetrically around 0 by triangular IT2FSs $\tilde{A}^1, \tilde{A}^2, \dots, \tilde{A}^M$ shown in Fig. 3, where M is an odd number.
- 2) The consequent parameters of the rules R^i and R^{M-i+1} satisfy that $w_0^i = -w_0^{M-i+1}$ ($i \neq \frac{M+1}{2}$), $w_0^{\frac{M+1}{2}} = 0$ and $w_1^i = w_1^{M-i+1}$.

D. Prior Knowledge of Special Points

In many practical systems the initial status is usually zero, and a lot of function curves pass through some special points such as origin. Hence, the prior knowledge of special points is widespread and important.

From a purely mathematic viewpoint, we formulate the prior knowledge of special points as follows

$$\hat{y}(x)|_{x=x_0} = y(x_0), \quad (6)$$

where $\hat{y}(x)$ is the output of the above-mentioned SISO IT2FLS. $y(x_0)$ is the known special point.

For example, when $\hat{y}(x)|_{x=0} = 0$ for the first-order SISO IT2FLS in Theorem 2.1, from (4) and (5) we have $y_l(x) = y_r(x) = w_0^{\frac{M+1}{2}}$. Then, it follows from (3) that $\hat{y}(x) = w_0^{\frac{M+1}{2}}$. Therefore, we can derive the constraint condition of the prior knowledge, i.e., $w_0^{\frac{M+1}{2}} = 0$.

III. DESIGN OF IT2FLSS VIA OPTIMIZATION ALGORITHMS

Suppose there exist r input-output data pairs $(x_1, y_1), \dots, (x_r, y_r)$. And our objective is that the parameters of the IT2FLS can be optimized subject to the following training criteria (7) such that the overall error measure E is minimal.

$$\min_{\theta} E = \min_{\theta} \sum_{k=1}^r (\hat{y}(x_k) - y_k)^2, \quad (7)$$

where θ is all the parameters which need to be optimized, and $k = 1, 2, \dots, r$. Next, according to this objective, we design the IT2FLS based on prior knowledge and sample data via constrained least squares algorithm, active-set algorithm and hybrid learning algorithm, respectively.

A. Design of IT2FLSs via Least Squares Algorithm

In this subsection, we design the IT2FLS via constrained least squares algorithm. Firstly, we should determine whether the output of the IT2FLS is linear with its consequent parameters.

Since no more than two fuzzy rules are fired, Equations (4) and (5) can be rewritten as

$$y_l(x) = \frac{\sum_{i=1}^M \mu_{\tilde{A}_l^i}(x) w_0^i}{\sum_{i=1}^M \mu_{\tilde{A}_l^i}(x)} + \frac{\sum_{i=1}^M \mu_{\tilde{A}_l^i}(x) x w_1^i}{\sum_{i=1}^M \mu_{\tilde{A}_l^i}(x)}, \quad (8)$$

$$y_r(x) = \frac{\sum_{i=1}^M \mu_{\tilde{A}_r^i}(x) w_0^i}{\sum_{i=1}^M \mu_{\tilde{A}_r^i}(x)} + \frac{\sum_{i=1}^M \mu_{\tilde{A}_r^i}(x) x w_1^i}{\sum_{i=1}^M \mu_{\tilde{A}_r^i}(x)}. \quad (9)$$

where

$$\mu_{\tilde{A}_l^i}(x) = \begin{cases} \underline{\mu}_{\tilde{A}_l^i}(x), & x \in S^{i-1}, \quad y^{i-1} \leq y^i; \\ \overline{\mu}_{\tilde{A}_l^i}(x), & x \in S^{i-1}, \quad y^{i-1} > y^i; \\ \underline{\mu}_{\tilde{A}_l^i}(x), & x \in S^i, \quad y^i \leq y^{i+1}; \\ \overline{\mu}_{\tilde{A}_l^i}(x), & x \in S^i, \quad y^i > y^{i+1}. \end{cases}$$

$$\mu_{\tilde{A}_r^i}(x) = \begin{cases} \underline{\mu}_{\tilde{A}_r^i}(x), & x \in S^{i-1}, \quad y^{i-1} \leq y^i; \\ \overline{\mu}_{\tilde{A}_r^i}(x), & x \in S^{i-1}, \quad y^{i-1} > y^i; \\ \underline{\mu}_{\tilde{A}_r^i}(x), & x \in S^i, \quad y^i \leq y^{i+1}; \\ \overline{\mu}_{\tilde{A}_r^i}(x), & x \in S^i, \quad y^i > y^{i+1}. \end{cases}$$

in which $y^i = w_0^i + w_1^i x$.

Then, it follows from (3), (8) and (9) that

$$\hat{y}(x) = \Phi^T(x) \omega, \quad (10)$$

where

$$\omega = [w_0^1, w_0^2, \dots, w_0^M, w_1^1, w_1^2, \dots, w_1^M]^T, \\ \Phi(x) = [\phi_0^1(x), \phi_0^2(x), \dots, \phi_0^M(x), \phi_1^1(x), \phi_1^2(x), \dots, \phi_1^M(x)]^T,$$

in which

$$\phi_0^i(x) = \frac{1}{2} \left(\frac{\mu_{\tilde{A}_l^i}(x)}{\sum_{i=1}^M \mu_{\tilde{A}_l^i}(x)} + \frac{\mu_{\tilde{A}_r^i}(x)}{\sum_{i=1}^M \mu_{\tilde{A}_r^i}(x)} \right), \\ \phi_1^i(x) = \frac{1}{2} \left(\frac{\mu_{\tilde{A}_l^i}(x)x}{\sum_{i=1}^M \mu_{\tilde{A}_l^i}(x)} + \frac{\mu_{\tilde{A}_r^i}(x)x}{\sum_{i=1}^M \mu_{\tilde{A}_r^i}(x)} \right).$$

From (10), we can conclude that the output of the IT2FLS is linear with its consequent parameters. Furthermore, the training criteria (7) can be rewritten as:

$$\min_{\omega} E = \min_{\omega} (\Phi \omega - \mathbf{y})^T (\Phi \omega - \mathbf{y}), \quad (11)$$

where

$$\mathbf{y} = [y_1, y_2, \dots, y_r]^T, \\ \Phi = [\phi^T(x_1), \phi^T(x_2), \dots, \phi^T(x_r)]^T \in \mathbb{R}^{r \times 2M}.$$

If the partitions of the IT2FLS are determined beforehand according to the conditions in the above theorems, we can use the sample data to train the consequent parameters of the IT2FLS via constrained least squares algorithm. Note that the given sample data (x_k, y_k) ($k = 1, 2, \dots, r$) should conform to the above-mentioned some or all prior knowledge. Also, in Theorems 2.1 – 2.4, if antecedent parameters are determined, the conditions on the consequent parameters can be expressed as linear inequalities and/or linear equalities with respect to the parameters. Then, the minimization problem (11) can be abstractly formulated as the following constrained least squares optimization problem,

$$\begin{cases} \min_{\omega} (\Phi \omega - \mathbf{y})^T (\Phi \omega - \mathbf{y}) \\ \text{subject to } C \omega \leq 0 \text{ and/or } C_{eq} \omega = 0 \end{cases} \quad (12)$$

where both C and C_{eq} are linear constraint matrixes.

Next, we deduce the definite form of $C \omega \leq 0$ or $C_{eq} \omega = 0$ according to the different prior knowledge.

1) *Monotonically increasing property*: The constraints on the consequent parameters of the IT2FLS in Conditions 2) and 3) of Theorem 2.1 can be expressed as

$$\begin{aligned} -w_1^i &\leq 0, \text{ for } i = 1, 2, \dots, M, \\ w_0^i - w_0^{i+1} + w_1^i x_k - w_1^{i+1} x_k &\leq 0, \\ \text{for } x_k \in S^i, i &= 1, 2, \dots, M-1. \end{aligned}$$

Here, assume there are n_i data pairs on interval S^i . Hence, there exist $M + \sum_{i=1}^{M-1} n_i$ linear inequality constraints for these consequent parameters, and the constraints can be written as the following matrix form.

$$\begin{bmatrix} V_0 \\ V_1 \\ \vdots \\ V_{M-1} \end{bmatrix} \omega \leq \mathbf{0}_{(M + \sum_{i=1}^{M-1} n_i) \times 1}, \quad (13)$$

where

$$V_0 = \begin{bmatrix} \mathbf{0}_M & -I_M \end{bmatrix},$$

$$V_i = \begin{bmatrix} & i & i+1 & & & & \\ 0 & \dots & 0 & 1 & -1 & 0 & \dots & 0 \\ \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & 1 & -1 & 0 & \dots & 0 \\ & M+i & M+i+1 & & & & & \\ 0 & \dots & 0 & x_{i,1} & -x_{i,1} & 0 & \dots & 0 \\ \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & x_{i,n_i} & -x_{i,n_i} & 0 & \dots & 0 \end{bmatrix},$$

in which x_{i,n_i} denotes the n_i th datum on interval S^i . $\mathbf{0}_M$ is M by M zero matrix, I_M is M by M identify matrix, $V_0 \in \mathbb{R}^{M \times M}$, and $V_i \in \mathbb{R}^{(M-1) \times 2M}$.

Remarks: Especially, when $U \in [0, +\infty]$, inequalities

$$\begin{aligned} w_1^i &\geq 0, \text{ for } i = 1, 2, \dots, M, \\ w_0^i &\leq w_0^{i+1}, w_1^i \leq w_1^{i+1} \text{ for } i = 1, 2, \dots, M-1, \end{aligned}$$

can ensure that Conditions 2) and 3) hold, and can be expressed as the following matrix form.

$$\begin{bmatrix} \mathbf{0}_M & -I_M \\ \Lambda_M & \mathbf{0}_{(M-1) \times M} \\ \mathbf{0}_{(M-1) \times M} & \Lambda_M \end{bmatrix} \mathbf{w} \leq \mathbf{0}_{(3M-2) \times 1}, \quad (14)$$

where

$$\Lambda_M = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{bmatrix} \in \mathbb{R}^{(M-1) \times M}.$$

2) *Monotonically decreasing property:* If $U \in [0, +\infty]$ or $U \in [-\infty, 0]$, the constraints on the consequent parameters of the IT2FLS in Conditions 2) and 3) of Theorem 2.2 can be expressed as $M + \sum_{i=1}^{M-1} n_i$ linear inequalities, and can be transformed into the following matrix form.

$$\begin{bmatrix} -V_0 \\ -V_1 \\ \vdots \\ -V_{M-1} \end{bmatrix} \mathbf{w} \leq \mathbf{0}_{(M + \sum_{i=1}^{M-1} n_i) \times 1} \quad (15)$$

Remarks: Especially, when $U \in [-\infty, 0]$, inequalities

$$\begin{aligned} w_1^i &\leq 0, \text{ for } i = 1, 2, \dots, M, \\ w_0^i &\geq w_0^{i+1}, w_1^i \leq w_1^{i+1}, \text{ for } i = 1, 2, \dots, M-1, \end{aligned}$$

can ensure that Conditions 2) and 3) hold, and can be expressed as the following matrix form.

$$\begin{bmatrix} \mathbf{0}_M & I_M \\ -\Lambda_M & \mathbf{0}_{(M-1) \times M} \\ \mathbf{0}_{(M-1) \times M} & \Lambda_M \end{bmatrix} \mathbf{w} \leq \mathbf{0}_{(3M-2) \times 1} \quad (16)$$

3) *Even symmetry:* The constraints on the consequent parameters of the IT2FLS in Condition 2) of Theorem 2.3 can be expressed as

$$\begin{aligned} w_0^i - w_0^{M-i+1} &= 0, \\ w_1^i + w_1^{M-i+1} &= 0, \\ w_1^{\frac{M+1}{2}} &= 0, \text{ for } i = 1, 2, \dots, \frac{M-1}{2}. \end{aligned}$$

Furthermore, the M inequalities can be written as the following matrix form.

$$\begin{bmatrix} I_{\frac{M-1}{2}} & \mathbf{0}_{\frac{M-1}{2} \times 1} & -I_{\frac{M-1}{2}}^R & \mathbf{0}_{\frac{M-1}{2} \times M} \\ \mathbf{0}_{\frac{M-1}{2} \times M} & I_{\frac{M-1}{2}} & \mathbf{0}_{\frac{M-1}{2} \times 1} & I_{\frac{M-1}{2}}^R \\ \mathbf{0}_{1 \times M} & \mathbf{0}_{1 \times \frac{M-1}{2}} & 1 & \mathbf{0}_{1 \times \frac{M-1}{2}} \end{bmatrix} \times \mathbf{w} = \mathbf{0}_{M \times 1} \quad (17)$$

where I_M^R is the $M \times M$ matrix which is obtained by rotating I_M 90 degrees clockwise.

4) *Odd symmetry:* The constraints on the consequent parameters of the IT2FLS in Conditions 2) of Theorem 2.4 can be expressed as M linear inequalities, and can be transformed into the following matrix form.

$$\begin{bmatrix} I_{\frac{M-1}{2}} & \mathbf{0}_{\frac{M-1}{2} \times 1} & I_{\frac{M-1}{2}}^R & \mathbf{0}_{\frac{M-1}{2} \times M} \\ \mathbf{0}_{1 \times \frac{M-1}{2}} & 1 & \mathbf{0}_{1 \times \frac{M-1}{2}} & \mathbf{0}_{1 \times M} \\ \mathbf{0}_{\frac{M-1}{2} \times M} & I_{\frac{M-1}{2}} & \mathbf{0}_{\frac{M-1}{2} \times 1} & -I_{\frac{M-1}{2}}^R \end{bmatrix} \times \mathbf{w} = \mathbf{0}_{M \times 1} \quad (18)$$

In conclusion, we make use of the following steps to construct the IT2FLS:

- step 1) Determine the membership functions of IT2FSs $\tilde{A}^1, \tilde{A}^2, \dots, \tilde{A}^M$ according to condition 1) in the Theorems 2.1 – 2.4 and the corresponding prior knowledge.
- step 2) Let $C\mathbf{w} \leq 0$ in (12) be (13) for monotonically increasing property; (15) for monotonically decreasing property. Let $C_{eq}\mathbf{w} = 0$ be (17) for even symmetry; (18) for odd symmetry. Solve the constrained least squares optimization problem (12). Obtain the consequent parameter vector $\mathbf{w}$.

B. Design of IT2FLSs via Active-set Algorithm

In this subsection, we use active-set algorithm [10] to optimize all the parameters of the IT2FLS subject to (7).

Let us define $\mathbf{v}_i = [a^{i+1}, \underline{a}^{i+1}, \underline{c}^i, c^i, b^{i+1}]^T$ ($i \neq M-1$) and $\mathbf{v}_{M-1} = [a^M, \underline{a}^M, \underline{c}^{M-1}, c^{M-1}]^T$, where M is the number of IT2FSs. Furthermore the antecedent parameters can be defined as $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{M-1}]^T \in \mathbb{R}^{5M-6}$. Then, we can use the active-set algorithm to determine all the antecedent parameters $\mathbf{v}$ and the consequent parameters $\mathbf{w}$ according to different prior knowledge.

1) *Monotonically increasing property:* According to Condition 1) in Theorems 2.1, the constraints on the antecedent parameters can be expressed as

$$\begin{aligned} a^{i+1} &\leq \underline{a}^{i+1} \leq \underline{c}^i \leq c^i, \text{ for } i = 1, 2, \dots, M-1; \\ c^i &\leq b^{i+1}, \text{ for } i = 1, 2, \dots, M-2; \end{aligned}$$

Furthermore, the $5M - 7$ inequalities can be written as the following matrix form.

$$\Lambda_{5M-6} \mathbf{v} \leq \mathbf{0}_{(5M-7) \times 1}, \quad (19)$$

Since b^i ($i = 2, \dots, M - 1$) is optimization parameters, in Condition 3) $x \in S^i$ should be changed into $x \in U$. Hence, V_i can be written as

$$V_i = \begin{bmatrix} & i & i+1 & & & & \\ 0 & \dots & 0 & 1 & -1 & 0 & \dots & 0 \\ \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & 1 & -1 & 0 & \dots & 0 \\ & M+i & M+i+1 & & & & & \\ 0 & \dots & 0 & x_1 & -x_1 & 0 & \dots & 0 \\ \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & x_k & -x_k & 0 & \dots & 0 \end{bmatrix},$$

According to Condition 3), since $b^i, b^{i+1} \in S^i$ and w^i are unknown variables, if we want to determine all the antecedent and the consequent parameters, there exist $2M - 4$ nonlinear inequality constraints as follows:

$$\begin{aligned} w_0^{i+1} - w_0^i + (w_1^{i+1} - w_1^i)b^i &\geq 0, \text{ for } i = 2, \dots, M - 1 \\ w_0^{i+1} - w_0^i + (w_1^{i+1} - w_1^i)b^{i+1} &\geq 0, \text{ for } i = 1, \dots, M - 2 \end{aligned} \quad (20)$$

Hence, the minimization problem (7) can be abstractly formulated as the following constrained optimization problem:

$$\begin{cases} \min_{\mathbf{v}, \mathbf{w}} \sum_{k=1}^r (\hat{y}(x_k) - y_k)^2 \\ \text{subject to} \quad (13), (19) \text{ and } (20). \end{cases} \quad (21)$$

2) *Monotonically decreasing property*: Similarly, when $u_l \geq 0$ and $u_r \leq 0$, according to the conditions in Theorems 2.2, the minimization problem (7) can be abstractly formulated as the following constrained optimization problem:

$$\begin{cases} \min_{\mathbf{v}, \mathbf{w}} \sum_{k=1}^r (\hat{y}(x_k) - y_k)^2 \\ \text{subject to} \quad (15), (19) \text{ and } \end{cases} \quad (22)$$

$$\begin{aligned} w_0^{i+1} - w_0^i + (w_1^{i+1} - w_1^i)b^i &\leq 0 \text{ for } i = 2, \dots, M - 1, \\ w_0^{i+1} - w_0^i + (w_1^{i+1} - w_1^i)b^{i+1} &\leq 0 \text{ for } i = 1, \dots, M - 2. \end{aligned}$$

3) *Even Symmetry*: According to Condition 1) in Theorems 2.3, the constraints on the antecedent parameters can be expressed as

$$\begin{aligned} a^{i+1} &\leq \underline{a}^{i+1} \leq \underline{c}^i \leq c^i, \text{ for } i = 1, 2, \dots, \frac{M-1}{2}; \\ c^i &\leq b^{i+1}, \text{ for } i = 1, 2, \dots, \frac{M-3}{2}; \\ a^i &= c^{M-i+1}, \underline{a}^i = \underline{c}^{M-i+1}, \text{ for } i = 2, \dots, M; \\ b^{\frac{M+1}{2}} &= 0, b^i = b^{M-i+1}, \text{ for } i = 2, \dots, M - 1, i \neq \frac{M-1}{2}; \end{aligned}$$

Furthermore, the $\frac{5M-9}{2}$ inequalities and the $\frac{5(M-1)}{2}$ equalities can be written as the following matrix form.

$$\begin{bmatrix} W & \mathbf{0}_{\frac{5M-7}{2} \times \frac{5M-7}{2}} \end{bmatrix} \mathbf{v} \leq \mathbf{0}_{\frac{5M-7}{2} \times 1}, \quad (23)$$

$$\begin{bmatrix} I_{\frac{5M-7}{2}} & \mathbf{0}_{\frac{5M-7}{2} \times 1} & I_{\frac{5M-7}{2}}^R \\ \mathbf{0}_{1 \times \frac{5M-7}{2}} & 1 & \mathbf{0}_{1 \times \frac{5M-7}{2}} \end{bmatrix} \mathbf{v} = \mathbf{0}_{\frac{5(M-1)}{2} \times 1}. \quad (24)$$

where

$$W = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{bmatrix} \in R^{\frac{5M-7}{2} \times \frac{5M-5}{2}}$$

Hence, the minimization problem (7) can be abstractly formulated as the following constrained optimization problem:

$$\begin{cases} \min_{\mathbf{v}, \mathbf{w}} \sum_{k=1}^r (\hat{y}(x_k) - y_k)^2 \\ \text{subject to} \quad (17), (23) \text{ and } (24). \end{cases} \quad (25)$$

4) *Odd Symmetry*: Similarly, according to Condition 1) in Theorems 2.4, the minimization problem (7) can be abstractly formulated as the following constrained optimization problem:

$$\begin{cases} \min_{\mathbf{v}, \mathbf{w}} \sum_{k=1}^r (\hat{y}(x_k) - y_k)^2 \\ \text{subject to} \quad (18), (23) \text{ and } (24). \end{cases} \quad (26)$$

In conclusion, we utilize the following steps to construct the IT2FLS via the active-set algorithm:

- step 1) Determine the initial value of the antecedent and the consequent parameters of the IT2FLS according to the conditions in Theorems 2.1 – 2.4 and the corresponding prior knowledge.
- step 2) Solve (21) for monotonically increasing property; (22) for monotonically decreasing property. Solve (25) for even symmetry; (26) for odd symmetry. And then obtain the antecedent and the consequent parameters of the IT2FLS.

C. Design of IT2FLSs via Hybrid Learning Algorithm

The fuzzy rules of IT2FLSs consist of the antecedent part and the consequent part. We use constrained least squares algorithm to determine the consequent parameters, and then use active-set algorithm to determine the antecedent parameters. The procedure is repeated on the specified demand. Next, we give the design steps of monotonic IT2FLSs and symmetric IT2FLSs, respectively.

1) Design steps of monotonic IT2FLSs:

- Step 1) Determine the number of the fuzzy system rules and the initial antecedent and consequent parameters of the IT2FLS, and then solve the following optimization problems to obtain the consequent parameters via constrained least squares algorithm. If a target function monotonically increases, then (27) should be solved.

Else if the target function monotonically decreases, then (28) should be solved.

$$\begin{cases} \min_{\mathbf{w}} (\Phi \mathbf{w} - \mathbf{y})^T (\Phi \mathbf{w} - \mathbf{y}) \\ \text{subject to (13).} \end{cases} \quad (27)$$

$$\begin{cases} \min_{\mathbf{w}} (\Phi \mathbf{w} - \mathbf{y})^T (\Phi \mathbf{w} - \mathbf{y}) \\ \text{subject to (15).} \end{cases} \quad (28)$$

Step 2) The consequent parameters are fixed, and only the antecedent parameters are adjusted by solving the following constraint optimization problem via the active-set algorithm.

$$\begin{cases} \min_{\mathbf{v}} \sum_{k=1}^r (\hat{y}(x_k) - y_k)^2 \\ \text{subject to (19).} \end{cases} \quad (29)$$

Step 3) If the overall error measure E is less than a desired level or the number of the iterative steps reaches a specified value, then stop. Otherwise return to Step 1.

2) *Design steps of symmetric IT2FLSs:*

Step 1) Determine the number of the fuzzy system rules and the initial antecedent and consequent parameters of the IT2FLS, and then solve the following optimization problems to obtain the consequent parameters via constrained least squares algorithm. If a target function is even symmetry, then (30) should be solved. Else if the target function is odd symmetry, then (31) should be solved.

$$\begin{cases} \min_{\mathbf{w}} (\Phi \mathbf{w} - \mathbf{y})^T (\Phi \mathbf{w} - \mathbf{y}) \\ \text{subject to (17).} \end{cases} \quad (30)$$

$$\begin{cases} \min_{\mathbf{w}} (\Phi \mathbf{w} - \mathbf{y})^T (\Phi \mathbf{w} - \mathbf{y}) \\ \text{subject to (18).} \end{cases} \quad (31)$$

Step 2) The consequent parameters are fixed, and only the antecedent parameters are adjusted by solving the following constraint optimization problem via the active-set algorithm.

$$\begin{cases} \min_{\mathbf{v}} \sum_{k=1}^r (\hat{y}(x_k) - y_k)^2 \\ \text{subject to (23) and (24).} \end{cases} \quad (32)$$

Step 3) If the overall error measure E is less than a desired level or the number of the iterative steps reaches a specified value, then stop. Otherwise return to Step 1.

IV. SIMULATION

In this section, we will give a simulation to show the usefulness of the prior knowledge and compare constrained least squares algorithm, active-set algorithm and hybrid learning algorithm when the output of a target function is corrupted by white noise.

Consider the following nonlinear function:

$$y = \left(\frac{x}{3}\right)^2 \tanh(|x|), \quad (33)$$

where $x \in U = [-3, 3]$. It is obvious that the curve of Function (33) passes through origin, is even symmetric, monotonically increases when $x \geq 0$, and monotonically decreases when $x \leq 0$.

In order to demonstrate the superiority of IT2FLSs, the output of the target function is corrupted by *Noise* which is the uniformly distributed noise on interval $[-n_b, n_b]$ ($n_b = 0.2$). From $\tilde{y} = y + \text{Noise}$, we can create 500-sample-data set $\mathcal{D}$ whose elements are $(x_i, \tilde{y}_i)$ ($i = 1, \dots, 500$), in which r_k training data were randomly chosen as the elements of training data subset $\mathcal{D}_k$. From (33), we can create 500-sample-data set $\mathcal{E}$ whose elements are (x_i, y_i) ($i = 1, \dots, 500$), in which r_k sample data were randomly chosen as the elements of test data subset $\mathcal{E}_k$. In this simulation, five cases are considered, that is, r_k equals 12, 25, 45, 70, 90, where $k = 1, \dots, 5$.

Firstly, we design a first-order SISO IT2FLS via the above-mentioned three optimization algorithms with and without Prior Knowledge description (PK), and then evaluate the approximation performance and the generalization performance of the IT2FLS, which are measured by Root Mean Square Error (RMSE) of the training data and the test data, and the two performance indices are formulated as follows:

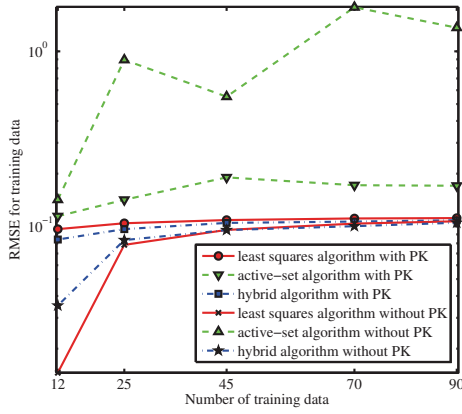
$$ap_k = \left(\frac{1}{r_k} \sum_{i=1}^{r_k} (\hat{y}(x_i) - \tilde{y}_i)^2 \right)^{\frac{1}{2}}, \quad (34)$$

$$gp_k = \left(\frac{1}{r_k} \sum_{i=1}^{r_k} (\hat{y}(x_i) - y_i)^2 \right)^{\frac{1}{2}}, \quad (35)$$

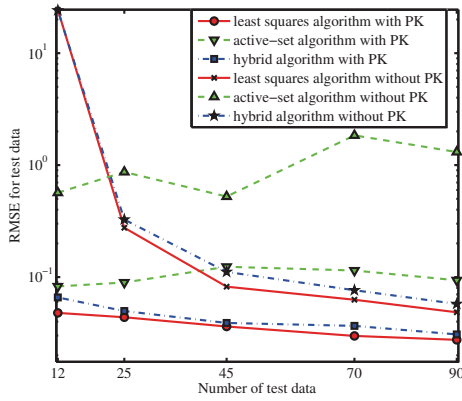
where r_k is the size of the training data subset $\mathcal{D}_k$ and the test data set $\mathcal{E}_k$ in case k ; ap_k can reflect the approximation ability of the IT2FLS for the training data; gp_k can reflect the generalization ability of the IT2FLS for test data. For the two indices, the less the values are, the better the performances are. Also, the statistical indices, ap_{sk} and gp_{sk} , which are the arithmetic mean of the corresponding indices obtained by the random training-test-data-selection process from 30 runs, are shown in Fig. 4 in semilogarithmic coordinates. For the sake of clarity, we plot the performance curves of the three algorithm with PK in cartesian coordinates once again shown in Fig. 5.

From Fig. 4, we observe that, generally, for generalization performance the algorithms with PK perform better than the ones without PK, and for approximation performance the algorithms without PK outperform the ones with PK except for the active-set algorithm. Furthermore, from Fig. 4 and Fig. 8, using the algorithms with PK can obtain better the overall performance, especially when the element number of a data subset is few.

From Fig.5, among the three optimization algorithms with PK, we can observe that the approximation performance and the generalization performance of the IT2FLS are the worst via the active-set algorithm. The IT2FLS achieves the best approximation performance via hybrid learning algorithm, and achieves the best generalization performance via constrained least squares algorithm. The least squares algorithm can overcome overfitting, and can obtain good results for generalization performance. For each iteration, the consequent parameters



(a)



(b)

Fig. 4. The performance curves of the three algorithms with and without prior knowledge description in semilogarithmic coordinates when $r_k \in \{12, 25, 45, 70, 90\}$ ($k = 1, \dots, 5$) and $n_b = 20\%$: (a) approximation performance (b) generalization performance.

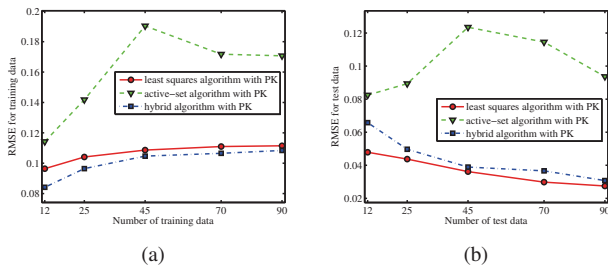


Fig. 5. The performance curves of the three algorithms with prior knowledge description when $r_k \in \{12, 25, 45, 70, 90\}$ ($k = 1, \dots, 5$) and $n_b = 20\%$: (a) approximation performance (b) generalization performance.

are updated via the constrained least squares estimator of the hybrid learning algorithm, and then the antecedent parameters are adjusted via the active-set algorithm part of the hybrid learning algorithm. In this case, by adjusting the antecedent parameters the overall error measure E is reduced further, but the algorithm needs iterative computation to achieve the smaller

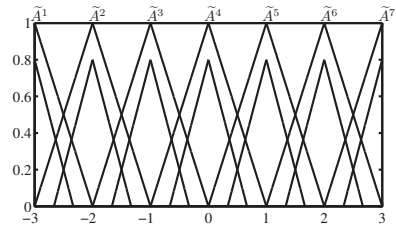


Fig. 6. Type-2 initial fuzzy partition of the IT2FLS.

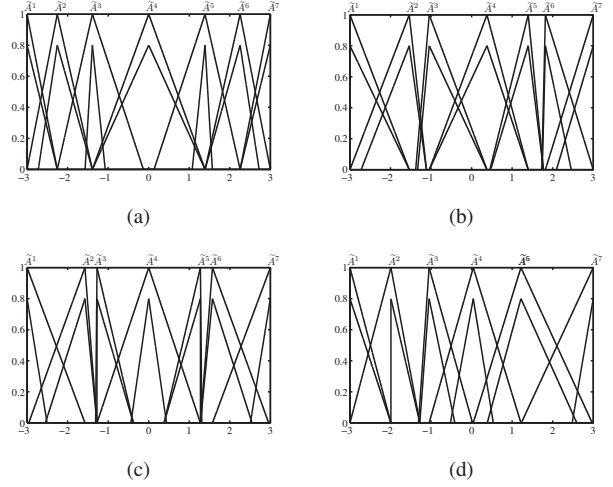


Fig. 7. Type-2 fuzzy partitions of the IT2FLS via different optimization algorithms: (a) active-set algorithm with PK (b) active-set algorithm without PK (c) hybrid learning algorithm with PK (d) hybrid learning algorithm without PK.

E . Hence the hybrid learning algorithm can better improve the approximation and the generalization performance than the active-set algorithm, but it consumes more computation time and more computer resources.

We use 7 fuzzy rules to describe the first-order SISO IT2FLS. For each optimization algorithm, the initial fuzzy partitions are shown in Fig. 6. After optimization computation with PK via active-set algorithm and hybrid learning algorithm, the fuzzy partitions of the IT2FLS are shown in Fig. 7(a) and 7(c), and the fuzzy partitions via the algorithms without PK are shown in Fig. 7(b) and 7(d). The consequent parameters of the IT2FLS via the three algorithms with and without PK are listed in Table I and Table II. Note that the fuzzy partitions in Fig. 7(a) and 7(c) and the consequent parameters satisfy the conditions in Theorem 2.1 – 2.3. Fig. 8 is one of identification results of the target function (33). From the figures, we can observe that the identification results are consistent with Fig. 4 and Fig. 5 on the whole.

In conclusion, from the standpoint of the approximation performance, generally speaking, the hybrid learning algorithm is the best among the three optimization algorithm. From the standpoint of the generalization performance and computational complexity, the constrained least squares algorithm is the best among the three optimization algorithm.

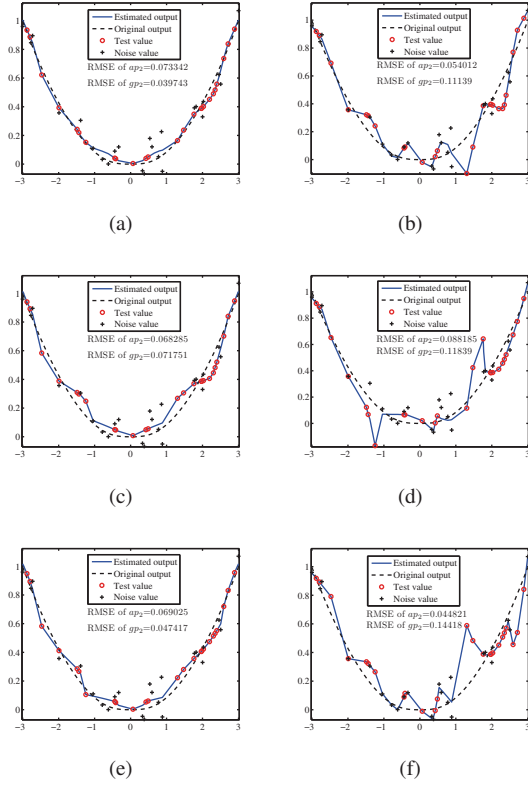


Fig. 8. One of identification results when the $r_2 = 25$ and $n_b = 20\%$: (a) constrained least squares algorithm with PK (b) constrained least squares algorithm without PK (c) active-set algorithm with PK (d) active-set algorithm without PK (e) hybrid learning algorithm with PK (f) hybrid learning algorithm without PK.

TABLE I
FUZZY RULES OF THE IT2FLS WITH PRIOR KNOWLEDGE

Antecedent part x	Consequent part y^i		
	least squares	active-set	hybrid learning
$\tilde{A}^1$	$0.39 - 0.20x$	$0.56 - 0.15x$	$0.32 - 0.23x$
$\tilde{A}^2$	0.39	$0.37 - 0.017x$	0.32
$\tilde{A}^3$	0.10	$0.27 - 0.017x$	0.11
$\tilde{A}^4$	0	0	0
$\tilde{A}^5$	0.10	$0.27 + 0.017x$	0.11
$\tilde{A}^6$	0.39	$0.37 + 0.017x$	0.32
$\tilde{A}^7$	$0.39 + 0.20x$	$0.56 + 0.15x$	$0.32 + 0.23x$

V. CONCLUSION

We have presented three efficient methods to design a first-order SISO IT2FLS for function approximation under noisy circumstances. The three kinds of prior knowledge – symmetry, monotonicity and special points are combined with the IT2FLS to solve the grey-box identification problem via constrained least squares algorithm, active-set algorithm and hybrid learning algorithm. From the simulation results and comparisons, we can conclude that the three kinds of prior knowledge can improve the generalization performance of the IT2FLS, and that the hybrid learning algorithm with PK performs best among the three optimization algorithms

TABLE II
FUZZY RULES OF THE IT2FLS WITHOUT PRIOR KNOWLEDGE

Antecedent part x	Consequent part y^i		
	least squares	active-set	hybrid learning
$\tilde{A}^1$	$0.57 - 0.13x$	$2.04 + 0.36x$	$0.94 - 0.005x$
$\tilde{A}^2$	$-0.37 - 0.37x$	$1.25 + 0.68x$	$-0.0065 - 0.16x$
$\tilde{A}^3$	$-0.69 - 0.78x$	$3.85 + 3.62x$	$-0.86 - 0.93x$
$\tilde{A}^4$	$-0.013 - 0.88x$	$-1.37 + 3.46x$	$0.042 - 1.005x$
$\tilde{A}^5$	$1.22 - 1.26x$	$-5 + 3.68x$	$1.10 - 0.95x$
$\tilde{A}^6$	$2.49 - 1.05x$	$0.38 + 0.0084x$	$0.013 + 0.53x$
$\tilde{A}^7$	$2.83 - 0.59x$	$-1.26 + 0.77x$	$-6.22 + 2.43x$

for approximation capability, and constrained least squares algorithm performs best for the generalization capability and computational complexity.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant No. 60975060.

REFERENCES

- [1] Wayne H. Joerding, and Jack L. Meador, "Encoding a priori information in feedback networks," *Neural Networks*, vol. 4, pp. 847–856, 1991.
- [2] Peter Lindskog and Lennart Ljung, "Ensuring monotonic gain characteristic in estimated models by fuzzy model structures," *Automatica*, vol. 36, pp. 311–317, Jun. 2000.
- [3] Chengdong Li, Jianqiang Yi and Dongbin Zhao, "Analysis and design of monotonic type-2 fuzzy inference system," *2009 IEEE International Conference on Fuzzy Systems*, Jeju Island, Korea, Aug. 2009, pp. 1193–1198.
- [4] Chengdong Li, Jianqiang Yi and Dongbin Zhao, "Design of interval type-2 fuzzy logic system using sampled data and prior knowledge," *ICIC Express Letters*, vol. 3, pp. 695–700, Sep. 2009.
- [5] J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, New Jersey : Prentice-Hall, 2001.
- [6] J. M. Mendel, "Advances in type-2 fuzzy sets and systems," *Information Sciences*, vol. 177, pp. 84–110, Jan. 2007.
- [7] J. S. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System," *IEEE transaction on systems, man, and cybernetics*, vol. 23, pp. 665 – 685, 1993.
- [8] J. R. Csato, O. Castillo, P. Melin, R. Antonio "A Hybrid Learning Algorithm for a Class of Interval Type-2 Fuzzy Neural Networks," *Information Sciences*, vol. 179, pp. 2175 – 2193, 2009.
- [9] G. M. Méndez, M. A. Hernandez "Hybrid Learning for Interval Type-2 Fuzzy Logic Systems Based on orthogonal least-squares and back-propagation methods," *Information Sciences*, vol. 179, pp. 2146 – 2157, 2009.
- [10] Fletcher, R. and M.J.D. Powell, "A Rapidly Convergent Descent Method for Minimization," *Computer Journal*, Vol. 6, pp. 163-168, 1963.