Multi-source Knowledge Based Unnormalized Interval Type-2 Fuzzy Logic Systems Design

Tiechao Wang and Jianqiang Yi
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

Chengdong Li
School of Info. and Electrical Engineering
Shandong Jianzhu University
Jinan 250101, China
Email: lichengdong@sdjzu.edu.cn

Abstract—In this paper we propose an effective method to design a Single-Input Single-Output (SISO) Unnormalized Interval Type-2 Takagi-Sugeno-Kang (TSK) Fuzzy Logic System (UIT2FLS) for noisy regression problems based on multi-source knowledge which includes here the information from sample data and the prior knowledge of bounded range, symmetry and monotonicity. The sufficient conditions are given which ensure that the prior knowledge can be embedded into the UIT2FLS, and then the UIT2FLS is designed so that the target function can be approached as accurately as possible via constrained least squares algorithm. The performance of the UIT2FLS is verified through comparisons with unnormalized type-1 Fuzzy Logic Systems (FLSs) and normalized interval type-2 FLSs under three different noisy circumstances. Simulation results verify the correctness of the sufficient conditions, and demonstrate that the UIT2FLS has the best overall performance.

Keywords – interval type-2 fuzzy logic system; constrained least squares algorithm; prior knowledge

I. INTRODUCTION

In the past decades, there had been many excellent studies on the theory and application of type-1 FLSs [1] – [5]. Type-2 FLSs are extensions of type-1 FLSs, and more complex than type-1 FLSs due to using type-2 fuzzy sets in the antecedent and the consequent part of fuzzy rules. Interval Type-2 Fuzzy Logic Systems (IT2FLSs), a special case of type-2 FLSs, have obvious advantages for handling different sources of uncertainties, such as noisy data.

Although interval type-2 FLSs provide a powerful framework to represent and handle the uncertainties, the problem of how to design interval type-2 FLSs is still a heavily researched issue, especially when sample data are insufficient and noisy. In order to overcome the problem, we can utilize some prior knowledge of plants or systems, such as bounded range, symmetry, monotonicity, etc., to compensate the insufficiency of the information from sample data or to make the data invalid which are out of accord with the known prior knowledge. Recently, many researchers have incorporated prior knowledge into neural networks [6], type-1 FLSs [7] – [10] and normalized IT2FLSs [11], [12]. Lindskog [7] proposed a fuzzy model structure to ensure monotonic gain characteristics in identified models. In [8], Won gave sufficient conditions of monotonicity for normalized type-1 TSK fuzzy systems.

In [11], [12], Li and Yi gave sufficient conditions to ensure that the prior knowledge of bounded range, symmetry and monotonicity can be incorporated into normalized IT2FLSs.

Unnormalized type-1 TSK FLSs are employed by K. Tanaka [13], and then Liang combined unnormalized TSK fuzzy systems with type-2 fuzzy logic in [4]. Unnormalizing the output of a TSK FLS can provide fast inference, reduce computational complexity, and may achieve the same performance as the normalized TSK in some specific applications [13]. At present, there are a few studies in the literature that incorporate prior knowledge. In [14], Wang and Yi gave the sufficient conditions to ensure that the prior knowledge of monotonicity and convexity can be integrated into UIT2FLSs, but did not present the method of designing an UIT2FLS. In this paper, we propose an effective method to design an UIT2FLS based on multi-source knowledge, and demonstrate the advantages of UIT2FLSs over type-1 FLSs and normalized IT2FLSs in simulation.

II. SISO ZEROTH-ORDER UIT2FLS

Consider an SISO zeroth-order type-2 FLS whose fuzzy rule base consists of $M$ fuzzy rules. The $i$th rule $R_i$ is denoted as

$$R_i : \text{IF } x \text{ is } \tilde{A}_i, \text{ THEN } Y_i = [\underline{w}_i, \overline{w}_i], \quad (1)$$

where $i = 1, 2, \ldots, M$; $x$ is an input variable; $\tilde{A}_i$’s are the antecedent IT2FSs which can be completely depicted by the lower and the upper membership functions $\mu_{\tilde{A}_i}(x)$ and $\overline{\mu}_{\tilde{A}_i}(x)$, which are the trapezoidal membership functions formulated by (4) and (5) shown in Fig. 1, or the quadratic polynomial membership functions formulated by (6) and (7) shown in Fig. 2. The firing set $F^i$ of rule $R_i$ is an interval type-1 set, i.e., $F^i = [\underline{f}_i(x), \overline{f}_i(x)]$, in which $\underline{f}_i(x) = \mu_{\tilde{A}_i}(x)$ and $\overline{f}_i(x) = \overline{\mu}_{\tilde{A}_i}(x)$.

Definition 2.1: If the computation method of the SISO zeroth-order type-2 FLS described by (1) follows that

$$\hat{y}(x) = \frac{1}{\sum_{i=1}^{M} \underline{f}_i y_i} \sum_{i=1}^{M} \overline{f}_i y_i, \quad (2)$$

then we refer to the fuzzy system as UIT2FLS.
the UIT2FLS is inferred as

\[
\hat{y}(x) = \frac{1}{2} \sum_{i=1}^{M} (f_i(x)w_i + \bar{f}_i(x)\bar{w}_i),
\]

where

\[
\hat{A}_i(x) = \left\{ \begin{array}{ll}
0 & x < a_i, \\
\frac{h(x-a_i)}{(a_i-b_i)} & a_i \leq x < b_i, \\
\frac{h}{b_i} & b_i \leq x < c_i, \\
\frac{h}{c_i-x} & c_i \leq x < d_i, \\
0 & x \geq d_i.
\end{array} \right.
\]

\[
\bar{A}_i(x) = \left\{ \begin{array}{ll}
0 & x < a_i, \\
\frac{h(x-a_i)}{(a_i-b_i)} & a_i \leq x < b_i, \\
\frac{h}{b_i} & b_i \leq x < c_i, \\
\frac{h}{c_i-x} & c_i \leq x < d_i, \\
0 & x \geq d_i.
\end{array} \right.
\]

According to Theorem 13.2 in [1], the final output \(\hat{y}(x)\) of the UIT2FLS is inferred as

\[
\hat{y}(x) = \frac{1}{2} \sum_{i=1}^{M} (f_i(x)w_i + \bar{f}_i(x)\bar{w}_i).
\]

III. CONSTRAINT CONDITIONS OF PRIOR KNOWLEDGE

In this section, we will give some sufficient conditions to ensure that some prior knowledge can be integrated into an UIT2FLS. We will consider the prior knowledge of bounded range, symmetry and monotonicity.

A. Prior Knowledge of Bounded Range

For the prior knowledge of bounded range, we have the following results for an SISO zeroth-order UIT2FLS.

**Theorem 3.1**: Assume that the input domain \(U = [\mu, \bar{\mu}]\) is partitioned by \(M\) IT2FSs \(\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_M\) shown in Fig. 3 or Fig. 4. Then, the output \(\hat{y}(x)\) of the UIT2FLS is bounded on \([\underline{\mu}, \bar{\mu}]\), if the following conditions are satisfied:

1) No more than two fuzzy rules are fired.
2) Inequalities (8) and (9) hold.

\[
\frac{1}{2} \left( \min_{p=1,\ldots,M-1} \{\mu_{p+1}\} \min_{i=1,\ldots,M} \{\mu_i\} \right) \leq \bar{\mu}, \quad (8)
\]

\[
\frac{1}{2} \left( \max_{p=1,\ldots,M-1} \{\mu_{p+1}\} \max_{i=1,\ldots,M} \{\mu_i\} \right) \leq \bar{\mu}, \quad (9)
\]

where

\[
\mu_{p+1} = \sup_{x \in S^p} \{\mu_{\tilde{A}_p}(x) + \mu_{\tilde{A}_{p+1}}(x) : x \in S^p\},
\]

\[
\mu_{p+1} = \inf_{x \in S^p} \{\mu_{\tilde{A}_p}(x) + \mu_{\tilde{A}_{p+1}}(x) : x \in S^p\},
\]

\[
\mu_{p+1} = \sup_{x \in S^p} \{\mu_{\tilde{A}_p}(x) + \mu_{\tilde{A}_{p+1}}(x) : x \in S^p\},
\]

\[
\mu_{p+1} = \inf_{x \in S^p} \{\mu_{\tilde{A}_p}(x) + \mu_{\tilde{A}_{p+1}}(x) : x \in S^p\},
\]

\[
S^p = [m_p, m_{p+1}] \quad (p = 1, 2, \ldots, M-1).
\]

**Proof**: Suppose that no more than two fuzzy IT2FSs are fired, and the ordinal numbers of the two fired rules are \(p\) and \(p+1\).
Then, from (3), the final output can be rewritten as
\[
\hat{y}(x) = \frac{1}{2} \sum_{k=p}^{p+1} \left( \mu_{\hat{x}_k}(x) w^k + \mu_{\bar{x}_k}(x) \bar{w}^k \right).
\] (10)

When \( x' \) is an input to the UIT2FLS, we get
\[
\hat{y}(x') = \sum_{i=p}^{p+1} \mu_{\hat{x}_i}(x') w^i
\]
\[
\geq \left( \mu_{\hat{x}_p}(x') + \mu_{\hat{x}_{p+1}}(x') \right) \min_{i=1,2,\ldots,M} \{ w^i \}
\]
\[
\geq \min_{p=1,\ldots,M-1} \{ w^{p_{\min}} \} \min_{i=1,\ldots,M} \{ w^i \},
\] (11)
and
\[
\hat{y}(y) = \sum_{i=p}^{p+1} \mu_{\hat{y}_i}(y) w^i
\]
\[
\leq \left( \mu_{\hat{y}_p}(y) + \mu_{\hat{y}_{p+1}}(y) \right) \max_{i=1,2,\ldots,M} \{ w^i \}
\]
\[
\leq \max_{p=1,\ldots,M-1} \{ w^{p_{\max}} \} \max_{i=1,\ldots,M} \{ w^i \}.
\] (12)

In the similar way, we have
\[
y_r \geq \min_{p=1,\ldots,M-1} \{ w^{r_{\min}} \} \min_{i=1,\ldots,M} \{ w^i \},
\] (13)
and
\[
y_r \leq \max_{p=1,\ldots,M-1} \{ w^{r_{\max}} \} \max_{i=1,\ldots,M} \{ w^i \},
\] (14)

From (11), (12), (13) and (14), it is obvious that
\[
\frac{1}{2} \left( \min_{p=1,\ldots,M-1} \{ w^{p_{\min}} \} \min_{i=1,2,\ldots,M} \{ w^i \} \right)
+ \min_{p=1,\ldots,M-1} \{ w^{p_{\min}} \} \min_{i=1,2,\ldots,M} \{ w^i \}
\leq \hat{y}(x) \leq \frac{1}{2} \left( \max_{p=1,\ldots,M-1} \{ w^{p_{\max}} \} \max_{i=1,2,\ldots,M} \{ w^i \} \right)
+ \max_{p=1,\ldots,M-1} \{ w^{p_{\max}} \} \max_{i=1,2,\ldots,M} \{ w^i \}.
\] (15)

As a result, if (8) and (9) hold, then \( \hat{b} \leq \hat{y}(x) \leq \bar{b} \).

Hence, the theorem holds. \( \square \)

**B. Prior Knowledge of Symmetry**

For the prior knowledge of symmetry, we have the following results for an SISO zeroth-order UIT2FLS.

**Theorem 3.2:** An UIT2FLS is even symmetry around \( a \), i.e., \( \hat{y}(x) = \hat{y}(2a - x) \), if the following conditions are satisfied:

1. The input domain of \( x \in [\underline{u}, \bar{u}] \) is partitioned symmetrically around \( a \) by IT2FSs \( \hat{A}^1, \hat{A}^2, \ldots, \hat{A}^M \) shown in Fig. 5 or Fig. 6, where \( M \) is an odd number.
2. The consequent interval weights of the rules \( R^i \)'s satisfy that \( [\mu^i, \bar{w}^i] = [\mu^{M+1-i}, \bar{w}^{M+1-i}] \) for \( i = 1, 2, \ldots, \frac{M-1}{2} \).

**Proof:** when \( x' \) is an input, assume that \( t \) fuzzy IT2FSs can be fired, and the ordinal numbers of the fired rules are \( i_1, i_2, \ldots, i_t \), here \( t \) is a positive integer.

As the fuzzy rule base is complete, and the input domain is partitioned symmetrically around \( a \). Therefore, when \( 2a - x' \) is input to the UIT2FLS, the ordinal numbers of the fired rules are \( M + 1 - i_1, M + 1 - i_2, \ldots, M + 1 - i_t \).

The conditions of Theorem 3.2 can also be stated as follows:
\[
\mu_{\hat{A}_p}(x') = \mu_{\hat{A}_M+1-p}(2a - x'),
\] (16a)
\[
\bar{w}_{\hat{A}_p}(x') = \bar{w}_{\hat{A}_M+1-p}(2a - x'),
\] (16b)
and
\[
\mu^{p_{\min}} = \mu^{M+1-i_{p}},
\] (17a)
\[
\bar{w}^{p_{\min}} = \bar{w}^{M+1-i_{p}},
\] (17b)
where \( p = 1, 2, \ldots, t \).

Since
\[
\hat{y}(x') = \sum_{p=1}^{t} \mu_{\hat{A}_p}(x') w^{p_{\min}}
\]
and
\[
\hat{y}(2a - x') = \sum_{p=1}^{t} \mu_{\hat{A}_p}(2a - x') w^{M+1-i_{p}},
\]
according to (16a) and (17a), we can derive that
\[
\hat{y}(x') = \hat{y}(2a - x').
\]
Similarly, according to (16b) and (17b), we can derive that
\[
\bar{y}(x') = \bar{y}(2a - x').
\]

Therefore,
\[
y(x') = \frac{1}{2} (\hat{y}(x') + \bar{y}(x'))
\]
\[
= \frac{1}{2} (\hat{y}(2a - x') + \bar{y}(2a - x'))
\]
\[
= y(2a - x')
\] (18)

From (18), Theorem 3.2 holds. \( \square \)
C. Prior Knowledge of Monotonicity

We consider the monotonicity of an SISO zeroth-order UIIT2FLS. Firstly, the definition of SISO monotonic fuzzy systems is given as follows.

Definition 3.1: Let \( x \) be an input of a fuzzy system defined on \( U = [a, b] \), where \( U \subset \mathbb{R} \). And, \( y = F(x) \) be the output of the fuzzy system in the set \( V \subset \mathbb{R} \). Then, \( F: U \rightarrow V \) is said to be monotonically increasing if \( a \leq x_1 < x_2 \leq b \) implies

\[
F(x_1) \leq F(x_2).
\]  

(19)

If Inequality (19) is reversed, we obtain the definition of SISO monotonically decreasing fuzzy systems.

For the prior knowledge of monotonically increasing property, we have the following results for an SISO zeroth-order UIIT2FLS.

Theorem 3.3: Assume that the fuzzy system is single-input single-output, and that the input domain \( U = [\underline{u}, \overline{u}] \) which consists of \( M - 1 \) intervals \( S^i = [m^i, m^{i+1}] \) is partitioned by \( M \) IT2FSs \( \tilde{A}^1, \tilde{A}^2, \ldots, \tilde{A}^M \) shown in Fig. 3 or Fig. 4. Then, the UIIT2FLS monotonically increases with respect to \( x \), if the following conditions are satisfied:

1) No more than two fuzzy rules are fired, i.e. \( a^1 = \tilde{a}^1 = \tilde{b}^1 = b^1 = m^1 = \underline{u}, m^M = \overline{u} \), \( M = \underline{u}, b^M = \overline{u} = \overline{u} = \overline{u} \), \( a^i \leq \tilde{a}^i, b^i \leq \tilde{b}^i, c^i \leq \tilde{c}^i, d^i \leq \tilde{d}^i, b^i \leq m^i \leq \tilde{c}^i \) for \( 2 \leq i \leq M - 1 \);
2) For any \( x \in S^i \), \( b^{i+1} = d^i, b^{i+1} = d^i, a^{i+1} = c^i, A^{i+1} = \tilde{A}^i \), so that \( \mu_{\tilde{A}^i}(x) + \mu_{\tilde{A}^i+1}(x) = \overline{u} \), and \( \mu_{\tilde{A}^i}(x) + \mu_{\tilde{A}^i+1}(x) = \overline{u} \) (i = 1, 2, ..., M - 1);
3) \( \underline{w}^i \leq \underline{w}^{i+1} \) and \( \overline{w}^i \leq \overline{w}^{i+1} \) (i = 1, 2, ..., M - 1).

Proof: Assume that no more than two fuzzy IT2FSs can be fired, and the ordinal numbers of the fired rules are \( k \) and \( k+1 \).

Let

\[
\tilde{y}^k(x) = \sum_{i=k}^{k+1} \mu_{\tilde{A}^i}(x) \underline{w}^i, \quad \overline{y}^k(x) = \sum_{i=k}^{k+1} \mu_{\tilde{A}^i}(x) \overline{w}^i.
\]

(20)

(21)

Hence, we have \( \overline{y}(x) = \frac{1}{2}(\underline{y}(x) + \overline{y}(x)) \).

1) When \( x \in S^k \), \( x' \in S^k \), \( x' > x \), from the second condition of this theorem, (20) can be rewritten as

\[
\underline{y}^k(x) = \mu_{\tilde{A}_k}(x) \underline{w}^k + \mu_{\tilde{A}_{k+1}}(x) \underline{w}^{k+1} = (\overline{u} - \mu_{\tilde{A}_{k+1}}(x)) \underline{w}^k + \mu_{\tilde{A}_{k+1}}(x) \underline{w}^{k+1} = \underline{w}^k + \mu_{\tilde{A}_{k+1}}(x) (\underline{w}^{k+1} - \underline{w}^k).
\]

(22)

Then, we get

\[
\underline{y}^k(x') - \underline{y}^k(x) = (\mu_{\tilde{A}_{k+1}}(x') - \mu_{\tilde{A}_{k+1}}(x)) (\underline{w}^{k+1} - \underline{w}^k).
\]

(23)

Notice that \( \mu_{\tilde{A}_{k+1}}(x) \) monotonically increases when \( x \in S^k \), i.e.,

\[
\mu_{\tilde{A}_{k+1}}(x') - \mu_{\tilde{A}_{k+1}}(x) \geq 0.
\]

If the third condition of this theorem holds, we can deduce that

\[
\underline{y}^k(x') - \underline{y}^k(x) \geq 0.
\]

(24)

In the same way, we can also obtain that

\[
\overline{y}^k(x') - \overline{y}^k(x) \geq 0.
\]

(25)

According to (24) and (25), when \( x \in S^k \), we can deduce that

\[
\overline{y}(x') \geq \overline{y}(x) \quad (26)
\]

2) When \( x \in S^k \), \( x' \in S^{k'} \), \( k' > k \), and \( k, k' = 1, 2, \ldots, M \), it is clear that \( x' > x \) holds. From Equations (20) and (21), when \( x \in S^k \), we can see that

\[
\overline{w}^k(x') \leq \overline{w}^k(x) \leq \overline{w}^{k+1}(x), \quad \underline{w}^k(x') \leq \underline{w}^k(x) \leq \underline{w}^{k+1}(x).
\]

With Condition 3) of this theorem holding, it follows that

\[
\overline{y}^1(x') = \overline{y}^2(x') \leq \ldots \leq \overline{y}^M(x'), \quad \underline{y}^1(x') \leq \underline{y}^2(x') \leq \ldots \leq \underline{y}^M(x'),
\]

where \( x^i \in S^i \), \( i = 1, \ldots, M - 1 \). So, in this circumstance, we can deduce that

\[
\overline{y}(x') \geq \overline{y}(x).
\]

(27)

Therefore, from the discussion above, we can conclude that the theorem holds.

Next, we give the sufficient conditions on the prior knowledge of monotonically decreasing property.

Theorem 3.4: Assume that the fuzzy system is single-input single-output, and that the input domain \( U = [\underline{u}, \overline{u}] \) which consists of \( M - 1 \) intervals \( S^i = [m^i, m^{i+1}] \) is partitioned by \( M \) IT2FSs \( \tilde{A}^1, \tilde{A}^2, \ldots, \tilde{A}^M \) as shown in Fig. 3 or Fig. 4. Then, the UIIT2FLS monotonically decreases with respect to \( x \), if the following conditions are satisfied:

1) No more than two fuzzy rules are fired, i.e. \( a^1 = \tilde{a}^1 = \tilde{b}^1 = b^1 = m^1 = \underline{u}, m^M = \overline{u} \), \( M = \underline{u}, b^M = \overline{u} = \overline{u} = \overline{u} \), \( a^i \leq \tilde{a}^i, b^i \leq \tilde{b}^i, c^i \leq \tilde{c}^i, d^i \leq \tilde{d}^i, b^i \leq m^i \leq \tilde{c}^i \) for \( 2 \leq i \leq M - 1 \);
2) For any \( x \in S^i \), \( b^{i+1} = d^i, b^{i+1} = d^i, a^{i+1} = c^i, \)

\[
\overline{A}^{i+1} = \tilde{A}^i,
\]

so that \( \mu_{\tilde{A}^i}(x) + \mu_{\tilde{A}^i+1}(x) = \overline{u} \), and \( \mu_{\tilde{A}^i}(x) + \mu_{\tilde{A}^i+1}(x) = \overline{u} \) (i = 1, 2, ..., M - 1);
3) \( \overline{w}^i \geq \overline{w}^{i+1} \) and \( \underline{w}^i \geq \underline{w}^{i+1} \) (i = 1, 2, ..., M - 1).

Proof: The theorem can be proved in the similar way as Theorem 3.3.

IV. DESIGN OF UIIT2FLS USING MULTI-SOURCE KNOWLEDGE

Suppose that there exist \( r \) input-output data pairs \((x_1, y_1), \ldots, (x_r, y_r)\). And our objective is that the parameters of the above-mentioned UIIT2FLS can be optimized subject to the following training criteria (28) such that overall error measure \( E \) is minimal.

\[
\min_E = \min_{w} \sum_{k=1}^{r} (\tilde{y}(x_k) - y_k)^2.
\]

(28)

Then, according to the objective, we design the UIIT2FLS based on sample data and prior knowledge via constrained least squares algorithm.
A. Design of UIT2FLSs Using the Constrained Least Squares Algorithm

Firstly, we should determine whether the output of the UIT2FLS is linear with its consequent parameters. Since (3) can be written as
\[
\hat{y}(x) = \phi^T(x)w,
\] (29)

where
\[
\phi_i(x) = \begin{cases} \frac{1}{2} \mu_{X_i}(x) & \text{if } i = 1, 2, \ldots, M, \\ \frac{1}{2} \nu_{X_i}(x) & \text{if } i = M + 1, \ldots, 2M, \end{cases}
\]
\[
\phi(x) = [\phi_1(x), \phi_2(x), \ldots, \phi_{2M}(x)]^T,
\]
\[
w = [w^1, \ldots, w^M, \overline{w}^1, \ldots, \overline{w}^M]^T,
\]
we can deduce that the output of the UIT2FLS is linear with its consequent parameters. Then, the overall error measure \( E \) can be rewritten as the following form:
\[
E = (\Phi w - y)^T(\Phi w - y),
\] (30)

where
\[
y = [y_1, y_2, \ldots, y_r],
\]
\[
\Phi = [\phi(x_1), \phi(x_2), \ldots, \phi(x_r)]^T.
\]

Since the sample data should satisfy one or several kinds of prior knowledge such as bounded range, symmetry and monotonicity, we can use the sample data to train the consequent parameters of the UIT2FLS when the antecedent parameters are determined beforehand. From (30), according to our objective, the UIT2FLS design problem can be transformed into the following constrained least squares optimization problem (31). Therefore, we can solve (31) to design the UIT2FLS utilizing the multi-source knowledge which includes the information from both sample data and prior knowledge.

\[
\begin{cases}
\min_w (\Phi w - y)^T(\Phi w - y) \\
\text{subject to } w \in \Omega
\end{cases},
\] (31)

where \( \Omega \) represents the constrained feasible parameter space.

Next, we will show that how to transform different kinds of prior knowledge description into the constraints of the consequent parameters of UIT2FLSs.

B. Constraints for Bounded Range

The constraints on the consequent interval weights of bounded UIT2FLSs in Theorem 3.1 can be expressed as
\[
\frac{1}{2}(\underline{w}_{\min} w^i + \overline{w}_{\min} \overline{w}^i) \geq \overline{b},
\]
\[
\frac{1}{2}(\underline{w}_{\max} w^i + \overline{w}_{\max} \overline{w}^i) \leq \underline{b},
\]
where
\[
\overline{w}_{\min} = \min_{p=1,\ldots,M-1} \{\underline{w}_{p}\},
\]
\[
\overline{w}_{\min} = \min_{p=1,\ldots,M-1} \{\overline{w}_{p}\},
\]
\[
\underline{w}_{\max} = \max_{p=1,\ldots,M-1} \{\underline{w}_{p}\},
\]
\[
\overline{w}_{\max} = \max_{p=1,\ldots,M-1} \{\overline{w}_{p}\},
\]
\[
i = 1, 2, \ldots, M.
\]

Therefore, there exist the \( 2M \) linear inequality constraints. It is convenient and brief to formulate these constraints as the following linear matrix inequality:
\[
\begin{bmatrix}
-\frac{1}{2} \underline{w}_{\min} I_M & -\frac{1}{2} \overline{w}_{\min} I_M \\
\frac{1}{2} \underline{w}_{\max} I_M & \frac{1}{2} \overline{w}_{\max} I_M
\end{bmatrix} w \leq \begin{bmatrix} -\overline{b} \\ \overline{b} \end{bmatrix},
\] (32)

where \( I_M \) is the \( M \times M \) identity matrix, \( \overline{b} = [\overline{b}, \ldots, \overline{b}]^T \in \mathbb{R}^M \), and \( \overline{b} = [\overline{b}, \ldots, \overline{b}]^T \in \mathbb{R}^M \).

C. Constraints for Symmetry

The constraints on the consequent parameters of symmetric UIT2FLSs in Theorem 3.2 can be expressed as
\[
\begin{align*}
\underline{w}^i - \overline{w}^{i+1} &= 0, \\
\overline{w}^i - \overline{w}^{i+1} &= 0, \quad (i = 1, \ldots, M-1), \\
\underline{w}^j - \overline{w}^j &= 0, \quad (j = 1, 2, \ldots, M).
\end{align*}
\]

Hence, there exist \( M-1 \) equality and \( M \) inequality constraints for these parameters, and the matrix form of the constraints can be expressed as
\[
\begin{bmatrix}
I_{M-1} & 0_{M-1 \times 1} & -I_{M-1} \\
0_{M-1} & I_{M-1} & 0_{M-1 \times 1} \\
0_{M-1} & 0_{M-1 \times 1} & I_{M-1} \\
0_{M-1} & 0_{M-1 \times 1} & -I_{M-1}
\end{bmatrix}
\times w = 0_{M-1 \times 1},
\] (33)
\[
[I_M - I_M] w \leq 0_{M \times 1},
\] (34)

where \( I_M \) denotes the \( M \times M \) matrix obtained by rotating \( I_M \) 90 degrees clockwise.

D. Constraints for Monotonicity

The constraints on the consequent parameters of monotonically increasing UIT2FLSs on \( U \) in Theorem 3.3 can be expressed as
\[
\begin{align*}
\underline{w}^i - \overline{w}^{i+1} &\leq 0, \\
\overline{w}^i - \overline{w}^{i+1} &\leq 0, \quad (i = 1, 2, \ldots, M-1), \\
\underline{w}^j - \overline{w}^j &\leq 0, \quad (j = 1, 2, \ldots, M).
\end{align*}
\]

Hence, there exist \( 3M - 2 \) inequality constraints for these parameters, and the matrix form of the constraints can be expressed as
\[
\begin{bmatrix}
Y_{11} & 0_{(M-1) \times M} \\
0_{(M-1) \times M} & Y_{22} \\
I_M & -I_M
\end{bmatrix} w \leq 0_{(3M-2) \times 1},
\] (35)

where
\[
Y_{11} = \min_{p=1,\ldots,M-1} \{Y_{p}\},
\]
\[
Y_{22} = \min_{p=1,\ldots,M-1} \{Y_{p}\},
\]
\[
\underline{w}_{\max} = \max_{p=1,\ldots,M-1} \{\underline{w}_{p}\},
\]
\[
\overline{w}_{\max} = \max_{p=1,\ldots,M-1} \{\overline{w}_{p}\},
\]
\[
i = 1, 2, \ldots, M.
\]
where
\[
Y_{11} = Y_{22} = \begin{bmatrix}
1 & -1 & 0 & \ldots & 0 & 0 \\
0 & 1 & -1 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & 1 & -1
\end{bmatrix} \in \mathbb{R}^{(M-1) \times M}.
\]

In a similar way, based on Theorem 3.4, there exist \(3M-2\) inequality constraints for the consequent parameters, and when \(\text{UIT2FLS}\) monotonically decrease, the matrix form of the constraints can be expressed as
\[
\begin{bmatrix}
Z_{11} \\
0_{(M-1) \times M} \\
Z_{22}
\end{bmatrix} \leq \begin{bmatrix}
0_{(M-1) \times M} \\
M_{1} \\
-M_{1}
\end{bmatrix},
\]
where \(Z_{11} = Z_{22} = -Y_{11} \in \mathbb{R}^{(M-1) \times M}.

**Remarks:** Assume there exists a function \(y = f(x)\), which is monotonically decreasing when \(x < 0\), and monotonically increasing when \(x > 0\). If the input variable \(x\) is partitioned as shown in Fig. 5, then we can deal with the circumstances as follows:
\[
\begin{bmatrix}
ZY \\
Z_{11} \\
Z_{22}
\end{bmatrix} \leq \begin{bmatrix}
0_{(M-1) \times M} \\
M_{1} \\
-M_{1}
\end{bmatrix},
\]
where
\[
ZY = \begin{bmatrix}
Z_{11} \\
0_{M \times M} \\
Z_{22}
\end{bmatrix} \in \mathbb{R}^{(M-1) \times M}.
\]
where \(Z_{11}, Y_{11} \in \mathbb{R}^{M \times M}\), and \(M\) is odd number.

V. SIMULATION

In this section, we use a simulation example to illustrate the advantages based on multi-source knowledge, as above mentioned, when the output of a target function is corrupted by white noise.

Consider the following SISO nonlinear function
\[
y = \left(\frac{x}{3}\right)^2 \tanh(|x|),
\]
where \(x \in U = [-3, 3]\). It is obvious that the function is even symmetry with respect to \(x = 0\), bounded on interval \([0, 1]\), monotonically decreases on interval \([-3, 0]\), and monotonically increases on interval \([0, 3]\).

We can create 500-training-data set \(D\) whose elements \((x_i, y_i) (i = 1, \ldots, 500)\) are generated by \(\tilde{y} = y + \text{Noise}\), where \(\text{Noise} \in [-n_b, n_b]\) is the uniformly distributed additive noise. From Equation (38), we can obtain 500 input-output data pairs \((x_i, y_i) (i = 1, \ldots, 500)\) which are the elements of test data set \(E\). In this simulation, we test three different levels of noisy disturbance including \(n_b = 20\%, 30\%, 40\%\). For each level of noisy disturbance, five cases are considered. In case \(k (k = 1, \ldots, 5)\), the elements of training data subset \(D_k\) and test data subset \(E_k\) are randomly chosen from \(D\) and \(E\), respectively. The numbers of the training data subset and the test data subset \(r_k\) are 25, 40, 60, 80, 100, and all the values of input variable \(x_i\) in \(D_k\) and \(E_k\) are different in each case \(k\).

In order to demonstrate the superiority of prior knowledge and type-2 FLSs, we use the following six FLSs to identify the target function: \(\text{UIT2FLS}\) with the monotonic, symmetric and bounded constraints (\(\text{T2PFLS}\)), \(\text{UIT2FLS}\) without the constraints (\(\text{T2NPFLS}\)), unnormalized type-1 FLS (\(\text{T1FLS}\)) with the constraints (\(\text{T1PFLS}\)), \(\text{T1FLS}\) without the constraints (\(\text{T1NPFLS}\)), normalized interval type-2 FLS (\(\text{NT2FLS}\)) with the constraints (\(\text{NT2PFLS}\)) and NT2FLS without the constraints (\(\text{NT2NPFLS}\)). For each FLS, we use seven membership functions equally distributed on input domain, and the fuzzy partitions which are shown in Fig. 7 satisfy the conditions on the parameters of the antecedent parts in Theorem 3.1 - 3.4.

Based on the training data and the test data, we use the following two RMSE performance indices in case \(k\),
\[
ap_k = \left(\frac{1}{r_k} \sum_{i=1}^{r_k} (\tilde{y}(x_i) - \tilde{y}_k)^2\right)^{\frac{1}{2}},
\]
\[
gp_k = \left(\frac{1}{r_k} \sum_{i=1}^{r_k} (\tilde{y}(x_i) - y_k)^2\right)^{\frac{1}{2}},
\]
where \(ap_k\) can reflect the approximation performance of the fuzzy systems for the training data, and \(gp_k\) can reflect the generalization performance of the fuzzy systems for the test data. From (39) and (40), we know that the less both indices are, the better the corresponding performances are. Furthermore, in order to avoid particularity, we calculate the statistical indices of the approximation performance and the generalization performance, \(ap_{sk}\) and \(gp_{sk}\), which are the arithmetic mean of the corresponding indices obtained from 50 runs.

As mentioned above, three different levels of noise and five cases for each level of noise are considered. The comparisons between the six FLSs are drawn with respect to the size of data subsets and the noise level for the statistical RMSEs of the approximation performance and the generalization performance, \(ap_{sk}\) and \(gp_{sk}\), which are shown in Fig. 8, Fig. 9 and Fig. 10, respectively. When the training data subset is comprised of 25 sample data, and the noise distributes uniformly in \([-40\%, 40\%]\), we obtain one of the identification results shown in Fig. 11. We can note that the data shown in Fig. 11 are consistent with the corresponding data in Fig.10.
From these figures, we can compare and analyze the approximation performances and the generalization performances in terms of both the size of the data subsets and the level of noise for the six FLSs as follows:

1. About approximation performance, from Fig. 8(a), Fig. 9(a) and Fig. 10(a), we observe that the unnormalized interval type-2 FLS without the prior knowledge constraints (T2NPFLS) performs better than the other five FLSs listed in the descending order of approximation performance, in general, as follows: NT2PFLS, T1NPFLS, T2PFLS, NT2PFLS and T1PFLS. Also, it is can be observed that the FLSs without the constraints have better approximation capabilities than the FLSs with the constraints, whose performance indices are almost the same. There exist the following two reasons to explain this. For one thing, the consequent parameters of T2PFLS, NT2PFLS and T1PFLS are not constrained by the prior knowledge, so the three FLSs have a larger degree of freedom than the counterpart of T2PFLS, NT2PFLS and T1PFLS. For another, type-2 FLSs have more parameters than type-1 FLSs. Thus interval type-2 FLSs without constraints have best approximation capabilities among the six FLSs. In terms of the size of a training data subset, note that the smaller the size of the subset is, the better the approximation performance is in each case. As the level of noise increases, the approximation performances deteriorate for the six FLSs in each case.

2. About generalization performance, from Fig. 8(b), Fig. 9(b) and Fig. 10(b), we observe that the unnormalized interval type-2 FLS with the prior knowledge constraints (T2PFLS) outperforms the other five FLSs listed in the descending order of generalization performance, in general, as follows: NT2PFLS, T1PFLS, T1NPFLS, T2NPFLS and NT2NPFLS. Also, it is can be observed that the FLSs with the constraints, whose performance indices are different slightly, have better generalization capabilities than the FLSs without the constraints. The reason for this is that the prior knowledge makes the FLSs with the constraints not be over-fitted to the target function so that their feasible parameters spaces are smaller than the counterpart of the FLSs without constraints. Thus, to some extent, T2PFLS, NT2PFLS and T1PFLS can overcome the effect of noise to approach better the original noise-free function. In terms of the size of a test data subset, note that the bigger the size of the subset is, the better the generalization performance is in each case. As the level of noise increases, the generalization performances deteriorate for the six FLSs in each case.

3. About T2PFLS, NT2PFLS and T1PFLS, from Fig. 8 – Fig. 10, we observe that the performance curves of the three FLSs are in close proximity both for approximation capability and for generalization capability, and as the level of noise increases, they become closer. The following reasons can explain the fact. Since the consequent weights of the three kinds of FLSs are constrained identically, the performances of T2PFLS and NT2PFLS are almost the same as that of T1PFLS. However, the interval type-2 FLSs are formed by the linear combination of the two type-1 bounding FLSs [3]. This implies that type-2 FLSs have more parameters and more freedom degrees than type-1 FLSs. Thus, the approximation performances and the generalization performances of the two kinds of FLSs are better slightly than the counterpart of T1PFLS.

4. From Fig. 8 – Fig. 10, we observe that unnormalized interval type-2 FLSs perform slightly better than normalized interval type-2 FLSs, and as the level of noise increases, the difference between the unnormalized T2FLSs and the normalized T2FLSs became smaller and smaller. On the other hand, since unnormalized T2FLSs do not need the operation of normalization, i.e., division operation, its elapsed time is much shorter than that of normalized T2FLSs. Here, elapsed time of a type-2 FLS is defined as the time spent on computing the output \( \hat{y}(x) \) of the type-2 FLS for a data subset. Without loss of generality, in the simulation, we use arithmetic means of all the elapsed time obtained from 50 runs. For each case \( k \) of T2NPFLS and NT2NPFLS, the average values of the elapsed time are shown in Table I. From the table, we can know that if
The FLSs with prior knowledge description have better simulation results have verified the validity of Theorem 3.1 monotonicity can be incorporated into UIT2FLSs. And then, that the prior knowledge of bounded range, symmetry and we utilize UIT2FLS, we would save about 11% of the elapsed time of normalized IT2FLS.

In conclusion, based on multi-source knowledge, not only can the UIT2FLS obtain better performance than unnormalized type-1 FLSs on the whole, but also the UIT2FLS slightly outperform and need less computation time than normalized IT2FLS.

**VI. CONCLUSIONS**

In this paper, we give some sufficient conditions on ensuring that the prior knowledge of bounded range, symmetry and monotonicity can be incorporated into UIT2FLSs. And then, we design an SISO zeroth-order UIT2FLS for noisy regression problems via constrained least squares algorithm. The simulation results have verified the validity of Theorem 3.1 – 3.4. The FLSs with prior knowledge description have better performances than the FLSs without prior knowledge description on the whole. In contrast with T1PFLS, T2PFLS have better overall performance for approximation capability and generalization capability. In contrast with NT2PFLS, T2PFLS can save more computation time, and have slightly better performance. Therefore, we can conclude that based on multi-source knowledge, the UIT2FLS performs best among all the six fuzzy logic systems.

**ACKNOWLEDGMENT**

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

**REFERENCES**


