Universal Approximation of TS Fuzzy Systems Constructed Dynamically–MISO Cases

ShiYu Yan, ZengQi Sun and Zhe Li

Abstract—Study on approximation capabilities of fuzzy systems has been an important and hot issue for many years. Many existing studies on fuzzy systems based on standard affine TS fuzzy models suffer seriously from curse of dimensionality though such models are relatively simple to construct; whereas other investigations based on hierarchical fuzzy models have to face the difficulty of relatively more complexity. Through utilizing the dynamic partition along all dimensions, this paper proposes a novel dynamically constructive method in MISO cases to construct fuzzy systems based on standard affine TS fuzzy models, which have simple structures and also can ease the serious curse of dimensionality greatly. Correspondingly, new sufficient conditions for a affine TS fuzzy system as an universal approximator are obtained. The situation in 2ISO cases is derived first which is then extended to general MISO cases. Theoretical analyses comparing to some typical methods and several numerical examples all confirm that the dynamic method in this paper can reduce the fuzzy rules number dramatically and therefore can ease the serious curse of dimensionality to some extent. Some conclusions and discussions on further work are also given.

I. INTRODUCTION

Universal approximation of fuzzy models has been studied greatly in the past decades, which can provide the theoretical basis for approximating actual targets by corresponding fuzzy models and guide people to establish the fuzzy models more properly and efficiently.

Because of its simplicity and usefulness for development of systematic approaches to stability analysis and controller design of fuzzy control systems via powerful conventional control theory and techniques, the standard affine TS fuzzy model has attracted a lot of researchers’ attentions and efforts including the studies on universal approximation of affine TS fuzzy models [1]-[13]. References [1]-[4] first gave the “existence proof”. Then the “constructive proof” issue, i.e., "given a real function, how to construct a TS fuzzy model that can TS fuzzy model that can approximate it with given accuracy", was studied by some scholars [5], [7], [9] -[12] with some drawbacks more or less. The rule consequent was assumed as the proportional affine function of input variables in [5], whereas [9] used the simplified fuzzy systems. Moreover, [7], [10]-[12] gave constructive proofs for the common TS fuzzy systems where each partition of the input space indicates a rule. However, almost all these above fuzzy approximation schemes suffered seriously from the curse of dimensionality[14], which greatly limited their applications in high-dimensional inputs cases.

To deal with the curse of dimensionality to some degree, some articles proposed another kind of method, clustering or scatter partitioning, to reduce fuzzy rules[15]-[19]. Since such method places the rules wherever they are useful or convenient to perform the modeling, the number of fuzzy rules needed is much lower. Nevertheless, it is the same reason that makes its coverage of input space is not thorough[20]. By offering another angle, hierarchical fuzzy model has been studied deeply and widely[21]-[28] since the initial work[22] and [23] in early 1990s. [24], [25], [26] all studied the approximation capabilities of distinct hierarchical fuzzy models and obtained the different results, respectively, which were better than that based on standard affine TS fuzzy model described above at the cost of either considerable free parameters or a great number of hierarchical layers. As an improvement, [21] proposed a novel method by introducing the concept of the natural hierarchical structure and obtained new achievements involved in much fewer parameters and number of hierarchical layers than that in previous work. Recently, [28] derived the new conclusions in discrete input space based on the flexibility of hierarchical representation, the original discovery in that paper. However, all these results based on hierarchical fuzzy models have to face the relatively more complexity than that based on standard affine TS fuzzy models when modeling because of its hierarchy.

Therefore, it should be an interesting and significant work to explore the much simpler fuzzy systems based on the standard affine TS fuzzy models, to approximate any given real continuous function defined on a compact set within any given approximation accuracy as well as easing the serious curse of dimensionality to some extent. Actually, the reason why most previous fuzzy approximation schemes based on standard affine TS fuzzy models would face such serious problem is the static partition of premise space, which leads to the same size of partitions either along different dimensions or in distinct parts inside the premise space. It is convenient to theoretical inference but with no ground in practice. As pointed out by [29], “the size of each partition is dependent only on the global property of the function to be approximated, that is, unrelated to its local property in the very region.” Thus, there are much less fuzzy rules
needed in [29] than that in [10]-[12] by using dynamically constructive method for partitioning in SISO cases. However, [29] failed to consider the MISO cases because it is not a straight extension. Most recently, [30] offered an attempt to extend [29] to MISO cases by using combined method with dynamic partition only along one dimension. Although the fuzzy system in [30] only needed nearly a half of number of rules comparing to system in [12] in MISO cases, the work [30] still should be further improved in that the method in [30] was not a really total dynamic one, since such method employed dynamic partition along only one dimension while other dimensions were still partitioned statically.

It is thus obviously necessary to further study the universal approximation of fuzzy systems based on the standard affine TS fuzzy models. The main contribution of this paper is to propose a novel totally dynamically constructive method to build certain fuzzy systems based on standard affine TS fuzzy models to further ease the curse of dimensionality greatly, which leads to new sufficient conditions for a affine TS fuzzy system as an universal approximator. According to [20], the dynamic method in this paper is neither a grid partitioning one nor a clustering one because on the one hand it does not use common fuzzy sets shared by all rules, but, with thorough space coverage on the other hand. Moreover, since the standard affine TS model is employed, we can use the very simple form for approximation with greatly much fewer fuzzy rules than that in former similar results based on standard affine TS models in MISO cases. When the function is second order differentiable, its TS model can be constructed directly. If not, certain polynomial function will be the bridge between the TS model and the original function by employing Weierstrass approximation theorem[31]. The theoretical reasoning, comparative studies and several numerical examples all illustrate that the number of fuzzy rules needed in Zeng’s model[12] or in Yan’s model[30] is nearly $2^n$ times or $2^{n-1}$ times of that in the dynamic model in this paper respectively.

II. MATHEMATICAL NOTATIONS

To express clearly, some important notations are illustrated here. Throughout this paper, we always use these notations without explanation.

1) $X$ is two-dimensional or multi-dimensional, $h(X)$ is any given bivariate or multivariate polynomial defined on a compact set; $F(X)$ is any given real continuous function(bivariate or multivariate one) defined on a compact set; $f(X)$ is the total output of the affine TS fuzzy system obtained by the dynamically constructive method to approximate $h(X)$ or $F(X)$ within any given approximation accuracy $\varepsilon$.

2) $||*||_U$ is defined as $||A(X)||_U = \sup_{X \in U}|A(X)|$.

3) $D^T_{h U} = \left|\begin{array}{c} \frac{\partial f(X)}{\partial x_1} |_{X^*} U \\ \frac{\partial f(X)}{\partial x_2} |_{X^*} U \end{array}\right|$, $D_{ij} = \left|\begin{array}{c} \frac{\partial^2 f(X)}{\partial x_i \partial x_j} |_{X^*} U \end{array}\right|$. 

4) If space $C$ is divided into several sub-spaces, $S_k^C$ represents the $kth$ sub-space of $C$.

5) The first order of Taylor expansion of a bivariate function $g(X) = g((x_1, x_2))$ at point $X^* = (x_1^*, x_2^*)$ is:

$$g(X^*) + \nabla g^T_{X=X^*}(X-X^*)$$

Or

$$g(X^*) + \frac{\partial g(X^*)}{\partial x_1}(x_1-x_1^*) + \frac{\partial g(X^*)}{\partial x_2}(x_2-x_2^*)$$

The corresponding Lagrange remainder is:

$$\frac{1}{2}(X-X^*)^T \nabla^2 g_{X=X^*}((X-X^*)^2)$$

Or

$$\frac{1}{2} \left[ \frac{\partial}{\partial x_1}(x_1-x_1^*) + \frac{\partial}{\partial x_2}(x_2-x_2^*) \right]^2 g(X)$$

Or

$$\frac{1}{2} \left( \frac{\partial^2 g(X)}{\partial x_1}(x_1-x_1^*)^2 + 2\frac{\partial^2 g(X)}{\partial x_1 \partial x_2}(x_1-x_1^*)(x_2-x_2^*) + \frac{\partial^2 g(X)}{\partial x_2^2}(x_2-x_2^*)^2 \right)$$

where $X^*$ locates in the minimum super sphere containing $X$ and $X^*$.

6) The first order of Taylor expansion of a multivariate function $g(X) = g((x_1, x_2, \ldots, x_n))$ at point $X^* = (x_{1^*}, x_{2^*}, \ldots, x_{n^*})$ is:

$$g(X^*) + \nabla g^T_{X=X^*}(X-X^*)$$

Or

$$g(X^*) + \sum_{i=1}^{n} \frac{\partial g(X^*)}{\partial x_i}(x_i-x_i^*)$$

The corresponding Lagrange remainder is:

$$\frac{1}{2}(X-X^*)^T \nabla^2 g_{X=X^*}((X-X^*)^2)$$

Or

$$\frac{1}{2} \sum_{i=1}^{n} \left( \frac{\partial^2 g(X)}{\partial x_i^2}(x_i-x_i^*)^2 + \sum_{j=1}^{n} \frac{\partial^2 g(X)}{\partial x_i \partial x_j}(x_i-x_i^*)(x_j-x_j^*) \right)$$

where $X^*$ locates in the minimum super sphere containing $X$ and $X^*$.

III. DYNAMICALLY CONSTRUCTIVE METHOD FOR AFFINE TS FUZZY SYSTEMS TO APPROXIMATE A POLYNOMIAL

To express clearly, we first take the two-dimensional situation(2ISO) as the example which is then extended to MISO cases.

A. Illustration of Dynamically Constructive Method–2ISO

The basic idea of our dynamically constructive method is to build the proper affine TS fuzzy system to approximate a polynomial $h(X)$ within the desired approximation accuracy by partitioning the premise space dynamically according to the local properties of the polynomial in each dimension. In the dynamically constructive method, there are two main parts which will be described in detail in the following two sub-subsections 1), 2).

1) Construct affine TS Fuzzy Model Based on Dynamic Partition in Premise Space: Here, we assume that the entire two-dimensional premise space has been dynamically divided into $m$ sub-spaces $U$. As for how to partition the premise space dynamically, it will be discussed in the next sub-subsection.

Let each sub-space represent a fuzzy rule, we construct a affine TS fuzzy model. All rules have the same form and the $i$th rule as an example is given below:
Rule i: IF \( X = (x_1, x_2) \) is \( A_i \), THEN
\[
f_i(X) = h(X_i^T) + \nabla h|_{X = X_i}(X - X_i^T)
= h(X_i^T) + \frac{\partial h(X_i^T)}{\partial x_1}(x_1 - x_i^1) + \frac{\partial h(X_i^T)}{\partial x_2}(x_2 - x_i^2),
\]
which is the first order of Taylor expansion of \( h(X) \).

where \( A_i \) represents the fuzzy set of the \( i \)th sub-space \( S_i^U \) and its membership function is defined as \( A_i(X) = \left\{ \begin{array}{ll} 0, & X \notin S_i^U \\ 1, & X \in S_i^U \end{array} \right. ; f_i(X) \) is the output of TS fuzzy model in the \( i \)th rule; \( X_i^T = (x_i^1, x_i^2) \) is the central point of the \( i \)th sub-space \( S_i^U \).

Thus, the corresponding Lagrange remainder is:
\[
h(X) - f_i(X) = \frac{1}{2} \frac{\partial^2}{\partial x_1^2}(x_1 - x_i^1)^2 + \frac{\partial^2}{\partial x_2^2}(x_2 - x_i^2)^2
= \frac{1}{2} \frac{\partial^2}{\partial x_1^2}h(X_i)(x_1 - x_i^1)^2 + \frac{\partial^2}{\partial x_2^2}h(X_i)(x_2 - x_i^2)^2
+ \frac{1}{2} \frac{\partial^2}{\partial x_1^2}h(X_i)(x_1 - x_i^1)^2 + \frac{\partial^2}{\partial x_2^2}h(X_i)(x_2 - x_i^2)^2
\]

where \( X \) locates in the minimum super sphere containing \( X \) and \( X_i^T \).

Remark 1: Obviously, the membership function \( A_i(X) \) is rigid, which means that the total output \( f(X) \) of affine TS fuzzy system is equal to the local output \( f_i(X) \) of that fuzzy system when the \( i \)th rule is fired, i.e., \( X \in S_i^U \). Also, the approximation error \( h(X) - f(X) \) is equal to Lagrange remainder (1) in \( S_i^U \). It is the so-called "Switching Fuzzy System" but not "General Fuzzy System". The improvement of softening such membership function will be further considered in the future work.

2) Dynamically Constructive Method for Partition in Premise Space: In this sub-subsection, we will show how to partition the premise space using dynamically constructive method, which is the most important step in the method.

The overall idea is to through many times of iteration, divide the entire premise space continuously and dynamically into a series of sub-spaces where the given approximation accuracy can be satisfied, i.e., the final partitions, according to the local properties of polynomial or function to be approximated. There are three important questions during such dynamic partition: \( i \) how to check whether a space or sub-space (E) can meet the given approximation accuracy, i.e., \( \forall X \in E, |h(X) - f(X)| \leq \varepsilon \); \( ii \) if not, how to find a proper sub-space \( S \) which satisfies the given accuracy contained in space \( E \) from the bottom left vertex of \( E \); \( iii \) based on \( S \), how to divide \( E \) into several sub-spaces, i.e., \( E = S \cup E/S \).

(Remark: \( E/S \) denotes the set of other sub-spaces except \( S \) and some or all these sub-spaces in \( E/S \) may not satisfy the given accuracy, need further partition)

\( i \) For the first question above, denote \( E = E_1 \times E_2 = [E_{1f}, E_{1b}] \times [E_{2f}, E_{2b}] \) where \( (x_{1f}, x_{2f}) \) is the central point, let \( |E_{1f} - E_{1b}| = q_1^f, |E_{2f} - E_{2b}| = q_2^f \) \( \forall X = (x_1, x_2) \in E \), \( |x_1 - E_{1f}| = d_1^f, |x_2 - E_{2f}| = d_2^f \) \( d_1^f, d_2^f \) are functions of \( x_1, x_2 \), then \( |x_1 - x_{1f}| = |d_1^f - d_1^f|, |x_2 - x_{2f}| = |d_2^f - d_2^f| \)

and from (1),
\[
|h(X) - f(X)| \leq \frac{1}{2} D_1^f (q_1^f)^2 - d_1^f)^2 + D_2^f (d_2^f)^2 \leq \varepsilon
\]

Therefore, if \( \frac{D_1^f}{q_1^f} (q_1^f)^2 + \frac{D_2^f}{q_2^f} (q_2^f)^2 \leq \varepsilon \), then such space \( E \) can meet the given approximation accuracy; otherwise, for prudence, we regard space \( E \) as the one needed further partition.

\( ii \) For the second question above, suppose \( S = S_1 \times S_2 = [E_{11}, S_{1b}] \times [E_{21}, S_{2b}], \) what we have to do is to determine the proper \( S_{1b} \) and \( S_{2b} \). Considering that the value of \( (E_{1f}, E_{2f}) \) is known, we just need to determine \( |E_{1f} - S_{1b}| \) and \( |E_{2f} - S_{2b}| \).

Similar to the solution to first question, let \( |E_{1f} - S_{1b}| = q_1^s, |E_{2f} - S_{2b}| = q_2^s \), \( \forall X = (x_1, x_2) \in E, \) \( |x_1 - E_{1f}| = d_1^s, |x_2 - E_{2f}| = d_2^s \) then \( q_1^s, q_2^s \) have to get the following inequality hold:
\[
\frac{D_1^s}{q_1^s} (q_1^s)^2 + \frac{D_2^s}{q_2^s} (q_2^s)^2 \leq \varepsilon.
\]

Besides, we should find the maximum value of \( (q_1^s, q_2^s) \), which in general, can minimize the number of final partitions.

Moreover, \( q_1^s, q_2^s \) obviously should satisfy the following inequalities:
\[
q_1^s \leq q_1^f \quad \text{and} \quad q_2^s \leq q_2^f
\]

In short, we can find the solution to the second question by solving the following constrained optimization problem, which can be easily solved by Matlab:
\[
\text{Max } q_1^s \cdot q_2^s
\text{ s.t. } \frac{D_1^s}{q_1^s} (q_1^s)^2 + \frac{D_2^s}{q_2^s} (q_2^s)^2 \leq \varepsilon
q_1^s \leq q_1^f, q_2^s \leq q_2^f
\]

\( iii \) For the third question above, every time, we divide \( E \) based on \( S \) from the point of \( (E_{1f}, E_{2f}) \) like the following way:
\[
E = E_1 \times E_2 \Rightarrow \begin{cases}
S = S_1 \times S_2 & \\
R^1 = S_1 \times E_2 & \\
R^2 = S_2 \times E_1
\end{cases}
\]

where \( S \cup S_1 = E_1, S \cup S_2 = E_2 \).

Obviously, in real cases, \( E \) can be divided into the above 3 sub-spaces at most, shown in (a) in Fig.1. Nevertheless, when \( S_1 = E_1 \), there are only 2 sub-spaces, \( S \) and \( R^1 \) but no \( R^2 \), shown in (b) in Fig.1; when \( S_2 = E_2 \), there are only 2 sub-spaces, \( S \) and \( R^1 \) but no \( R^2 \), shown in (c) in Fig.1.
To sum up, such total process of dynamic partition can be described by the following algorithm (Dynamic Partition Algorithm–2ISO). There are some illustrations involved in algorithm in advance:

Without loss of generality, assume the entire premise space is represented by \( U = [a, b]^2 \).

The queue \( Q \) in the following algorithm is used to store all the sub-spaces to be checked or partitioned contained in the space \( U \). When \( Q \) becomes empty, the process of dynamic partition is over and we can obtain the final partition result.

\[
Q^h = Q^h_1 \times Q^h_2 = [Q^h_{11}, Q^h_{12}] \times [Q^h_{21}, Q^h_{22}] \text{ represents the first element of } Q. \]

During each iteration in algorithm, the current \( Q^h \) in \( Q \) is checked, and saved as one of the final partitions in premise space if \( Q^h \) satisfies the given accuracy; otherwise, \( Q^h \) is further divided. It should be noted that \( Q^h \) as the first element of \( Q \), is always renewed and changed when passing each iteration in algorithm.

During each iteration, \( Q^t = Q^t_1 \times Q^t_2 = [Q^t_{11}, Q^t_{12}] \times [Q^t_{21}, Q^t_{22}] \) always represents the desired result of partitioning the current \( Q^h \), i.e., the sub-space which meets the given accuracy. Obviously, each \( Q^t \) is one of the final partitions.

\( Q^h/Q^t \) represents the set of other sub-spaces except \( Q^t \) after partitioning \( Q^h \) in each iteration. That is, \( Q^h = Q^t \cup Q^h/Q^t \).

**Dynamic Partition Algorithm–2ISO**

**Initial:** Build a new empty queue \( Q \) and add \( U \) into \( Q \). Now, \( Q^h_{1f} = Q^h_{2f} = a, Q^h_{1b} = Q^h_{2b} = b \).

**Loop:**

while \( Q \) is not empty do

if \( \forall X \in Q^h, |h(X) - f(X)| \leq \varepsilon \) then

i) Save \( [Q^h_{11}, Q^h_{12}] \times [Q^h_{21}, Q^h_{22}] \) as one of the final partitions in premise space, and remove such first element from \( Q \).

ii) Renew the first element of \( Q \), if exists.

end if

else

i) Using the aforementioned solution for the second question, i.e., solving (3), find the proper sub-space \( Q^t \) contained in space \( Q^h \) from the point \((Q^t_{11}, Q^t_{21}), (Q^t_{12}, Q^t_{22})\), in which the given approximation accuracy can be satisfied.

ii) By the aforementioned solution for the third question, divide \( Q^h \) into several sub-spaces (including \( Q^t \)) based on \( Q^t \).

iii) Save \( Q^t \) as one of the final partitions in premise space, and remove current first element \( Q^h \) from \( Q \).

iv) Add \( Q^h/Q^t \) into the front of \( Q \) and renew the first element of \( Q \).

end if

end while

An example of the process of dynamic partition in premise space is shown in Fig.2.

**B. Extend Dynamically Constructive Method to MISO cases**

In this sub-section, we directly present the extension result from 2ISO case to MISO case (n-dimensional). The similar derivation is omitted due to the space limitation.

1) **Construct affine TS Fuzzy Model Based on Dynamic Partition in Premise Space:**

Rule \( i: \) IF \( X = (x_1, \ldots, x_n) \) is \( A_i \),

THEN \( \sum f_i(X) = h(X_i^0) + \sum_{X \to X_i^0} (X - X_i^0) = h(X_i^0) + \sum_{j=1}^n \frac{\partial f_i(X_i^0)}{\partial x_j} (x_j - x_{i,j}^0) \),

which is the first order of Taylor expansion of \( h(X) \).

\[
h(X) - f_i(X) = \frac{1}{2} \left( \frac{\partial^2 f_i(X_i^0)}{\partial x_j^2} (x_j - x_{i,j}^0) \right)^2 + \sum_{j=1}^{n-1} \sum_{j > j} \frac{\partial^2 f_i(X_i^0)}{\partial x_j \partial x_k} (x_j - x_{i,j}^0)(x_k - x_{i,k}^0),
\]

where \( X_i^0 \) locates in the minimum super sphere containing \( X \) and \( X_i^0 \).

2) **Dynamically Constructive Method for Partition in Premise Space:**

Let \( E = \prod_{i=1}^n [E_{i1}, E_{i2}] \) where \((x_{E1}, \ldots, x_{En})\) is the central point. For \( i = 1, \ldots, n, |E_{ij} - E_i| = q^E_i, \forall X = (x_1, \ldots, x_n) \in E, |x_i - E_{ij}| = d_i \), then \( |x_i - x_{Ei}| = \frac{q^E_i}{2} - d_i \) and from (4),

\[
|h(X) - f(X)| \leq \sum_{i=1}^n \frac{d_i}{8} (q^E_i)^2 + \sum_{j=1}^{n-1} \sum_{j > j} \frac{d_i}{4} q^E_i q^E_j \leq \varepsilon
\]

When \( \sum_{i=1}^n \frac{d_i}{8} (q^E_i)^2 + \sum_{j=1}^{n-1} \sum_{j > j} \frac{d_i}{4} q^E_i q^E_j \leq \varepsilon \) holds, such space \( E \) can meet the given approximation accuracy; otherwise, for prudence, we regard space \( E \) as the one needed further partition.

Similar constrained optimization problem is the following one:

\[
\text{Max } \prod_{i=1}^n q^E_i
\]

s.t. \( \sum_{i=1}^n \frac{d_i}{8} (q^E_i)^2 + \sum_{j=1}^{n-1} \sum_{j > j} \frac{d_i}{4} q^E_i q^E_j \leq \varepsilon \)

\( q^E_i \leq q^E_i, (i = 1, \ldots, n) \)
The way we divide $E$ based on $S$:

$$E = \prod_{i=1}^{n} E_i$$

where $\overline{S_i} = E_i$, $i = 1, \ldots, n$.

Also, without loss of generality, assume that the entire premise space is represented by $U = [a, b]^n$.

$Q^h = \prod_{i=1}^{n} [Q_{i,j}^h, Q_{i,b}^h]$ represents the first element of $Q$. During each iteration, $Q^s = \prod_{i=1}^{n} Q_i^s = \prod_{i=1}^{n} [Q_{i,j}^s, Q_{i,b}^s]$ always represents the desired result of partitioning current $Q^h$.

$Q_{i,Q_i}^h$ represents the set of other sub-spaces except $Q_i^h$ after partitioning $Q_i^h$ in each iteration. That is, $Q^h = Q^s \cup Q_{i,Q_i}^h$.

**Dynamic Partition Algorithm—MISO**

**Initial:** Build a new empty queue $Q$ and add $U$ into $Q$. Now, $U$ is the only one and also the first element in $Q$, i.e.,

$$Q^h = \prod_{i=1}^{n} [Q_{i,j}^h, Q_{i,b}^h] = U = [a, b]^n.$$  

**Loop:**

while $Q$ is not empty do

if $\forall X \in Q^h$, $|h(X) - f(X)| \leq \varepsilon$ then

i) Save $\prod_{i=1}^{n} [Q_{i,j}^h, Q_{i,b}^h]$ as one of the final partitions in premise space, and remove such first element from $Q$.

ii) Renew the first element of $Q$, if exists.

else

i) Solve (6) to find the proper sub-space $Q_i^h$ contained in space $Q^h$ from the point $(Q_{i,j}^h, \ldots, Q_{i,b}^h)$, in which the given approximation accuracy can be satisfied.

ii) By the aforementioned method, divide $Q_i^h$ into several sub-spaces including $Q_i^s$, in which the given approximation accuracy can be satisfied.

iii) Save $Q_i^s$ as one of the final partitions in premise space, and remove current first element $Q_i^h$ from $Q$.

iv) Add $Q_{i,Q_i}^h$ into the front of $Q$ and renew the first element of $Q$.

end if

end while

**IV. NEW SUFFICIENT CONDITION FOR AFFINE TS FUZZY SYSTEM AS AN UNIVERSAL APPROXIMATOR**

Naturally, the following lemma is obvious.

**Lemma 1 (MISO):** For any given polynomial $h(X)$ defined on a compact set and approximation accuracy $\varepsilon > 0$, we can construct a TS fuzzy model by the dynamically constructive method in the above section. Its output is $f(X)$, so that $|h(X) - f(X)| \leq \varepsilon$.

Given Lemma 1 and Weierstrass Approximation Theorem, the following main result can be obtained.

**Theorem 1 (MISO):** For any given real continuous multivariate function $F(X)$ defined on a compact set and approximation accuracy $\varepsilon > 0$, we can construct a TS fuzzy model, whose output is $f(X)$, so that $|F(X) - f(X)| \leq \varepsilon$.

**Proof:** First, by Weierstrass Approximation Theorem, there exists a polynomial $h(X)$ such that $|F(X) - h(X)| \leq \varepsilon_1, \varepsilon_1 \leq \varepsilon$.

By Lemma 1, we can construct a TS model, satisfying $|h(X) - f(X)| \leq (\varepsilon - \varepsilon_1)$, where $f(X)$ is the output of TS model. Thus,

$$|F(X) - f(X)| \leq |F(X) - h(X)| + |h(X) - f(X)| \leq \varepsilon.$$ 

In particular, when the function needed to be approximated is second order differentiable, we can construct its TS model directly by the dynamically constructive method.

**Theorem 2 (MISO):** For any given real continuous multivariate function $F(X)$ defined on a compact set and approximation accuracy $\varepsilon > 0$, if $F(X)$ is second order differentiable, then directly using the dynamically constructive method in the above section, we can construct a TS fuzzy model, satisfying $|F(X) - f(X)| \leq \varepsilon$, where $f(X)$ is the output of TS fuzzy model.

**Proof:** Omission.

**V. COMPARATIVE STUDIES ON THE SUFFICIENT CONDITIONS**

In this section, both qualitative and quantitative comparative studies illustrate the much improvement of the dynamically constructive method in MISO cases.

**A. Qualitative Analysis and Explanation**

From [12], there are at least the following shortcomings which lead to the great number of superfluous rules:

1) The ways of partitioning along different dimensions are all the same and static;

2) The partition of each dimension is even and uniform, i.e., all the intervals along each dimension are the same;

3) The grid partitioning are employed, i.e., all the rules share the same fuzzy sets;

4) Only the globe properties but no local properties of function to be approximated are used when dividing premise space into sub-spaces.

The four aspects are all not sensible in real cases because generally the change tendencies of a real continuous function in different parts of input space or along different dimensions or even along the same dimension are definitely distinct.

Combining the static method in [12] and the dynamic method in [29], a combined constructive method was proposed in Yan’s work[30]. Such model only needs almost a half of fuzzy rules in Zeng’s model in MISO case. However, because its dynamic design is just employed in one dimension while other dimensions are partitioned statically, which definitely cause that there are still a lot of unnecessary fuzzy rules.

Since dynamic partition is employed in all dimensions, the dynamically constructive method here effectively utilizes different local properties of function in different local parts and overcomes the aforementioned four shortcomings. Therefore,
it is reasonable to believe that the dynamically constructive model will be much better than Zeng’s[12] and Yan’s[30].

B. Quantitative Comparative Studies

Let the polynomial to be approximated be \( h(X) \), input space be \([a, b]^n\), approximation accuracy be \( \varepsilon \).

According to [12], the interval length of each partition, i.e., each sub-space, along each dimension is \( L_{\text{Static}} = \frac{b-a}{n_0+1} \) and the number of sub-spaces is

\[
N^{r}_{\text{Static}} = \frac{(b-a)^n}{(n_0+1)^n} = (n_0 + 1)^n, \quad (7)
\]

The rule number needed in all is

\[
N^{r}_{\text{Static}} = (n_0 + 2)^n \geq N^{r}_{\text{Static}} = (n_0 + 1)^n, \quad (8)
\]

where \( n_0 \) is the same as its definition in [12] and then the following inequality about the relation between \( n_0 \) and \( \varepsilon \) holds according to [12]:

\[
\varepsilon \geq \frac{(b-a)^2 \sum_{i=1}^{n} \sum_{j=1}^{n} D_i^{[a,b]} D_j^{[a,b]} n_0 + 1}{2(n_0 + 1)^2} \quad (9)
\]

According to [30], the total needed fuzzy rules number is almost a half of that in Zeng’s model, that is

\[
N^{r}_{\text{Combined}} \approx \frac{1}{2} \times N^{r}_{\text{Static}} = \frac{(n_0 + 2)^n}{2}
\]

Now, we will check whether generally the interval length of sub-space along each dimension in the dynamic method can be twice of that in Zeng’s.

Suppose \( q_i^* = 2 \times L_{\text{Static}} = \frac{2(b-a)}{n_0+1}, (i = 1, \ldots, n) \) and \( Q^r \) is the corresponding sub-space. Then

\[
\sum_{i=1}^{n} D_i^{[a,b]} q_i^* q_j^* \leq \sum_{i=1}^{n} D_i^{[a,b]} q_i^* q_j^* \leq \frac{(b-a)^2}{2(n_0+1)^2} \sum_{i=1}^{n} \sum_{j=1}^{n} D_i^{[a,b]} D_j^{[a,b]} = \frac{(b-a)^2}{2(n_0+1)^2} \sum_{i=1}^{n} \sum_{j=1}^{n} D_i^{[a,b]} D_j^{[a,b]} \leq \varepsilon \quad (10)
\]

From (6) and (10), the above assumption that \( q_i^* = 2 \times L_{\text{Static}} = \frac{2(b-a)}{n_0+1}, (i = 1, \ldots, n) \) can satisfy the first inequality in constrained condition in (6).

Other hand, from the above description of the dynamically constructive method, the interval length along each dimension of sub-space is determined by solving a constrained optimization problem shown in (6) and generally the constrained conditions \( q_i^* \leq q_i^*, (i = 1, \ldots, n) \) always hold. Therefore, in most cases, the interval length along each dimension of sub-space in the dynamically constructive method \( L_{\text{dynamic}} \) can satisfy that \( L_{\text{dynamic}} \geq 2 \times L_{\text{Static}} = \frac{2(b-a)}{n_0+1} \).

Then the number of fuzzy rules in the dynamically constructive method, i.e., the number of sub-spaces, defined as

\[
N^{r}_{\text{dynamic}} \text{ satisfies the following inequalities:}
\]

\[
N^{r}_{\text{dynamic}} = \frac{(b-a)^n}{(n_0+1)^n} = \frac{(n_0 + 1)^n}{2^n} \quad (11)
\]

\[
= \frac{1}{2^n} \times N^{r}_{\text{Static}} \leq \frac{1}{2^n} \times N^{r}_{\text{Combined}} \quad (12)
\]

The above quantitative comparative results (11) and (12) show that generally the number of fuzzy rules needed in the dynamically constructive model is greatly fewer than that in Zeng’s model[12] or in Yan’s model[30] in MISO cases, especially when \( n \) is large.

In short, Conclusion of the above comparison is in Fig. 3.

VI. EXAMPLES

There are three examples employed to illustrate the following items:

1) The great improvement of the dynamically constructive method based on affine TS fuzzy model comparing to Zeng’s[12] and Yan’s[30];(Example A, B, C)
2) The effectiveness of the dynamic method to ease the serious curse of dimensionality to some degree;(Example C)

A. Affine TS model approximates twice differentiable real continuous function (2ISO)

Assume the twice differentiable real continuous function to be approximated is: \( F(X) = \sin(x_1 + x_2) + e^{x_1 x_2}, X \in [-0.5, 0.5]^2 \); approximation accuracy is \( \varepsilon = 0.029 \). The comparative results are shown in TABLE 1 where \( 20 \leq \frac{1}{\text{2D}} = 42.25 \) and \( 20 \leq \frac{0.1}{\text{2D}} = 45.5 \). The different partitions of premise space in distinct models and approximation errors (1600 points) can be found in Fig. 4(a) and Fig. 4(b), respectively.
TABLE I
COMPARATIVE RESULTS IN EXAMPLE A (2ISO)

<table>
<thead>
<tr>
<th></th>
<th>Fuzzy Rules Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zeng’s model[12]</td>
<td>169</td>
</tr>
<tr>
<td>Yan’s model[30]</td>
<td>91</td>
</tr>
<tr>
<td>The dynamic model</td>
<td>20</td>
</tr>
</tbody>
</table>

TABLE II
COMPARATIVE RESULTS IN EXAMPLE B (3ISO)

<table>
<thead>
<tr>
<th></th>
<th>Fuzzy Rules Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zeng’s model[12]</td>
<td>343</td>
</tr>
<tr>
<td>Yan’s model[30]</td>
<td>196</td>
</tr>
<tr>
<td>The dynamic model</td>
<td>36</td>
</tr>
</tbody>
</table>

B. Affine TS model approximates a polynomial (3ISO)

Assume the polynomial to be approximated is: \( h(X) = 1 + (x_1 + 2x_2 - x_3)^2, X \in [-0.45, 0.45]^3 \). approximation accuracy is \( \varepsilon = 0.4 \). The comparative results are shown in TABLE II where \( 36 \leq \frac{343}{2^9} = 42.875 \) and \( 36 \leq \frac{196}{2^9} = 49.2 \). The different partitions of premise space in distinct models and approximation errors can be found in Fig. 5.

C. Illustration of easing the curse of dimensionality (2ISO/3ISO/4ISO/5ISO)

Consider the following 4 polynomials, i.e., 2ISO/3ISO/4ISO/5ISO, with the same approximation accuracy \( \varepsilon = 0.5 \) and comparison results in different methods are in TABLE III.

\[
\begin{align*}
    h_{2ISO}(X) &= \frac{1}{2}(x_1 + x_2)^3, X \in [-0.4, 0.4]^2 \\
    h_{3ISO}(X) &= \frac{1}{3}(x_1 + x_2 + x_3)^3, X \in [-0.4, 0.4]^3 \\
    h_{4ISO}(X) &= \frac{1}{4}(x_1 + x_2 + x_3 + x_4)^3, X \in [-0.4, 0.4]^4 \\
    h_{5ISO}(X) &= \frac{1}{5}(x_1 + x_2 + x_3 + x_4 + x_5)^3, X \in [-0.4, 0.4]^5
\end{align*}
\]

From TABLE III, the number of rules in Zeng’s or Yan’s method increases exponentially with the increase of dimensionality. However, number of rules in the dynamically constructive method increases much more slowly, which shows the effectiveness of the dynamically constructive method to ease the serious curse of dimensionality to some extent.

VII. CONCLUSION AND DISCUSSION

A dynamically constructive method based on standard affine TS model is proposed in MISO cases, which is adaptive to the different properties of the function in different areas along each dimension. As a result, the number of rules in the TS model is greatly reduced comparing to [12] and [30]. Besides, because of the conciseness of standard affine TS model, the fuzzy systems obtained in this paper have simpler form than that in [21]. Moreover, unlike clustering method, the proposed method has a thorough coverage of input space.

As well known, there are many applications referring to the universal approximation theory of affine TS fuzzy systems in various fields. For example, it is very difficult, if not impossible, to directly control a very complex nonlinear system although its original nonlinear expression is known. It is necessary to turn such known complex model into approximative affine one which is generally easier to control. Nevertheless, many of the existing results in this field are either not practical for serious curse of dimensionality or a bit complex to operate. In this paper, one can easily construct a relatively simpler and more practical affine fuzzy system within the desired approximation accuracy.

However, there are also some problems needed further study in this paper:

1) In this method, all membership functions are rigid, which can be regarded as "Switching Fuzzy System". It would be even better to introduce soft membership
function, i.e., “General Fuzzy System”, because the consideration of quantitative influence of soft membership function is helpful for better partition and even fewer rules within the same approximation accuracy.

2) The dynamic partition algorithm may lead to a few of very small sub-spaces because some $Q_i$ may be very short during dynamic partition. Such case may be improved by some other better partition method, e.g., employ other different shapes of sub-spaces.

3) This method is used only for approximating to known original nonlinear functions. Although it is worth to further study, it should be extended to construct fuzzy model from real training data pairs.

In the future, we will study the above problems deeply and reduce the fuzzy rules number further more in sufficient conditions. Meanwhile, since MIMO cases can be regarded as the combination of several MISO cases, it will be also studied later.

VIII. ACKNOWLEDGMENT

The authors thank Dr. HongBo Li, Dr. Jian Ma, Mr. Shuai Feng and Mr. Kai Liu for their useful suggestions.

REFERENCES


TABLE III
THE ILLUSTRATION OF EASING THE CURSE OF DIMENSIONALITY (MISO)

<table>
<thead>
<tr>
<th>$h_{SISO}(X)$</th>
<th>Number of fuzzy rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n=2, 3, 4, 5$</td>
<td></td>
</tr>
<tr>
<td>$h_{SISO}(X)$</td>
<td>Zeng's model[12]</td>
</tr>
<tr>
<td>$h_{SISO}(X)$</td>
<td>Tan's model[30]</td>
</tr>
<tr>
<td>$h_{SISO}(X)$</td>
<td>The dynamic model</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>64</td>
<td>6</td>
</tr>
<tr>
<td>1296</td>
<td>1288</td>
</tr>
<tr>
<td>32768</td>
<td>129</td>
</tr>
</tbody>
</table>


