Identification and Control of Smooth Fuzzy Systems

by

Ebrahim Sadjadi

A dissertation submitted by in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

Computer Science and Technology

Universidad Carlos III de Madrid

Advisor(s):

Jose Manuel Molina Lopez
Jesus Garcia Herrero

Tutor:

Jesus Garcia Herrero

July 2019

To my Family

ACKNOWLEDGEMENTS

I would like to express my gratitude, firstly, to the advisors, Dr. Jesus Garcia Herrero and Dr. Jose Manuel Molina López, for all their support during my PhD research. Working with them was a great opportunity for me in my life to learn from them and develop my knowledge and skills.

I also would like to thank the committee members Dr. Antonio Berlanga de Jesús, Dr. Miguel Angel Patricio Guisado, and Dr. Javier Bajo Pérez, for the long discussions about the project and thesis and their motivational ideas and concerns during the period of the research and the pre-defense.

The continuous support of the staff members at Department of Informática is greatly appreciated: Rocio Pañero Martin and Maria Isabel Sánchez.

The support of all my friends could impact effectively in facilitation of the difficulties of going through the long months of carrying out the project, and all the writing and rewriting during the period of time. Especially, I would like to recognize the colleagues and friends: Dr. Abchouyeh, Dr. Javier Serrano and Dr. Hatami.

Last but not least, I greatly thank my family for their kind support during my studies.

PUBLISHED CONTENT

- E. N. Sadjadi, J. G. Herrero, J. M. Molina, Z. Hatami On approximation properties of the smooth fuzzy models. International Journal of Fuzzy Systems, Springer, 2018, DOI 10.1007/s40815-018-0500-9
- E. N. Sadjadi, M. B. Menhaj, J. G. Herrero, J. M. Molina Lopez. "How Effective are Smooth Compositions in Predictive Control of TS Fuzzy Models?", International Journal of Fuzzy Systems, 2019

OTHER RESEARCH MERITS

- Tutorial Accepted for IEEE PES ISGT ASIA 2019, entitled "Service-Dominant Logic as a Foundation for Business Model Innovation in Smart Grid", JINJIANG HOTEL in Chengdu, 21-24 May 2019

Abstract

During the years, we are witnessing a rapid change in the modeling and control of complex processes, which has necessitated employment of the approximate reasoning capacities of humans in the model identification and closed loop control of the uncertain and imprecision systems. One of the manifest of such intelligent schemes of modeling and control is the utilization of fuzzy logic based schemes, which has been facilitated the employment of computational capacities of the hardware.

Although the intelligent methods could empower the designers to reach high speed of computation and safe process control strategies, they are not perfect and bring the imperfections. They made the closed loop behavior of the system not to be continuous neither smooth due to the application of min-max composition in the fuzzy structure.

This thesis discusses on the alternative method of fuzzy modelling and control for the nonlinear processes, utilizing the smooth compositions. We introduce the modeling capacity of the smooth fuzzy models and then expand the formulation for the adaptive identification methods for the processes with the objective of incorporation to the model based predictive control schemes.

The smooth fuzzy compositions construct an overall nonlinear smooth and continuous model of the system. Hence, in the optimization based manipulations and control algorithms the model will require fewer computations in optimization phase rather than the classical fuzzy min-max based modeling scheme. It also provides an improvement in modeling accuracy and would be attractive for application to the systems with hybrid and switched dynamics with the limited number of discontinuity to obtain a continuous fuzzy model. The smoothness property has also impacted the closed loop behavior of the system largely.

Although, the combination of the iterative identification and model predictive control of the nonlinear processes have been directed many works in the academia and industry during the years; however, the smooth fuzzy structure will facilitate the employment of the experimental information of the system to closed loop structure with the minimum level of variations. To guarantee the stability of the scheme, we have considered the possibility of reaching the control horizon beyond the specific level to drag the system states inside the basin of attraction. Moreover, due to the smoothness of the scheme, the convergence of the results in face of uncertainties and disturbances will be faster, in comparison to the counterpart classical fuzzy schemes. It also can be easily tuned for the non-minimum phase and open loop unstable processes.

The performance of the theoretical studies has been examined using several simulations to demonstrate the outperform of the proposed schemes to the traditional fuzzy structures.

Index

Chapter 1

Introduction

The demand for the high quality products and energy saving methods has made the designers and practitioners in the industries to employ the integrated systems and tighter process design procedures to attain the safety standards and respect the environmental regulations. It has made the control problems evolve to be more complicated and challenging over the last few decades.

To facilitate the solution finding combat of the sophisticated problems in the process industry, the designers always care about the implementation issues and endeavor to find the paths that the available computing facilities would be capable of analysis of the dynamic process for monitoring of the full process taking from the raw material pressure and concentration to the finished product outflow rate and the involved energy and temperature transformation.

Taking into consideration that nearly all the variables are changing in time, the industrial processes are time varying, and beside that, almost all the transformation of processes are conducted and controlled by multivariable with nonlinear interactions, which induce the significant nonlinear track to the system input-output trajectory.

The primer objective of the employment of the controllers in the industrial processes is to handle and modify the nonlinear trajectory of the system appropriately, and to influence the parameters efficiently to make the process run through the desired fashion, which can be interpreted as,

- Safe operation of the process, with avoidance of the hazardous points of operation
- keeping the production rates to the specific points, in view of the disturbances, faults and parametric changes, aging, etc.
- Keeping the quality of products always reach and surplus the specified level, in view of different faults in the system, temperature change, etc.

However, the scale to which the system can operate for fulfillment of the mentioned objectives normally depends on different factors. Firstly, it depends on the nature of the process and how the intrinsic parameters will allow us have the freedom to impact the functioning of the process. Second, the available hardware, on which the one can trust for better design or modification of the dynamic process. Third factors, is the deepness of the knowledge of the process varying characteristic and monitoring of the system variables, upon which the operator can take the best decision and implement the best running policy.

It is so clear and understandable that the first and second factors are highly difficult to overcome. It means that, for instance, the control engineer cannot change intrinsic process characters, (to quit neither to expedite the time intervals for reaching the vapor from one unit to the next one); or the modification of the available hardware of the industry proceed fast with the advent of the digital computer innovations, are costly and in many cases impossible. Hence, the main challenge of the industrial process control will lie on the capacity to understand the process and the capability to infer from the available information of the processes, to take the best decision for fulfillment of the control process objectives.

This fact highlights more the current trend in process identification and model building for further control objectives and process behavior analysis, with the rational consideration of the uncertainties, parametric changes and noises. This is why model based control by concept has become one of the most important dilemma in the process control domain recently.

However, when it comes to the process control of the input-output data based models, the intelligent methods are always between the most demanded control strategies for the process adaptive and optimization based steering techniques to overcome the cumbersome situations.

The intelligent control schemes are normally considered the methods that combine the application of machine leaning techniques with the control theory [1]. They normally generalize the methods the human brains employs for functioning, leaning, detection, etc., to bring more flexibility, robustness and adaptive capacity to the computing based systems. Hence, the intelligent control systems are deemed to achieve the control goals, and the plants' model description, when the parameters, variables, and ambient are not completely recognized and defined with the proper level of robustness and adaptiveness.

1.1 Artificial Intelligence in Control

As stated above, the recent demand of the industrial processes for more reliable control schemes to raise the capacity of the production lines and make the performance more flexible in view of the faults and disturbances, has made the intelligent schemes of modelling and control very attractive for the designers and practitioners, especially that they rely more on the available input-output data and trust on the experience of the practitioners for dealing with the process challenges.

During the years, the intelligent schemes for modelling and control have become known more with the neural networks and fuzzy logic based algorithms which could show the high level of performance and robustness due to their capacity for the complex and uncertain conditions over the wide operating domain.

Artificial neural networks normally are designed to learn and infer from the stored data, which are supposed to emulate the human brain on the experiential basis while the fuzzy logic are assumed to function based on the linguistic expressions to emulate the human brain. In this debate, the soft computing algorithms are normally compared to the hard computing capacity, while the soft computing schemes are recognized by the fuzzy logic, neural networks and other verities of search techniques (includes genetic algorithms, convolutional fuzzy networks, etc.) in soft computing represents the hard computing schemes.

The peculiarities of the hard computing are the precision, and category, whereas soft computing are recognized by their approximation and dis-positionality. While, the mentioned attributes also exist in the hard computing, however, the nature and tolerance of the imprecision and uncertainty are quite different and are known to reach an acceptable level to make the machine learning techniques useful and handy, especially in neuro-computing and curve fitting.

The difference of soft computing and hard computing has opened a new era where many researchers have tried to take advantage of their combination, and many control techniques have appeared having the neural tools and fuzzy logic for complement of the classical technologies.

The combination of the soft computing and hard computing skills have facilitated employing the experimental information of the system to run the system tracks the specific trajectory or gain the desired features in operation. This is while, the hard computing skills alone could rely mostly just on the mathematical models derived from the physical laws. This could enable the controller system operates safely in the time varying environment, shows robust properties to the changes in the plant's dynamics and parametric variations, speeds up the stabilizing control behavior in the ill-defined and hazardous operations, and considers the structural restrictions and constraints.

However, all the AI techniques could not bring the same capacities to the controlled systems. For instance, while the neural network model of the system on the back propagation basis could save the continuity and the smoothness of the model, the fuzzy structure of min-max basis made the model loses such feature of continuity and smoothness. It has made the controller design more complicated which could not rely on the derivate of the model. We will discuss more on such attributes of the fuzzy model in the coming parts of this thesis.

Fuzzy Logic

As mentioned before, since the introduction, fuzzy logic has become one of the most attractive areas for research and also application to the industry. The control based on the linguistic approach has been appeared first in the works of Takagi and his colleagues [3] upon the fuzzy set definition of Zadeh in [4]. Since then, they have been employed for the control of processes in the diverse areas of application.

Despite the success of fuzzy logic in model building and control of the nonlinear processes, the models on the basis of the skills and knowledge of operators are static, whose good performance is appreciated to the lengthy process of trial and error and the effort involved for the proper rule selection, especially when it comes to the nonlinear systems. Besides, almost all the industrial processes are time-variant, which has directed many of the research works for the robust controller design and fuzzy controller development upon the for customized models which will be detailed in Chapters 3 and Chapter 5.

Adaptive fuzzy controller and adaptive fuzzy models are normally assumed to modify the fuzzy set definition, fuzzy membership function and the scaling factors. In this regards, such controllers (or models) also have been known as the self-tuning controllers [5] (or models). The modification of the fuzzy set parameters and fuzzy membership functions tune again the already designed fuzzy controllers (or models). There are also fuzzy model and controllers that alter the fuzzy rules in the process of modification, which often are called self-organizing controllers [6] (or models).

The direct self-organizing fuzzy controllers, employ the system input-output data to evaluate the system performance and modify the controller rules, without any modification of the process model, while the indirect fuzzy controllers [7], make the alteration and modification in the system model before finding the next proper controller input to the real system.

1.2 Motivations for Research

As stated above, an interesting character of the artificial intelligence techniques that has made the soft computing methods so interesting is their capacity to overcome the imprecision and uncertainty and resolve the modelling and control problems on the basis of the experience of the operators in a robust and cost effective manner. However, there are differences between the popular artificial intelligence techniques. It is to say, the neural network based models and controllers upon the back propagation technique are differentiable, continuous and smooth, in contrast to the fuzzy logic based models and controllers, on the basis of widely used min-max compositions.

Recently, smooth fuzzy compositions have been introduced which could narrow this gap, and make the fuzzy models differentiable, continuous and smooth as well. They could show superior performance to the widely used min-max compositions for modeling highly nonlinear industrial processes and for one-step ahead model predictive controller (MPC) design.

Hence, the principal goal of carrying the present research is to,

- Make contribution on the smoothness properties of fuzzy models, to find out why the smooth compositions show superior performance to the min-max compositions
- Incorporate smooth fuzzy framework for adaptive modelling of the nonlinear uncertain system in the Takagi-Sugeno structure
- Incorporate smooth fuzzy framework for long horizon MPC control of the nonlinear uncertain models in the Takagi-Sugeno structure

We will make novel fusion of ideas drawn from the fields of MPC, identification and fuzzy logic, on the basis of smooth composition to treat the nonlinear systems and will demonstrate through the theoretical studies and the comprehensive level of simulations the higher remarkable attributes of the smooth fuzzy compositions compared to the classical fuzzy ones, in their optimum functioning and the speed of convergence, with the capability of application to the unstable systems and the delayed processes.

1.3 Contributions of The Thesis

The research work during the period of the PhD program has been targeted to cover the following topics:

Fuzzy Modeling

The properties of smooth fuzzy models have been investigated and their approximation characters have been explored. We found that this modeling method offers a continuous and derivative model for representing nonlinear dynamic systems. Such properties will be employed in the subsequent sections for optimization based algorithms on the purpose of system identification and model predictive controller design. Such structure also can be utilized for making a continuous model for the switched and hybrid systems with limited number of switches.

System Identification

We have employed the smoothness properties of the fuzzy models introduced above to make a smooth fuzzy structure and thereby, obtain the optimum model of the dynamic system considering the possible time variations of the system parameters and the disturbances. We have tested the algorithm for the test problem as well as the highly nonlinear dynamic system of CSTR under uncertainties.

Although, we are witnessing different methods of identification of fuzzy models for the nonlinear processes, however, they are mostly on the basis of min-max compositions, which are not differentiable which lead to the construction of a non-smooth and non-continuous optimization problem. Hence, the proposed method facilitates the implementation issues and shows the faster convergence rate.

Model Predictive Control

Model predictive control is based on the employment of the optimization methods through the application of the modern computer-based hardware. Since normally the fuzzy models are constructed based on the min-max compositions, which are non-differentiable, hence, the use of such fuzzy models generally necessitates to resort to the techniques for solving the NP optimization problems, which are quite difficult. Hence, we have proposed the smooth fuzzy model structure to make up a control algorithm, which functions very efficient based on the mathematical derivative of the model. Therefore, the proposed model predictive control scheme targeted the long horizon. The ability of employing long horizon control for MPC can guarantee the stability of the obtained controller.

Adaptive Models

In the thesis we have developed identification and the MPC controller schemes based on the iterative methods of optimization. Hence, both the modeling and control structures are able to consider the parametric changes into account and thereby, give rise to adaptive control schemes for the uncertain systems. This propose advantages to the traditional approach of the fuzzy systems.

Robustness Properties

In the industrial processes, the improvement of the performance often is attained at the cost of deterioration of the robustness properties. Hence, in this work, we tried to demonstrate the acceptable betterment in the robustness properties of the smooth fuzzy models and controller for the working conditions considering the noises and disturbances different from the condition where the model has been identified or the controller has been tuned.

1.4 Overview of The Thesis

The current thesis comprises six chapters to present the investigation has been done during the PhD program. An overview of the contents of each chapter of this thesis is as follows.

- Chapter 1

  This chapter provides background information on our research project and examines the important contemporary challenges of industrial process control. The concepts of model fuzzy control and identification are briefly introduced. The emphasis in the second case has been on drawing the attention of the reader to the most important trends and to highlight the system control applications, with particular emphasis to fuzzy logic.

- Chapter 2

  Fuzzy modelling is a procedure for developing fuzzy membership functions and fuzzy rules from a given data set. This chapter begins by introducing the basic concepts and definitions involved in fuzzy modelling. These concepts are used to examine various fuzzy modelling approaches that have been proposed in the literature.

- Chapter 3

  In this chapter the different fuzzy composition schemes are being presented. We will also review the state of art and earlier contribution and results on the smooth fuzzy compositions and smooth fuzzy systems.

- Chapter 4

  This chapter starts by describing the Newton based fuzzy identification method and will continue with describing self-learning model scheme. Then, we will apply the method to model a benchmark example and will extend it to a CSTR system.

- Chapter 5

  This chapter emphasizes the development of the conceptual framework for a fuzzy model predictive control strategy based on the fuzzy modeling approach presented in Chapter 4. The focus will be on the analytical approaches of designing a long range predictive control algorithm. Compared to the earlier works, we intend to extend the prediction and control horizons used by the controller. The much lower computational requirements of the analytical approach provide it with a distinct advantage over the numerical approach.

  Employing the analytical approach, we will not use any fuzzy model linearization, hence, we will expect better performance, which has been explored explicitly. Since the performance of the controller in the presence of noise and under process conditions quite different from that used for the learning and identification of the process, the robustness of the model has been explored both analytically and in experience. We note that, the fuzzy model obtained through the algorithm is assumed to be dynamic, according to the system last data, which makes the whole structure, sensitive to the system disturbance, parameter variations and uncertainties.

- Chapter 6

  This chapter summarizes the most important findings of this research project and makes recommendations for further research work in the same area.

Chapter 2

Fuzzy Logic

In this chapter, we set forth the basic mathematical ideas used in fuzzy modeling and control. We will start explanation of these ideas easily and their applications will be expanded upon in later chapters.

2.1 Introduction

Fuzzy set theory is assumed globally to be the way of representation of the imprecise data. After the introduction by Zadeh, it has been widely utilized for the mathematical interpretation of the uncertainty and vagueness to formulate the imprecise reality.

The application of Fuzzy set theory has empowered the engineers to formulate the oral description of the dynamical behavior of the systems to develop advanced control methods. The integration of the experimental information and uncertain expectation of the real world as the component of the algorithms, has extended the capability of the engineers to think beyond the physical laws and hence, they could overcome several difficulties in the system modelling and control domain. First, the environmental effects and the ambient disturbances will be accounted easier in the fuzzy model based description of the industrial processes. This feature becomes more prominent when it comes to the distributed, stochastic and the nonlinear systems. Second, in the modeling of the large scale system, the are many unknown variables governed by the unknown causal relationships, that are either difficult to measure and take into account in the mathematical formulations, or are expensive computationally or physically to measure and embed in the system´s model.

Hence, when the tradeoff between the precision and the significance of the statement for formulation of the system behavior goes toward losing the exclusive characteristic of the system, in view of the (in)compatibility of the variable and the complexity of the model, the role of fuzzy logic in the system model making becomes more prominent. Indeed, fuzzy is the nonlinear mapping of the input to the output of the system, under the shadow of the experience and knowledge of the operator, to consider the different environmental effects and the ambient disturbances into account. In fact, the capacity of defining different regions of normal and abnormal working conditions has made the fuzzy logic so rich to be able to overcome the enormous possibilities of the system functioning. This richness is along the capacity of the fuzzy models to be enough accurate to include the linguistic and rule-based form of the desired operation of the system, instead of being constructed upon the pure mathematical formulations, and reasonably complex to provide solution for the available computing capacity of the industrial hardware.

From the structural point of view, fuzzy model is a combination of the fuzzification of the real worlds variables through the membership functions, setting fuzzy rules for the given experimental input-output data and the defuzzification step. Hence, we intend to introduce the basic concepts and definition of fuzzy modeling paradigm in this chapter. This chapter concludes with the comparison of the different modeling approaches, through the relational matrix based style and fuzzy rule based style, proposed over the years in the literature.

2.2. Preliminaries

Fuzzy Set

A Fuzzy set is assumed to be the extension of the ordinary crisp set [7,8], where each element of the set is a member of fuzzy set upon a certain degree of membership. The degree of membership is defined to take value between 0 and 1; (while in the crisp set theory an element is definitely a member or is not a member). Hence, the fuzzy set F in a universe of discourse $U = \{u_i, i = 1, \cdots, n\}$ by its membership function will be mathematically written as,

$$\mu_F = U \rightarrow [0,1] \tag{2.1}$$

If $\mu_F(u_i)$ takes the boundary values of 0 or 1, then, the set will be changed to be ordinary rather than fuzzy. However, considering U to be continuous, then, a set F will be defined as below to be fuzzy,

$$F = \int_U \mu_F(u)/u \tag{2.2}$$

Similarly, in the discrete domain, the set F should be presented as following to be fuzzy,

$$F = \sum_{i=1}^{n} \mu_F(u_i)/u_i \tag{2.3}$$

Linguistic Variables

A linguistic variable normally is considered to be a variable of the fuzzy number, where the fuzzy number is predefined to be from a normal and convex fuzzy set. Also, a variable of values defined by the linguistic terms is known to be a linguistic variable. For instance, if we define the term "very cheap" for a price below about 40 euros, "cheap " for a price about 50 euros, "moderate" for a price close to 60 euros, "expensive " for a prices about 70 euros and " very expensive " for a price above about 80 euros, then the linguistic values of price in the domain zeros to hundred and the associated variable could be

T(price) = {very cheap, cheap, moderate, expensive, very expensive}

Operation on Sets

The operations for the fuzzy sets are being done through the membership functions of the associated sets. To clear up, suppose A and B be two fuzzy sets in the a universe of discourse U defined, respectively, by two membership functions μ_A and μ_B.

Definition 2.1. Union: the membership function $\mu_{A\cup B}$ corresponding to the union $A \cup B$ for all $u \in U$ can be defined as $\mu_{A\cup B}(u) = \max\{(\mu_{A(u)}, \mu_B(u)\}$.

Definition 2.2. Intersection: the membership function $\mu_{A\cap B}$ corresponding to the intersection $A \cap B$ for all $u \in U$ can be defined as $\mu_{A\cap B}(u) = \min\{(\mu_{A(u)}, \mu_B(u)\}$.

Definition 2.3. Fuzzy relation: A fuzzy relation as an n-array $\{[(u_1, \cdots, u_n), \mu_R(u_1, \cdots, u_n)] | (u_1, \cdots, u_n) \in U_1 \times \cdots \times U_n\}$, is a fuzzy set in domain of $U_1 \times \cdots \times U_n$.

Definition 2.4. Sup-star composition: The sup-star composition of R and S expressed as $R \circ S$ for two fuzzy relations R and S over the domain $U \times V$ and $U \times V$, will be:

$$R \circ S = \{\left[(u, w), \sup_v(\mu_R(u, v) * \mu_S(v, w))\right], u \in U, v \in V, w \in W \} \tag{2.4}$$

where * is considered as minimum value of the arguments. The Sup-star composition is itself a fuzzy relation, and is considered as a class of composition of triangular norms (t-norms) and co-norms (s-norms). We will return to this point and talk more on the fuzzy interface and compositions later in chapter 3.

Approximate Reasoning

The approximate reasoning, based on the definition, is make inference of the available data, which is carried out mostly upon the generalization of the results. To illustrate our interpretation of the approximate reasoning here, assume the fuzzy sets A, A', B and B' as:

Premise 1: x is A'

Premise 2: if x is A then y is B

Simply speaking the reasoning upon the Premise 1 and Premise 2 will be,

Consequent: y is B'

However, naturally the reasoning cannot be deduced always by inspection neither in a unique way. Normally, the approximate reasoning is carried out upon the compositional rule for the inference.

Definition 2.5. Sup-star compositional: For the fuzzy relation R in $U \times V$, and considering x as a fuzzy set of U, the sup-star compositional rule for the inference expresses that the fuzzy set y in V will be induced from x upon the formula:

$$y = x \circ R \qquad (2.5)$$

Similar to the above definition, different compositional rule of inference has been developed and introduced, upon the employed operations on the fuzzy sets. We will talk more on the fuzzy operations and compositions in the next chapter.

Fuzzy Logic System

From the engineering view point of practical usage of the fuzzy logic, the operator is supposed to set the gauges and scale up the controllers upon the proper level upon the deduction of the fuzzy reasoning schemes. However, both the input and out of the fuzzy systems are number and the operator needs to have the right number to employ to make the desired functioning of the system. Therefore, it is required to have "fuzzifier" to convert the numbers in crisp set of the real world to the fuzzy numbers and "defuzzifier" to do the converse in returning back to the real world [9]. The combination of four elements of fuzzifier, rule base, inference engine and defuzzifier is called to be a fuzzy logic system (FLS), which is depicted in Figure 2.1.

Functioning of the fuzzifier is to scale up the input variables into the proper value in the universe of discourse, upon a predefine mapping. This is very similar to the fuzzification step, where the numerical data is translated into the linguistic values [10,11].

Converse to fuzzification, the defuzzification is to evaluate the output variable in a crisp value, from the result of the reasoning in the inference engine. In the industrial processes, normally, the defuzzification involves the scaling step, where the range of output variables becomes a numerical value upon the domain of the involved application.

The fuzzy rule presents the base of logic for description of the relations between input and output of the fuzzy system. Such rules are the principle component in the fuzzy logic paradigm, that emulates the human brain behavior for proper interpretation of situations and surroundings for

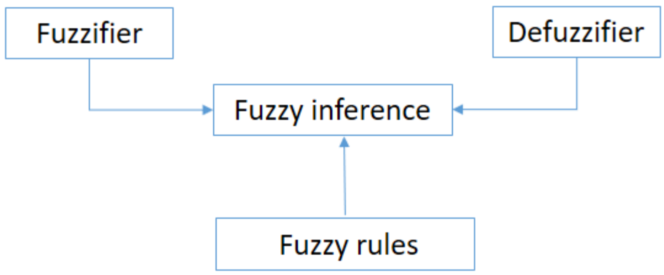


Figure 2.1. Principal structure of the fuzzy logic systems.

intelligent functioning for the logical decision making, pattern recognition, approximate reasoning, predictions, etc. The fuzzy rule can be manifested through the relational matrixes or be transparent in the IF-THEN style.

<u>Fuzzy Rules</u>

The fuzzy rules normally are defined in the IF-THEN style:

R^l: IF x_1 is F_1^l and $\cdots$and IF x_n is F_n^l THEN y is G^l (2.6)

where $\bar{x} = (x_1, \cdots, x_n)^T \in U$ and $y \in Y$ are supposed to be, respectively, the input vector and the output vector to the fuzzy system, F_i^l and G^l are assumed, respectively, as labels for the fuzzy sets in domains X_i and Y, and $l = 1, \cdots, r.$ is the rule number. In (2.6), corresponding to each fuzzy IF-THEN rule, a fuzzy implication $F_1^l \times \cdots F_n^l \rightarrow G^l$ is defined on per.

Different fuzzy implication rules have been introduced in the previous works for interpretation of the multi-variable fuzzy logic systems. We summarize the most common fuzzy implication rule as follows,

Minimum operation:

$$\mu_{F_1^l\times\cdots\times F_n^l\rightarrow G^l}(\bar{x}, y) = \min[\mu_{F_1^l\times\cdots\times F_n^l}(\bar{x}), \mu_{G^l}(y)] \quad (2.7)$$

Product operation:

$$\mu_{F_1^l\times\cdots\times F_n^l\rightarrow G^l}(\bar{x}, y) = \mu_{F_1^l\times\cdots\times F_n^l}(\bar{x}) \cdot \mu_{G^l}(y) \quad (2.8)$$

Arithmetic operation:

$$\mu_{F_1^l\times\cdots\times F_n^l\rightarrow G^l}(\bar{x}, y) = \min[1, 1 - \mu_{F_1^l\times\cdots\times F_n^l}(\bar{x}) + \mu_{G^l}(y)] \quad (2.9)$$

Max-min composite operation:

$$\mu_{F_1^l\times\cdots\times F_n^l\rightarrow G^l}(\bar{x}, y) = \max\{\min\left[\mu_{F_1^l\times\cdots\times F_n^l}(\bar{x}), \mu_{G^l}(y)\right], 1 - \mu_{F_1^l\times\cdots\times F_n^l}(\bar{x})\} \quad (2.10)$$

where $\mu_{F_1^l\times\cdots\times F_n^l}(\bar{x})$ is defined as,

$$\mu_{F_1^l \times \cdots \times F_n^l}(\bar{x}) = \mu_{F_1^l}(x_1) * \cdots * \mu_{F_n^l}(x_n).$$
(2.11)

Here the symbol "*" corresponds to the conjunction "and" in (2.10), which is known to be called as the t-norm. We will talk on t-norms, their properties and their different operations later in this thesis.

Fuzzy Inference

The fuzzy inference is supposed to make a mapping of fuzzy sets in U to the fuzzy sets in Y. The fuzzy inference is the exposition of the defined IF-THEN rules, which perform the calculations using the composition of operations in the fuzzy sets.

To clear up, assume A_x as the fuzzy set in U; and for each $R^{(l)}$ of (2.10) the fuzzy inference defines the fuzzy set $A_x \circ R^{(l)}$, into Y. This mapping is computed upon the composition of operations for interpretation of the rule of inference as:

$$\mu_{A_x \circ R^{(l)}}(y) = \sup_{\bar{x} \in U} \left[\mu_{A_x}(\bar{x}) * \mu_{F_1^l \times \cdots \times F_n^l \to G^l}(\bar{x}, y) \right]$$
(2.12)

where $\mu_{F_1^l \times \cdots \times F_n^l \to G^l}(\bar{x}, y)$ is the fuzzy implication rule from the formula (2.7)-(2.11).

In the general case, the final fuzzy set, $A_x \circ (R^{(1)}, \cdots, R^{(M)})$, is calculated upon the M rules defined by the fuzzy rules which is obtained using the fuzzy disjunction:

$$\mu_{A_x \circ (R^{(1)}, \cdots, R^{(M)})}(y) = \mu_{A_x \circ R^{(1)}} \otimes \cdots \otimes \mu_{A_x \circ R^{(M)}}(y).$$
(2.13)

Here the symbol "⊗" is called to be the t-conorm. We will talk on t-conorm, their properties and their different operations later in this thesis.

Fuzzifier

The fuzzifier is to map the crisp point $\bar{x} = (x_1, \cdots, x_n)^T \in U$ into the fuzzy set A_x in U. Upon the widely used fuzzifier scheme, for A_x with the fuzzy singleton with support $\bar{x}$ we would have

- $\mu_{A_x}(x') = 1$ for $x' = \bar{x}$ and
- $\mu_{A_x}(x') = 0$ for all other $x' \in U$ with $x' \neq \bar{x}$

In the other style, the definition is as follows,

- $\mu_{A_x}(\bar{x}) = 1$ for $x' = \bar{x}$ and
- $\mu_{A_x}(x')$ decreases from 1 as x' moves away from $\bar{x}$, $x' \neq \bar{x}$ (2.14)

Defuzzifier

The defuzzifier is assumed to map fuzzy sets in output into the crisp set of the output. Two commonly used choices of the deffuzifier are,

- Maximum defuzzifier:

$$y = \arg sup_{y' \in Y} \left[\mu_{A_x \circ (R^{(1)}, \cdots, R^{(M)})}(y') \right]$$
(2.15)

- Centre-average (or centroid) defuzzifier:

$$y = \frac{\sum_{l=1}^{M} y'(\mu_{A_x \circ R^{(l)}}(y^l))}{\sum_{l=1}^{M}(\mu_{A_x \circ R^{(l)}}(y^l))}$$
(2.16)

where y^l is the point of Y where $\mu_{G^{(l)}}(y)$ reaches the maximum.

2.3. Fuzzy Model Selection

In the last section we have discussed the four principal elements of the fuzzy logic system. Hence, for each element, there are a certain degree of freedom to shape up the fuzzy model through the proper selection of schemes from the available possibilities. To summarize, it is notable to say that, we could decide which type of fuzzification to utilize (i.e. singleton or nonsingleton), we could decide which shape of the membership functions to use, (triangular, trapezoidal, Gaussian, etc.), we choose their parameters to be fixed during a training procedure or varies, and defuzzifier scheme to be maximum or centre-average. Besides, one can choose from the available compositions (e.g. max-mm, max-product) and select the most appropriate inference scheme (minimum, product, etc.) to gain the highest level of performance from the fuzzy model building, according to the desired objective. -

The final formulation to obtain the output of the fuzzy logic systems, can be rewritten through the the product operator for inference operation, with the employment of the center-average defuzzifier, product inference and singleton fuzzifier, as follows,

$$y = \frac{\sum_{l=1}^{M} y'(\prod_{i=1}^{n} \mu_{F_i^l}(x_i))}{\sum_{l=1}^{M}(\prod_{i=1}^{n} \mu_{F_i^l}(x_i))}$$
(2.17)

where y' is the point where μ_{G^l} reaches the maximum value. This is the formulation that manifests the fuzzy logic system in a global prospect of function approximation. We will return to this formulation in Chapter 4 to deduce the properties of the fuzzy logic model upon the application of the smooth fuzzy compositions.

Relational Fuzzy Model

Relational fuzzy models rather than the IF-THEN rule based fuzzy models, rely on the relational arrays and matrixed for manifestation of the input-output relationship. The matrix considers any possible combination of input-output variable and assigns a value to that. This value is supposed to be between zero and one, and represents the possibility or truth of that particular relationship. Hence, the strongest relationship is valued one and the zero is assigned to the weakest relationship of the considered input-output relationship. The system description for the relational fuzzy model is as follows,

The common identification schemes for identification of the relational fuzzy models is on the basis of Cartesian product of input-output data, known as the linguistic approach. In this scheme the logical examination is used for determination of the degree of fulfillment of the rule for a particular data point. Hence, if the degree of fulfillment is more than the predefined threshold, it would be considered as a valid rule and subsequently, a rule table will be constructed with the entries of 1 and zero, corresponding to location of the valid rules. Although the relational models have shown the satisfactory results, however, they suffer from the incapacity for the on-line model modification and the requirement of huge space for the saving the arrays and tables for the high dimensional and large scale systems.

Conclusions

In this chapter different fuzzy modeling schemes have been introduced. They can be constructed either in the known IF-THEN style or be manifested in the relational matrix based schemes. We have tried to discuss briefly on different elements in the fuzzy model structure of fuzzifier, defuzzifier, fuzzy inference and approximation and the related compositions and operators. The measures to employ before taking decision on the intended shape of the fuzzy model will greatly depend on the intended application. However, in the control of industrial processes, model accuracy, on-line learning and modification capability, the computational requirements, and the convergence rate for the parameters normally become prominent. We will talk more on such attributes in the following chapters. Also, we will study how fuzzy compositions will bring about different attributes, by which will impact on the capacity of the fuzzy logic to be incorporated to the different application in system identification and control.

Chapter 3

Smooth Fuzzy Models (State of Art)

In this chapter, we review different fuzzy compositions and study their characters. Then, we review the state of art of smooth fuzzy system development for system modeling, identification and control schemes. These concepts will be employed for further theory development in the subsequent chapters.

3.1. Introduction

Although the most common operations for expression of AND, OR, and NOT are minimum, and maximum, however, there are more definitions of the operators for interpretation of the operations on the fuzzy sets. Hence, the subject of the next sections is these operations [8-11].

Definition 3.1: Triangular norm. A mapping

$$T: [0,1] \times [0,1] \to [0,1] \tag{3.1}$$

is called as triangular norm (t-norm as abbreviation) iff for each argument, it is symmetric, associative and non-decreasing. Besides, the value $T\ (a,1) = a$ for all $a \in[0,1]$. In other interpretation, any t-norm T would satisfy the below properties:

Commutative:

$$T(a,b) = \mathrm{T(b,a)},\ \forall\ a,b \in[0,1]. \tag{3.2}$$

Associative:

$$T\big(a,T(b,c)\big) = \mathrm{T(T(a,b,c))},\ \ \forall\ a,b,c \in[0,1]. \tag{3.3}$$

Monotone:

$$T(a,b) \leq T(c,d), if\ a \leq c, b \leq d \tag{3.4}$$

One Identity:

$$T(a,1) = \mathrm{a}, \forall \mathrm{a} \in (0,1) \tag{3.5}$$

Normally, the t-norms are expressed as $a * b = T(a,b)$, which consider to be functions of two variables. The basic t-norms defined in the literature are:

1- *Min t-norm*

$$T(a,b) = \min(a,b)$$

2- *Product t-norm*

$$T_P(a,b) = a \cdot b$$

3- *Lukasicwicz t-norm*

$$T_L(a,b) = \max(a+b-1,0)$$

4- *Weak t-norm*

$$T(a,b) = \begin{cases} \min(a,b) & \max(a,b) = 1 \\ 0 & o.w \end{cases}$$

5- *Hamacher t-norm*

$$T_H(a,b) = \frac{ab}{\gamma + (1-\gamma)\ \| a+b-ab \|} \quad \gamma \geq 0$$

6- *Doubois and Prade t-norm*

$$T_D(a,b) = \frac{ab}{\max(a,b,\alpha)} \quad \alpha \in (0,1)$$

7- *Yager t-norm*

$$T_Y(a,b) = 1 - \min\left\{1, \sqrt[p]{(1-a)^p + (1-b)^p}\right\} \ p > 0$$

8- *Frank t-norm*

$$T_F(a,b) = \begin{cases} \min(a,b) & if\ \lambda = 0 \\ T_P(a,b) & if\ \lambda = 0 \\ T_L(a,b) & if\ \lambda = 0 \\ 1 - \log_\lambda\left[1 + \frac{(\lambda^a - 1)(\lambda^b - 1)}{\lambda - 1}\right] & o.w \end{cases}$$

All t-norms may be extended, through associativity, to $n > 2$ arguments. A t-norm T is called strict if T is strictly increasing in each argument.

Definition 3.2 (Archimedean): A t-norm T is consider as Archimedean iff it is continuous and $T(a,a) < a$ for all $a \in (0,1)$.

Proposition 3.1: Every Archimedean t-norm T can be expressed by the continuous and decreasing function $f: [0,1] \to [0,\infty]$ and $f(1) = 0$ where

$$T(a,b) = f^{-1}(\min\{f(x) + f(y), f(0)\}). \tag{3.6}$$

In this case, the function f is called the additive generator for the t-norm T.

Proposition 3.2: All continuous t-norms introduced above, other than minimum, are Archimedean.

Definition 3.3: If $T(a,b) = 0$ holds for some $a, b \in (0,1)$, then the t-norm T is called nilpotent.

Example 3.1: Every Lukasiewicz t-norm is the prototype of a nilpotent t-norm.

Definition 3.4: An Archimedean t-norm is said strict when $T(a,a) = 0$ only for $a = 0$.

Example 3.2: Every t-norm of product is the prototype of a strict t-norm.

Definition 3.5: Consider two t-norms T_1, and T_2. Then, T_1 is said to be weaker than T_2 (with the notation $T_1 \leq T_2$) if $T_1(a,b) \leq T_2(a,b)$ for all $a, b \in [0,1]$.

Definition 3.6: Triangular conorm. A mapping

$$S: [0,1] \times [0,1] \to [0,1] \tag{3.7}$$

is a called triangular co-norm (or in abbreviation t-conorm) provided that in each argument, it be symmetric, associative, non-decreasing. Moreover, $S\ (a,0) = a$ for all $a \in [0,1]$. As a result, any t-conorm S will satisfy the properties:

Commutative:

$S(a,b) = \mathrm{S(b,a)}, \forall\ a,b \in [0,1].$ (3.8)

Associative:

$$S(a,S(b,c)) = \mathrm{S(S(a,b,c))},\ \forall\ a,b,c \in [0,1]. \tag{3.9}$$

Monotone:

$S(a,b) \le S(c,d), if\ a \le c, b \le d$ (3.10)

One Identity:

$S(a,0) = \mathrm{a}, \forall \mathrm{a} \in (0,1)$ (3.11)

Theorem 3.1: For the t-norm T,

$$S(a,b) = 1 - T(1-a, 1-b) \tag{3.12}$$

The basic t-conorms are:

1- *Max s-norm*

$$T(a,b) = \max(a,b)$$

2- *Probabilistic s-norm*

$$S(a,b) = \mathrm{a + b - ab}$$

3- *Lukasicwicz s-norm*

$$S(a,b) = \max(a+b, 1)$$

4- *Strong s-norm*

$$S(a,b) = \begin{cases} \max(a,b) & \min(a,b) = 0 \\ 1 & o.w \end{cases}$$

5- *Hamacher s-norm*

$$S_H(a,b) = \frac{a+b-(2-\gamma)ab}{1-(1-\gamma)ab} \quad \gamma \ge 0$$

6- *Yager s-norm*

$$S_Y(a,b) = \min\left\{1, \sqrt[p]{a^p + b^p}\right\} \quad p > 0$$

Until now in this section, we have studied the most popular and widely used t-norms and t-co-norms. However, not all the presented operators are differentiable, which harden application of fuzzy logic in the optimization part of the different identification, data classification and control algorithms. It is widely accepted that application of derivative based optimization algorithms facilitates highly the application to the different fast and reliable problems of the real world. Hence, in the next section we will focus on the smooth fuzzy compositions and try to look to such problems through the specific lenses.

3.2 Smooth Fuzzy Compositions

As stated above, the smoothness and continuity of the fuzzy compositions brings about the capacity to employ the gradient-based algorithms for optimal tuning of the parameters of the fuzzy model, identification or controller design, with higher speed of convergence and with smoother convergence behavior. The other benefit employing the gradient based methods is that even if we come to the sub-optimal solution, the overall algorithm will be stable and feasible.

Therefore, in this section we want to review a general framework for smooth fuzzy compositons, including the recent contributions on the matter, and we will review the utilization of the smooth fuzzy relational compositions in system modeling and control. Let's start with presenting the new smooth t-conorms and t-norms from [12,13]:

Smooth t-norm

$$I: T_S(a,b) = 1 - \cos\left(\frac{2}{\pi}\cos^{-1}(1-a)\ \cos^{-1}(1-b)\right) \quad (3.13)$$

$$II: T_S(a,b) = \frac{4}{\pi}\tan^{-1}\left(\tan\left(\frac{\pi}{4}a\right)\tan\left(\frac{\pi}{4}b\right)\right) \quad (3.14)$$

$$III: T_S(a,b) = 1 - \frac{2}{\pi}\cos^{-1}\left(\sin\left(\frac{\pi}{2}a\right)\sin\left(\frac{\pi}{2}b\right)\right) \quad (3.15)$$

$$\text{IV}: T_S(a,b) = \cos\left(\cos^{-1}a + \cos^{-1}b - \frac{2}{\pi}\cos^{-1}a\ \cos^{-1}b\right) \quad (3.16)$$

Smooth s-norm

$$I: S_S(a,b) = \frac{r.d.\beta^{-\log_\beta(d) - \log_\beta(r)} - 1}{(\beta-1)}, r = (\beta-1)a+1, s = (\beta-1)b+1, \beta \in (1,\infty) \quad (3.17)$$

$$II: S_S(a,b) = 1 - \frac{4}{\pi}\tan^{-1}\left(\tan\left(\frac{\pi}{4}(1-a)\right)\tan\left(\frac{\pi}{4}(1-b)\right)\right) \quad (3.18)$$

$$III: S_S(a,b) = \frac{2}{\pi}\cos^{-1}\left(\cos\left(\frac{\pi}{2}a\right)\cos\left(\frac{\pi}{2}b\right)\right) \quad (3.19)$$

$$\text{IV}: S_S(a,b) = \cos\left(\frac{2}{\pi}\cos^{-1}a\ \cos^{-1}b\right) \quad (3.20)$$

The above-mentioned smooth compositions are differentiable almost everywhere. After presenting these new smooth compositions in [12], they have been use to make fuzzy model of the dynamic systems [13]. Following that other authors have employed them for one step ahead model predictive control of dynamic system [14]. However, all the contributions in the field have employed the fuzzy relational modeling framework.

Relational fuzzy models indeed can be easily developed and modified online, however, they have some structural drawbacks that has obstacles their applications in the industrial systems.

Firstly, since they are being manifested with the elements of zero and one assigned to the proper cells in the arrays and matrixes, therefore, their application is limited just to the systems of limited number of variables, since handling of such matrixes are not easy for the large scale and complex systems. Besides, the computation requirement of such arrays of zero and one is not comparable to the computational requirement of the fuzzy model defined by the common IF-THEN rules.

The second difficulty is that analysis of the result of the controllers by handing the input-output of the system is not easy through the relational fuzzy models. For instance, the impression of the operator through the inspection of the matrixes for fault detection, stability analysis and other structural analysis are much more cumbersome rather than investigation of the common IF-THEN rules. From this point of view, it is worth mentioning that one of the primer ideas of employing the fuzzy models is to resort to the linguistic rules and statement for expression of the system behavior, which will get lost in handling the matrixes of zero and one. Therefore, many has underlined that fuzzy relational models cannot put the human in loop to interact with the skills and knowledge of the operators, as it is expected from the fuzzy modeling paradigm.

Hence, in this section we will review the literature developed so far, for application of smooth compositions in system identification and controller design to get ready and convey our motivation of making the main contributions of the current thesis, which will be presented in the following chapters.

3.3 Fuzzy Relational Model Identification

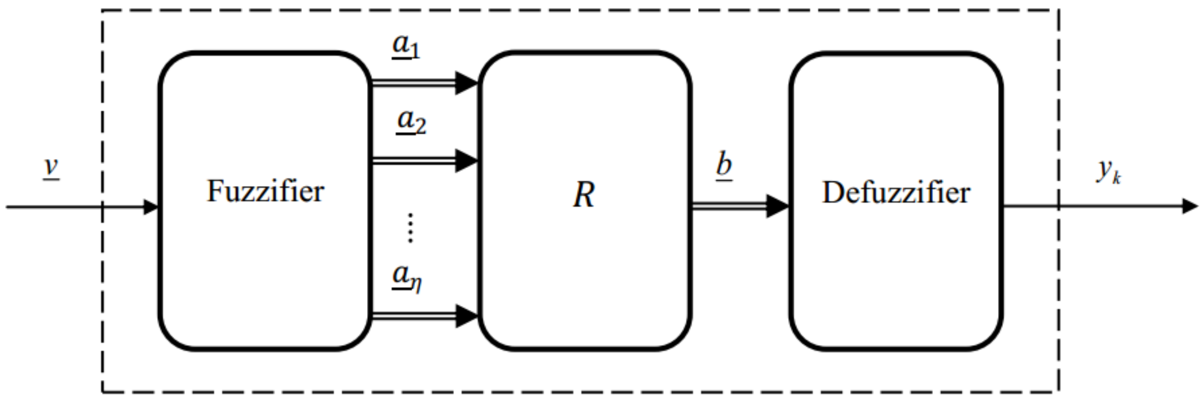


Figure 3.1 Main parts of an Fuzzy Relational Method

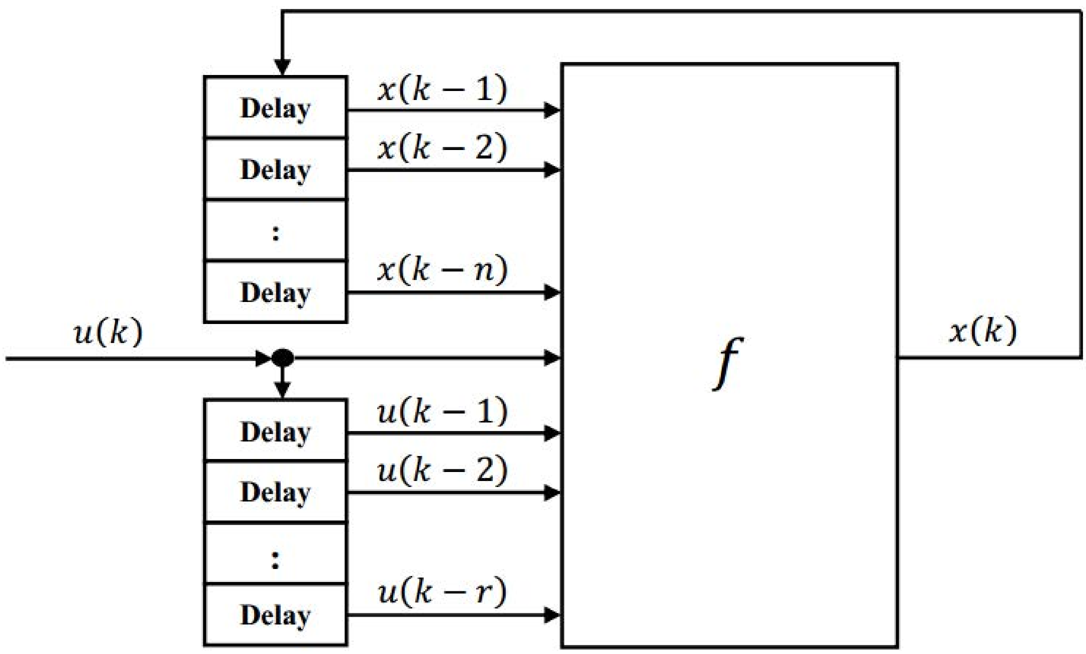


Figure 3.2. Block diagram of the dynamic system for system identification in [12].

Ashtiani and Menhaj [12, 13] have considered the actual input to the system comprised of the input vector and the delayed states of the dynamic system as demonstrated in the function block of Fig. 3.1. The same structure is shown in Fig. 3.2 with the outer box and the general input vector of $\underline{v} = [x(k-1), \cdots, x(k-n), u(k-1), \cdots, u(k-r)]$ of the length $\eta = n + r$. Correspondingly, they have considered an $(\eta + 1)$ dimensional fuzzy relational matrix. Hence, the structure of the model in its general format has been considered as, fuzzification Block, Fuzzy Relational Matrix and Defuzzifier.

As they elaborated in [13], Although the fuzzy relational matrix in many of the fuzzy relational models are considered to be only two-dimensional, Ashtiani and Menhaj have considered the fuzzy relational matrix with the dimension above two in order to make the model properly capture the real system.

They tried to apply fuzzy composition in a N-dimensional structure, to capture the multi-dimension structure of the fuzzy matrix, in a manner to be rational and understandable for proper globalization of the basic fuzzy matrix, and expand the two dimensional fuzzy compositions for higher dimensions for expression of the multilevel systems. Hence, compared to the N-dimensional fuzzy relational compositions in the earlier works that Ashtiani and Menhaj [12, 13] has called the straight fuzzy relational compositions, the recent authors have proposed the structural fuzzy relational compositions through the successive application of the fuzzy smooth compositions. They claimed that their fuzzy relational model can handle the multilevel fuzzy relational networks too, although they did not bring any simulation or other kind of evidence to back up the claim.

The operation of the innermost block to gain the output is represented by successive application of fuzzy relational s-t composition "∘":

$$\underline{b} = R \circ \underline{a}_1 \circ \underline{a}_2 \circ \cdots \circ \underline{a}_\eta. \tag{3.21}$$

The identification algorithm they proposed is based on calculation of the derivative of the functions and can be optioned online. Therefore, very much like the other fuzzy relational model identification, they employed a productive correction term to adjust the elements of the Fuzzy relational matrix iteratively. However, according to the virtue of the new s- and t- norms the derivation of the correction term for their modified fuzzy relational model takes more work than the ordinary ones. The benefit of their approach is that the iterative optimization based algorithm can be calculated from the defined error-dependent cost function explicitly upon Newton based optimization algorithm. They applied the algorithm to the gas furnace data from Box and Jenkins. In their contribution they considered a combination of the deterministic and stochastic optimization algorithms and allowed the learning rate to be dynamic for speeding up in the learning process.

Although in the tuning procedure they employed derivative-based algorithms and very well used the assumption that the t-norm and the t-conorm are differentiable or at least piecewise differentiable with a finite number of non-differentiable points, where for the non-differentiable point, a virtual value could be approximated by its neighbor points, their algorithm cannot easily be employed for the large scale systems since they suffer from the curse of dimensionality. Actually, since they model the whole dynamic of the system in an N-dimensional matrix, an increase in the number of the linguistic terms would result in the exponential expansion of the number of the size of the matrixes, which avoid making practical application of the algorithm and also their easy and fast tuning. This is why they in practice faced a difficult problem to solve and hence, employed a mixture of deterministic and stochastic optimization to speed up the identification process. However, implementation of such methods of problem solving is tremendous in practice.

3.4. One Step Ahead Predictive Control for Relational Fuzzy Model

Askari and Menhaj [14] have employed fuzzy relational modeling scheme described above for one step ahead prediction and control of complicated and nonlinear systems. They developed a MPC schemes based on fuzzy relational nonlinear models and considered non-quadratic cost-function directly from the input-output data. They have demonstrated the efficiency of the proposed fuzzy relational model predictive control scheme in contrast to the PID and sliding mode control methods, subject to different uncertainties toward the control of nonlinear systems.

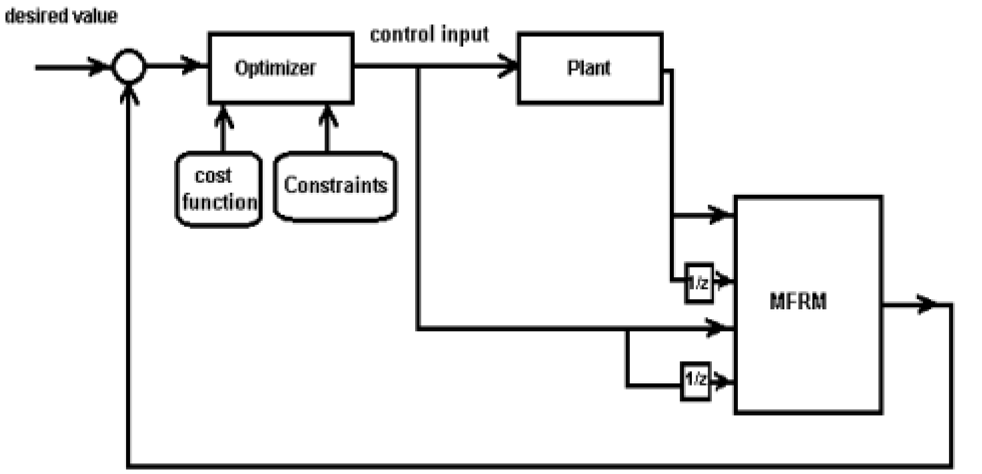
In the modeling part they have employed the same scheme of Aghili and Menhaj [12,13] for the iterative system identification process.

They considered the cost function as:

$$J = \frac{1}{2}(y_{desired} - y_{model})^2 = \frac{1}{2}\left(cte - \frac{\sum_{i=1}^{q} \underline{b}(i)c(i)}{\sum_{i}^{q} \underline{b}(i)}\right)^2 = \frac{1}{2}$$

$$\left(cte - \frac{\sum_{i=1}^{q}\left(R(k_1,\cdots,k_{\eta+1})\circ\underline{a}_1(k_{\eta+1})\circ\underline{a}_2(k_\eta)\circ\cdots\circ\underline{a}_\eta(k_2)\right)c(k_1)}{\sum_{i=1}^{q}\left(R(k_1,\cdots,k_{\eta+1})\circ\underline{a}_1(k_{\eta+1})\circ\underline{a}_2(k_\eta)\circ\cdots\circ\underline{a}_\eta(k_2)\right)}\right)^2 \quad (3.22)$$

where $c(k_1)$ s have been values of the weighted averages for the fuzzy model. As the normative of model predictive control, they made the formulation to optimize the cost function respect to the first system´s input $\underline{a}_1$ and repeated this iteration for the next time step. They could obtain the excellent tolerance of the control system to the different levels of noise, disturbance and up to 25% uncertainty.



3.3 Block diagram of Fuzzy MPC (MFRM stands for modified fuzzy relational model [14].)

But, there were some weak points, due to the triangular membership function they used for formulation of the problem. Therefore, although they employed the smooth fuzzy compositions to solve the optimization problem using the derivative based methods, but ultimately they solved the problem through a non-derivative optimization method and could not enjoy the actual computational benefit of the scheme. Secondly, the algorithm they proposed had a short prediction horizon which puts the stability of the algorithm at risk, especially for real time applications of complex systems.

Conclusions

In this chapter we have discussed on the properties of fuzzy compositions and presented different fuzzy compositions. Furthermore, we introduced the smooth fuzzy operators and reviewed their practical application in dynamical system modeling and control.

Since, the attention on smooth fuzzy models have emerged just recently, the literature developed for such systems and their practical applications are just in the infantile period. In the next chapter we will elaborate more on the smoothness properties of the smooth fuzzy operators and will employ them subsequently for identification and long range fuzzy model predictive control of the nonlinear systems.

The "what" is in constant flux, the "why" has a thousand variations.

Chapter 4

On Approximation Properties of Smooth Fuzzy Models

In this chapter, we pose the basic idea behind the smooth fuzzy compositions and prove their approximation properties. We will start explanation of the smoothness and continuity and then extend it to the fuzzy compositions and models.

4.1 Introduction

It is well known that fuzzy systems can uniformly approximate any real continuous function on a compact domain to the desired degree of accuracy. The universal approximation properties of the fuzzy systems, with Gaussian membership functions, product t-norm and centroid defuzzification has been proved in the literature [15, 16]. The results for the Gaussian, triangular or trapezoidal membership functions, any t-norm and any practical defuzzification also has been proved in [17]. Over the years, this topic has been extended for the accurate approximation of a smooth function beside its derivatives [18, 19], however, it is unanswered that whether a fuzzy system with arbitrary continuous membership functions (not necessarily Gaussian, triangular or trapezoidal) can accurate approximation a function smoothly, i.e. not only the smooth function is approximated but also its derivatives [15].

Hence, in this chapter, we will work on this topic and will demonstrate that from the application of smooth compositions and the arbitrary continuous membership functions, the fuzzy model can approximate the real plant and its derivatives. The importance of the topic is that, through this stream, we would be able to model two (or more) different states of a discontinuous or a switched system by a single fuzzy model with the minimum variation. The other contribution will be that we can be sure that employing the smooth compositions in the design of the fuzzy models and controllers, the plant can damp the uncertainties and parameter variation and noises fast. Such extensions will be topic of the next chapters.

However, for the present chapter, the structure is as follows. First we review mathematical smoothness and continuity properties. Then, we study the general structure of fuzzy systems. Based on the results of the two beginning sections, we formulize the smoothness property of a special class of fuzzy systems which is the main result of the chapter. Following that we bring an example to demonstrate the practical functionality and properties of the obtained results and the proposed theorems. Finally, we draw conclusions.

4.2 Preliminaries

In this section for the convenience of the readers we review some mathematical backgrounds from [8-11].

Definition 4.1: A function $f(x)$ is continuous at the point c if and only if $f(x)$ is defined at c and for any $\epsilon > 0$ there exists a $\delta > 0$ such that $|f(x) - f(c)| < \epsilon$ if $|x - c| < \delta$.

Definition 4.2: A function $f(x)$ has gap discontinuity at c if $f(c)$ is undefined.

For instance, $\frac{f_1(x)}{f_2(x)}$ has gap discontinuity at c if $f_2(c) = 0$.

Definition 4.3: A function $f(x)$ has jump discontinuity at c if $f(c)$ is defined and $\lim_{x \to c^+} f(x) \neq \lim_{x \to c^-} f(x)$.

The function $f(x) = \begin{cases} 4, x < 0 \\ 5, x \geq 0 \end{cases}$ for example has a jump discontinuity at $x = 0$.

<u>Structure of fuzzy systems</u>

We consider the multiple input single output systems to facilitate our theory development. Nevertheless, our results can be extended for the multiple - input- multiple output systems; since the multiple outputs can be decomposed readily into several single output systems.

Consider the problem of approximation for a nonlinear function of the following form:

$$f: R^n \to R \tag{4.1}$$

$$y = f(x_1, x_2, \cdots, x_n) \tag{4.2}$$

For every input variable of the system we consider an interval where there is the highest probability that the variable lies in this interval. Then, we divide the interval into 2N+1 regions and assign a membership function to each region.

Next step in constructing fuzzy system is to assign rules for the data in the different regions of the input and output domains. We consider,

$R^{(i)}$: if x_1 is M_1^i and x_2 is M_2^i and $\cdots$ and $x_n$is M_n^i then (4.3)

$g(x_1, \cdots, x_n)$ is b_i under the probability $\mu_i, i = 1, \cdots, r$

Here the function $g(x_1, \cdots, x_n)$ is about to approximate the function $f(x_1, \cdots, x_n)$ in the corresponding interval. The rules generated for the fuzzy system in this way, have two "if" and "then" parts. There are different ways for interpretation of the relations and making mathematical inference on the fuzzy values using the compositions of t-norm and s-norm in the fuzzy systems′ domain. The different types of the fuzzy compositions introduced in the literature, are summarized in Chapter 2 and Chapter 3.

We remind that the actual output of the model is determined based on the centroid defuzification formula, which is given simply by,

$$g(\bar{x}) = \frac{\sum_{i=1}^{r} b_i \ \mu_i}{\sum_{i=1}^{r} \mu_i} \tag{4.4}$$

where μ_i is considered to be at the center of the region B^i at every time instant of dynamics of the system, $\bar{x} = [x_1, \cdots, x_n]$ and r is the total number of the fuzzy rules for approximation of the plant.

4.3 Approximation Properties

On the purpose of explaining the approximation procedure, we consider equation (3.16) and equation (3.20) as the formulation of t-norm and t-conorm under study, and denote them by smooth compositions T_{s-IV} and S_{s-IV}, respectively. The approach can be extended to other types of the smooth compositions. For the system defined by the function $f(x_1, \cdots, x_n)$ introduced above, we assume r=2, with three state variables, then, the fuzzy model will be written as,

$$g(x_1, \cdots, x_n) = \frac{N(x_1,\cdots,x_n)}{D(x_1,\cdots,x_n)} = \frac{b_1 * \mu_1 + b_2 * \mu_2}{\mu_1 + \mu_2} \tag{4.5}$$

where $\mu_i(\bar{x}, \alpha_i)$are the membership functions from the system state vector $\bar{x} = [a, b, c]$, $i = 1, \cdots, r$ and α is the design parameter.

$$\mu_i(\bar{x}, \alpha_i) = S_{s-IV}\left(T_{s-IV}\left(\mu_i\ (a,\cdot), \mu_i(b,\cdot), \mu_i\ (c,\cdot)\right)\right) = \tag{4.6}$$

$$S_{s-IV}(T_{s-IV}(T_{s-IV}(\ \mu_i(a,\cdot), \mu_i(b,\cdot)), \mu_i\ (c,\cdot)))$$

Let $\Lambda_1 = T_{s-IV}\big(\mu(a,\cdot), \mu(b,\cdot)\big)$, and $\Lambda_2 = T_{s-IV}\big(\ \Lambda_1, \mu(c,\cdot)\big)$, and upon Eq (3.16),

$$\Lambda_1 = \cos\left(\cos^{-1}\mu_i(a,\cdot) + \cos^{-1}\mu_i(\ b,\cdot) - \frac{2}{\pi}\cos^{-1}\ \mu_i(a,\cdot)\ \cos^{-1}\mu_i(\ b,\cdot)\right)$$

$$\Lambda_2 = \cos\left(\cos^{-1}\Lambda_1\ +\ \cos^{-1}\mu_i(\ c,\cdot) - \frac{2}{\pi}\cos^{-1}\Lambda_1\ \ \cos^{-1}\mu_i(c,\cdot)\right).$$

Based on Eq (3.20), $\mu_i(\cdot, \alpha_i) = \cos\left(\frac{2}{\pi}\ cos^{-1}\mathit{\Lambda}_1\ cos^{-1}\mathit{\Lambda}_2\ \right)$, hence, we define,

$$\theta = \frac{2}{\pi}\ cos^{-1}\mathit{\Lambda}_1\ cos^{-1}\mathit{\Lambda}_2$$

(4.7)

Therefore,

$\mu_i(\cdot, \alpha_i) = \cos(\theta) =\ real\ (\exp(j\theta))$,

where $j = \sqrt{-1}$. For the suitable selection of exponential function $G(x, \alpha_i)$ and θ we write it more simple as,

$$\mu_i\ (\cdot, \alpha_i) =\ G(\cdot, \alpha_i) := real\ (\exp(j\theta)) \tag{4.8}$$

Therefore, we can generalize the procedure and write,

$$g(x_1, \cdots, x_n) = \frac{N(x_1,\cdots,x_n)}{D(x_1,\cdots,x_n)} = \frac{\sum_i^r b_i G(\bar{x},\alpha_i)}{\sum_i^r G(\bar{x},\alpha_i)}. \tag{4.9}$$

If we consider a box $[-N, N], \cdots, [-N, N]$ along a dense grid with the steps $\Delta\, \alpha_1 = \cdots = \Delta\alpha_r = h$ and correspondingly $b_i = b(\alpha_i)$, we can write the summation as the integration,

$$N(\bar{x}).h^n = \int_{-N}^{N} \cdots \int_{-N}^{N} b(\bar{\alpha}).G(\bar{x},\bar{\alpha})\ .d\bar{\alpha} \tag{4.10}$$

$$D(\bar{x}).h^n = \int_{-N}^{N} \cdots \int_{-N}^{N} G(\bar{x},\bar{\alpha}).d\bar{\alpha}\ . \tag{4.11}$$

Now, as $h \to 0$ and $N \to \infty$, we will have the multi-dimensional integrals,

$$N_\infty(\bar{x}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} b(\bar{\alpha}).G(\bar{x},\bar{\alpha})\ .d\bar{\alpha} \tag{4.12}$$

$$D_\infty(\bar{x}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} G(\bar{x},\bar{\alpha}).d\bar{\alpha}\ . \tag{4.13}$$

The value of the last integral is independent of the system states vector $(x_1, \cdots, x_n)$. Hence, it sums up to a constant value C for $D_\infty(\bar{x})$. Therefore, to find the approximation of the function $f(\cdot)$ we just need to find the weights $b(\bar{\alpha})$, such that

$$C\ \cdot g(\bar{x}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} b(\bar{\alpha}).G(\bar{x},\bar{\alpha})\ .d\bar{\alpha} \tag{4.14}$$

The right side of this equation is the convolution of the function $b(\bar{\alpha})$ and the real function $G(\bar{x},\bar{\alpha})$. We can transform the convolution to the frequency domain, and use the Fourier transformation to find the weights as,

$$\hat{b}(\bar{\omega}) = \frac{C\cdot g(\bar{\omega})}{G(\bar{\omega})} \tag{4.15}$$

and then use the inverse Fourier Transformation to get the desired function $b(\bar{\alpha})$.

To come up to the results, here we review some theorems from the signal and system literature.

Theorem 4.1: Let $\mathcal{F}$ and $\mathcal{R}$ be continuous real-valued functions and assume that $\mathcal{F}$ or $\mathcal{R}$ is zero outside some bounded set. If $\mathcal{F} \in C^k$ and $\mathcal{R} \in C^l$, then $\mathcal{F} * \mathcal{R} \in C^{k+l}$.

Proof: see [15].

*Theorem 4.2***:** (Derivative Theorem) If $\mathcal{F}$is a rough function, and $\mathcal{R}$ is a smooth function, then the convolution $\mathcal{F} * \mathcal{R}$ will be smoother than $\mathcal{F}$.

Proof: The theorem can be deduced from Theorem 4.1, see also [15].

Theorem 4.3: If $\mathcal{F}$is a rough function and $\mathcal{R}$ is n-times differentiable, then the convolution $\mathcal{F} * \mathcal{R}$ will be n-times differentiable.

Proof: The theorem can be deduced from Theorem 4.1, see also [15].

Corollary 4.1: The convolution $\mathcal{F} * \mathcal{R}$ is at least as smooth as the function $\mathcal{F}$ and the function $\mathcal{R}$ separately.

Theorem 4.4: The fuzzy model obtained by the arbitrary membership function and the smooth s-norm and t-norm compositions is continuous, n-time differentiable and smoother than a periodic cosine function.

Proof: The fuzzy model is the convolution of the function $b(\bar{\alpha})$ by the function $G(\bar{x},\bar{\alpha})$ weighted by the constant value C, according to the Eq (4.14). Since the function $G(\bar{x},\bar{\alpha})$ is a cosine function whatever the membership functions are, hence, based on Theorems 4.1-4.3, we can conclude Theorem 4.4. It is to say, the model will be smoother than cosine function, whatever the function $f(\cdot)$ is.

Remark 4.1: Theorem 4.4 applies independent of the shape and nature of the plant, according to the derivative theorem stated above. In other words, Theorem 4 applies even if the plant has a rough or discontinuous dynamics.

Remark 4.2: The interpretation of theorem 4 in control application will be that, the control surface which the smooth fuzzy system produces will be smooth. Even if the system has a discrete state or systematic transition, based on this theorem, the transition in the system will happen with the minimum level of abrupt changes and variations. Moreover, the control system will show a better robustness to the uncertainties and disturbances in the region around the steady state point, trying to stay on the smooth surface.

Now we look at the properties of the estimation of derivatives of the plant.

4.4 Estimation of the Dynamic System Derivatives

We first consider the first derivative of the model. Taking the first derivative we have,

$$g_1(\bar{x}) = \frac{N_1(\bar{x})}{D(\bar{x})} - \frac{N(\bar{x})D_1(\bar{x})}{D^2(\bar{x})} \tag{4.16}$$

$$N_\infty(\bar{x}) = \sum_i^r b_i G_1(x, \alpha_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} b(\bar{\alpha}). G_1(\bar{x}, \bar{\alpha}) \, .d\bar{\alpha} \tag{4.17}$$

$$D_\infty(\bar{x}) = \sum_i^r G_1(x, \alpha_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} G_1(\bar{x}, \bar{\alpha}).d\bar{\alpha} \, . \tag{4.18}$$

Again using the same procedure, we come to,

$$C \cdot g_1(\bar{x}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} b(\bar{\alpha}). G_1(\bar{x}, \bar{\alpha}) \, .d\bar{\alpha}. \tag{4.19}$$

If we consider the m-th higher derivatives, similarly, we arrive to

$$C \cdot g_m(\bar{x}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} b(\bar{\alpha}). G_m(\bar{x}, \bar{\alpha}) \, .d\bar{\alpha}. \tag{4.20}$$

Theorem 4.5: The fuzzy model obtained by the smooth compositions is continuous and m-time differentiable.

Proof: Considering that the function $G_m(\bar{x}, \bar{\alpha})$ is a cosine function in origin and m-times differentiable, also Theorem 2, hence we always can approximate the desired function up to m-th Derivative $g_m(\bar{x})$ with the desired accuracy using the smooth fuzzy model (m is an arbitrary number).

Theorem 4.6: The approximation function $g(\bar{x})$ is defined everywhere in the domain of the states with the possible finite numbers of jump discontinuities.

Proof: Based on the definition, the smooth compositions are smooth with the possible number of discontinuities over their domains. Hence, the integration $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} b(\bar{\alpha}). G \ (\bar{x}, \bar{\alpha}) \ d\,\bar{\alpha}$ in Eq (4.1) always can be calculated by the grid based sum of integration for the appropriate small h and large N.

Theorem 4.7: Consider the smooth fuzzy system defined above with the parameters $\bar{\alpha} = [\alpha_1, \cdots, \alpha_r]$ for the rules. The smooth fuzzy model $g(\bar{x})$ is continuous if $\exists \alpha_{i,}, i \in [1, r]$ such that $\max(\mu(\alpha_i)) > 0$, i.e. there exists at least one input fully covered by the membership functions.

Proof: We describe the case for $i = 2$ which is extendable to the cases with the higher number of rules. Consider the integral part as stated above for the case $i = 2$, when there is a discontinuity for $\alpha_2 \in [c_1, c_2]$.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} b(\bar{\alpha}).G_D(\bar{x},\bar{\alpha})\ d\,\bar{\alpha} = \int_{-\infty}^{\infty}\int_{-\infty}^{c_1} b(\alpha_1,\alpha_2).G_D(\bar{x},\alpha_1,\alpha_2)\ d\,\alpha_1 d\alpha_2$$
$$+\int_{-\infty}^{\infty}\int_{c_2}^{\infty} b(\alpha_1,\alpha_2).G_D(\bar{x},\alpha_1,\alpha_2)\ d\,\alpha_1 d\alpha_2. \quad (4.21)$$

The above integrations are calculable at every point, as the last input is supposed to be fully covered by the membership functions, and is continuous.

From the Fourier Transformation viewpoint, described above, it worth mentioning that the Fourier Transformation exists only if the jump discontinuity at $\alpha_2 = c$ cannot change the value of any of the integrals, i.e. , $c_1 = \lim_{\alpha_2 \to c^-}$, $c_2 = \lim_{\alpha_2 \to c^+}$ since at this case, the inverse of the Fourier transformation will converge to the mid value level at the point of discontinuity.

Corollary 4.2: Consider the smooth fuzzy system with one input. The jump discontinuity of the mapping function $f(x_1, \cdots, x_n)$ will not impact on the smoothness property of the resulted smooth fuzzy model.

Proof: For the one input case, the above formulation will be,

$$\int_{-\infty}^{\infty} b(\bar{\alpha}).G_D(\bar{x},\alpha_1)\ d\,\alpha_1 = \quad (4.22)$$

$$\int_{-N}^{c-\epsilon} b(\bar{\alpha}).G_D(\bar{x},\alpha_1)_{\epsilon\to 0}\ d\,\alpha_1 + \int_{c+\epsilon}^{+N} b(\bar{\alpha}).G_D(\bar{x},\alpha_1)_{\epsilon\to 0}\ d\,\alpha_1$$

for a suitable choice of number N. It is obvious that the value of the integral will not be affected by the point discontinuity of the system.

Theorem 4.8: (Main Theorem) Let d and n be integers, and $N > 0$ and $\epsilon > 0$ be real numbers. Assume the function $f(x_1, \cdots, x_n)$ is a D-times differentiable function on $[-N, N]^n$. Then, using the smooth fuzzy compositions, one can construct a fuzzy model $g(x_1, \cdots, x_n)$ to approximate the function $f(x_1, \cdots, x_n)$ and its derivatives up to D-th order with the desired accuracy ϵ.

Remark 4.3: The results we presented here compared to the earlier works by Kreinovich [19] on smoothness properties for the fuzzy models brings much lesser restrictions; As in this chapter we have not put any restriction on shape of membership function, (to be or not be in Gaussian Form) to gain the smoothness property, compared to their work.

Now we show the effectiveness of the obtained results by an illustration.

4.5 Illustrative Examples

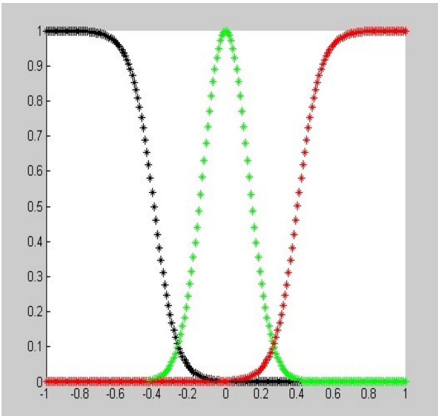
To demonstrate application of the proposed approach, we take the simple model as Table 4.1, where each rule consequent is shown based on the crisp number.

The table represents the logical rule that orchestrate switching between the different states of the finite state machine [20-22]. Such kind of logical rules, when coupled with the controller and the plants modelled with continuous or difference equations are generally called hybrid or switched systems which have the increasing popularity for modeling and control of the devices with digital components, e.g. relays, switches, stepper motors, so on [22]. Traditionally fuzzy controllers for hybrid and switched systems are designed such that every subsystem is being considered by a separate fuzzy structure. What follows is an evidence that using the smooth fuzzy schemes, it would be possible to model and control the different discrete states of the system by a single fuzzy model structure such that the augmented continuous and discrete dynamics of the system changes between the augmented continuous and discrete states of the model smoothly.

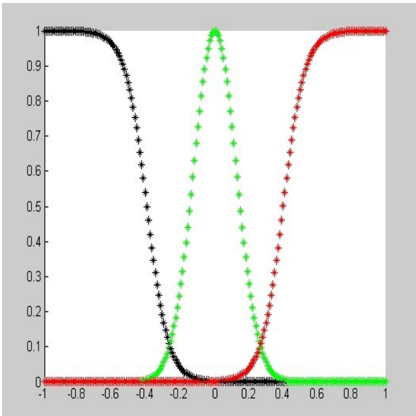
Table 4.1: logical rules of the switched system

$x_1 \backslash x_2$	X_{21}	X_{22}	X_{23}
X_{11}	**1**	**2**	**3**
X_{12}	**4**	**5**	**6**
X_{13}	**7**	**8**	**9**

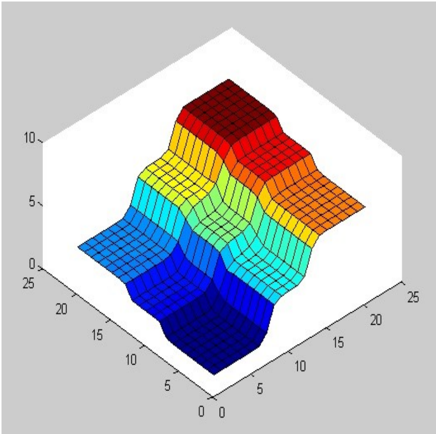
We first consider fuzzy membership functions shown in figure 4.1, where they cover all the domain of the system states definition. Consequently, we have used first the conventional fuzzy inference for the fuzzy model and compared that to the smooth fuzzy structure. They are equal in the functioning for mapping the input-output relation.



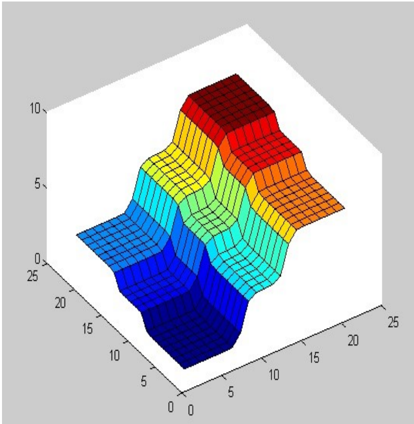
Membership function for the state x1



Membership function for the state x2



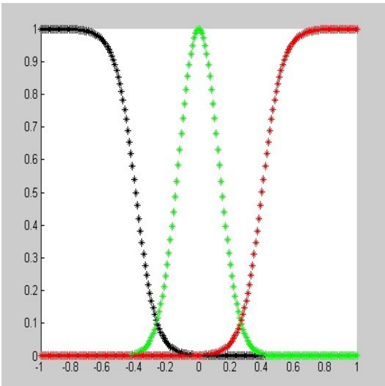
Output surface of the classical fuzzy systems



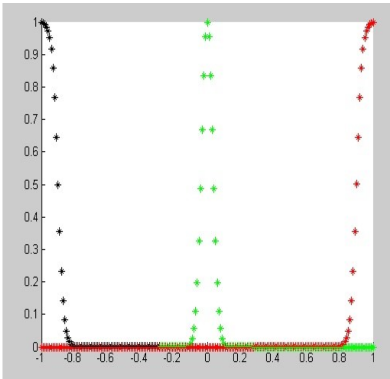
Output surface of the smooth fuzzy

Figure 4.1: Case 1: when membership functions of both states cover the space.

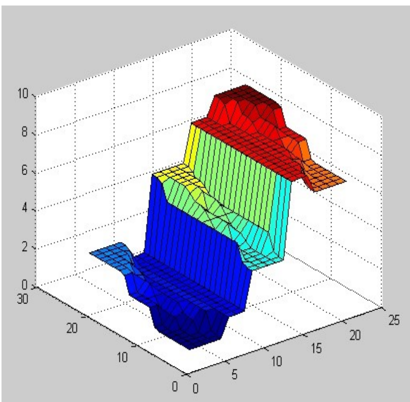
We then considered fuzzy membership functions shown in figure 4.2 and figure 4.3, where the membership function of the first state covers all the domain of system state definition, and the membership function of the second state does not cover all the second state space. Again, we have used the conventional fuzzy inference for the fuzzy model and compared that to the smooth fuzzy model. This is clear that the classical fuzzy model has great value of variation in the output. This is while the smooth fuzzy model has a minimum variation which is to say its performance is almost similar to the case 1, when the membership functions covered all the state space.

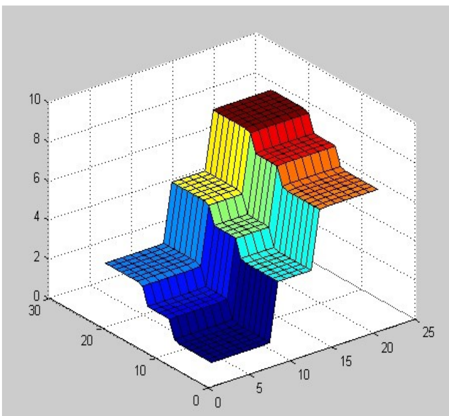


Membership function for the state x1



Membership function for the state x2





The output surface of the classical fuzzy

The output surface of the smooth fuzzy system

Figure 4.2: Case 2: when membership functions of just one state covers the state space

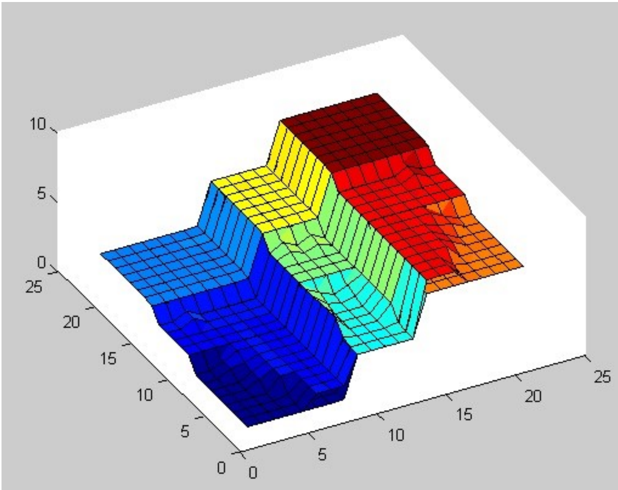
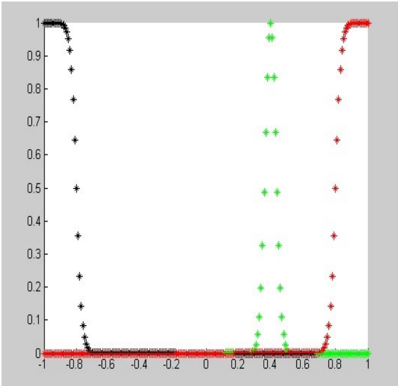


Figure 4.3: The comparison of the output of the fuzzy function with the smooth compositions vs the classical compositions

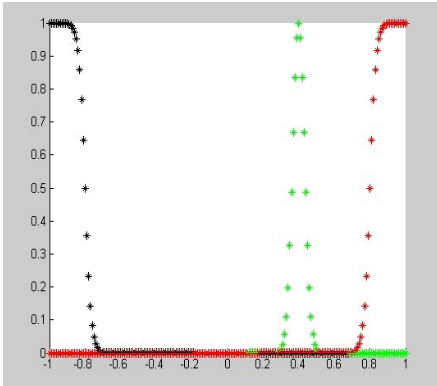
Lastly, we consider the fuzzy membership functions shown in figure 4.4, where none of the state spaces are covered by the membership functions. It is clear that both of the fuzzy models show high amount of variation and discontinuity.

As it is clear from the results of simulations, when the membership functions cover at least one of the state space variables, the smooth fuzzy model shows a smooth and minimum variation behavior for modeling of the input-output mapping, compared to the classical fuzzy model. In control applications, this feature can be used to damp the effect of the parameter variations and noise in the system and using the smooth compisitions one can run the system to return to the stable states after the disturbance with the minimum turbulences.

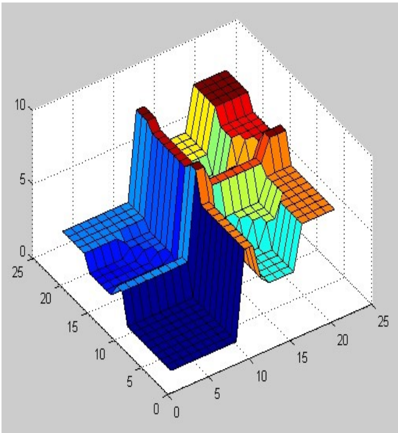
The inspection of the results in the example could clear up that i) converse to the earlier contributions, we are able to model two (or more) different states of a discontinuous or a switched system by a single smooth fuzzy model. ii) Based on the result of the case when the membership function of the second state in the simulation does not cover all the second state space, we claim that the smooth fuzzy models can uniformly approximate any real continuous function on a possible non-compact domain to the desired degree of accuracy, which is new in the fuzzy modeling literature. iii) As it has demonstrated in the simulation, for the



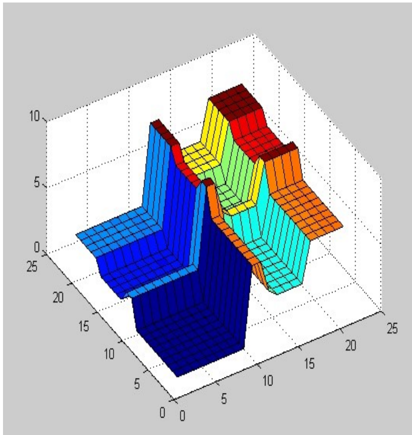
Membership function for the state x1



Membership function for the state x2



Output surface of the classical fuzzy systems



Output surface of the smooth fuzzy systems

Figure 4.4: Case 3: when membership functions of none of states cover the space

case when the second state in the simulation does not cover all the second state space, the error between the smooth fuzzy model and the plant, in comparison to the error value with the same definition in the classical fuzzy model of the plant, has declined much more - to the minimum possible value. Hence, we claim that employing the smooth compositions in the design of connectivist fuzzy modeling and controller schemes, we will not have high value of real plant-fuzzy model difference, neither the un-modeled dynamics, and therefore, will not need to restrict ourselves to the conservative methods of robust or adaptive control schemes. iv) As it is demonstrated by the simulation, the smoothness property of such fuzzy model structure could encompass the change in the discrete modes of the switched system by showing the minimum amount of variations and errors. Hence, we can generalize it and expect to be able to damp the uncertainties and parameter variations of the systems and environmental noises also very fast through the same smoothness properties of the fuzzy model. Some simulations in the earlier publications on smooth fuzzy modeling and control [12, 13] have demonstrated the robustness properties of such connectivist smooth fuzzy modeling and control schemes, however, they lack providing a theoretical analysis. The current chapter can back up their results by giving a clue that why such robustness properties exist.

Conclusions

In this chapter it is shown that we can model the dynamics and derivative of the continuous systems using the smooth fuzzy structure. We did not put any limitation on shape of the membership functions, in contrast to the earlier works, where the special Gaussian membership function has been considered. As a result of this finding, what we need to do in design of smooth fuzzy systems, to approximate dynamics of the system along its derivatives, would be to stick to the common practice of choosing the centric point and do not care of shape of the membership function.

We backed up our theories by an example where each rule's consequent has been shown based on a crisp number. It can be seen as a Mamdani model with the height defuzzification, or the discrete state models of hybrid and switched systems.

In our analysis and transformation, to run the approximation error and its derivative tends to zero, we need to increase the number of partitions in the dense grid as well as the fuzzy rules. It means that in the practical applications, we will have growing numbers of fuzzy rules to make use of the smooth approximation properties. Therefore, there is a trade-off between the accuracy of the fuzzy model and the modeling complexity. Hence, it is required to think about a method for finding the minimal number of fuzzy rules for a given accuracy of the fuzzy model in the future researches. One suggestion will be to discard the rules which have weak contribution to the output. The interested reader is referred to [23] for the fuzzy region assignment schemes.

No one can lose either the past or the future - how could anyone be deprived of what he does not possess?

Chapter 5

Fuzzy Model Identification and Self Learning with Smooth Compositions

This chapter develops a smooth model identification and self-learning strategy for dynamic systems taking into account possible parameter variations and uncertainties. Compared to the earlier works on the smooth fuzzy modelling structures, we could reach a desired tradeoff between the model optimality, without the need of having computations with large matrixes. It potentially will lower the computational load. The proposed method has been evaluated on a test problem as well as the nonlinear dynamic of a chemical process.

5.1. Introduction

The fuzzy structures presented in the literature are either defined by the rule- based models or they are fuzzy relational models [8, 11]. Rule based models describe the process behavior by a set of IF-THEN mechanisms. On the other hand, in the fuzzy relational models the input-output mapping is presented through a matrix. This matrix conveys any possible combination of the input-output mapping with a value between zero and one, which is the scale of truth (or probability) of each possible input to output relation.

It worth reminding the slight alteration of the definition of a smooth fuzzy topology built from the employment of the smooth fuzzy norms by fuzzy relations which is associated to the concept of composition of binary numbers and relations in the earlier works [12,13], rather than the topology built from the employment of the same norms in the IF-THEN model, which more relates to the concept of fuzzy numbers as introduced by Zadeh. This is to say, the main difference of two approaches of the relational smooth fuzzy models and IF-THEN smooth fuzzy models is that whether or not it is more practical that the functions be presented through fuzzy numbers of the fuzzy topology or one should restrain to only the constant zero and one fuzzy sets 0 and 1 of the smooth fuzzy relations; We think the first one is preferable and will contribute on development of the IF-THEN smooth fuzzy modeling scheme in this chapter.

Alongside, the other difficulty in smooth fuzzy relational models is that they suffer from the lack of analyzability. Hence, our other motivation has been not only to develop a new TS fuzzy modeling framework using the smooth compositions, but also to construct models which could be used more efficient for study on the numerical behavior, speed of convergence, and the stability of the algorithm. We will apply the method on the nonlinear dynamic of a continuous stirred tank reactor (CSTR) system. Indeed, the nonlinearities, uncertainties or the environmental parametric changes in the dynamic of a CSTR may make the control process to fail [24-28]. We have demonstrated the application of the algorithm to CSTR with the varying parameters and with the uncertain parameters can assist in accomplishment of a precise and effective modeling task without direct intervention of an operator. The simulation results show that the proposed adaptive identification algorithm can handle all the difficult types of such nonlinear system's behavior during the manipulation.

Hence, the rest of this chapter is as follows. First we review fuzzy IF-THEN structures for process modeling and introduce the smooth compositions based on the literature. Then, we employ them for generation of the adaptive fuzzy modeling scheme. Subsequent to that, we introduce the self-learning structure for smooth fuzzy models, to make it sensitive to the parameter variations of the process. After that, we apply the developed structure for a benchmark example and then on a practical example of CSTR, in two different working modes, and also with uncertainty in the parameters. Then, we conclude the chapter.

5.2. Smooth Fuzzy Structures for Process Modeling

This basic structure of smooth fuzzy models has been developed so far, in the past chapter, for designing fuzzy relational dynamic systems, and here we want to employ them for rule-based fuzzy model identification and gaining self -learning dynamics, to be described in the sequent.

Generation of Smooth Rules-Based Fuzzy Models

The aim of this section is to find the optimum parameters for the membership functions to shape it up correspondingly, such that the fuzzy model can make the best approximation of the nonlinear system using the smooth fuzzy compositions. For this aim, first we define the error function as,

$$e(k) = \underline{y}(k) - y(k) \tag{5.1}$$

$$E(k) = \frac{1}{2T}\sum_{t=0}^{T}\big(e(k+t)\big) \tag{5.2}$$

where T is the horizon of training , $y(k)$ is target value of the fuzzy model and $\underline{y}(k)$ is the output of the fuzzy model. The goal is to use this error function to find the optimal shape of the membership functions. Hence, the variables to find are the centers and the widths of the input and output membership functions in the model definition. To simplify the procedure, we consider the normal membership functions with the variables update algorithm, as

$$c_{ij}(k+1) = c_{ij}(k) - \alpha_c \frac{\partial E(k)}{\partial c_{ij}} \tag{5.3}$$

$$\delta_{ij}(k+1) = \delta_{ij}(k) - \alpha_\delta \frac{\partial E(k)}{\partial \delta_{ij}} \tag{5.4}$$

$$d_i(k+1) = d_i(k) - \alpha_b \frac{\partial E(k)}{\partial b_i} \tag{5.5}$$

where $\theta_{ij} = [c_{ij}, \delta_{ij}]$are the parameters of the normal membership functions that give shape to the membership functions, α_c, α_δ and α_b are the step lengths in the gradient based optimization and $i = 1 \cdots, r, j = 1, \cdots, n$ are the numbers of the system rules and the system inputs, and d_i are the parameters to be used in the defuzzification formula, respectively. In order to derive the error derivatives, we study the estimation process in more details. To begin with, we write the gradient descent method formula as follows,

$$\frac{\partial E}{\partial \theta_{ij}} = \frac{\partial E}{\partial \underline{y}} \frac{\partial \underline{y}}{\partial \acute{y}_i} \frac{\partial \acute{y}_i}{\partial \acute{x}_{ij}} \frac{\partial \acute{x}_{ij}}{\partial \mu_{ij}} \frac{\partial \mu_{ij}}{\partial \theta_{ij}}. \tag{5.6}$$

In order to complete the formulation, we need to take the partial derivative of each variable separately.

1- We define the fuzzy variables $\{\acute{x}_1, \cdots, \acute{x}_r\}$ at every time instant as

$\acute{x}_i = [\acute{x}_{i1}, \acute{x}_{i2}, \cdots \acute{x}_{in}] = [\mu_{i1}(x_1), \mu_{i2}(x_2), \cdots, \mu_{in}(x_n)], i = 1 \cdots, r$

where $\mu_{ij}(\cdot)$ is the value of ith membership function for jth input fuzzy set, presented in Equation (4.6).

For making gradient descent method formula, $\frac{\partial \mu_{ij}}{\partial \theta_{ij}}$ can be written as,

$$\frac{\partial \mu_{ij}(\cdot)}{\partial c_{ij}} = exp\left(\frac{-1}{2}\left(\frac{x_{ij}-c_{ij}}{\delta_{ij}}\right)^2\right)\left(\frac{x_{ij}-c_{ij}}{\delta_{ij}^2}\right) \tag{5.7}$$

$$\frac{\partial \mu_{ij}(\cdot)}{\partial \delta_{ij}} = exp\left(\frac{-1}{2}\left(\frac{x_{ij}-c_{ij}}{\delta_{ij}}\right)^2\right)\left(\frac{(x_{ij}-c_{ij})^2}{\delta_{ij}^3}\right) \tag{5.8}$$

2- The estimation of the system output based on the compositional rule inference, can be written as,

$$\acute{y}_i \quad = s - norm\left(t - norm(\acute{x}_i, R_i(\acute{x}, y))\right) \tag{5.9}$$

for the i-th rule R_i, $i = 1, \cdots, r$.We will use the abbreviation S: $s - norm$ and T: $t - norm$ in the followings.

In order to simplify the explanation of the procedure of taking the derivation of $\frac{\partial \acute{y}_i}{\partial \acute{x}_{ij}}$, we assume a system with $j = 2$, and put, $\acute{x}_i = [\acute{x}_{i1}, \acute{x}_{i2}]$ and $c = R_i(\acute{x}, y)$. Then, based on the properties of t-norms, we have,

$$\acute{y}_i \quad = S\big(T(T(\acute{x}_{i1}, \acute{x}_{i2}), c)\big) = S\ (T(\acute{x}_{i1}, c), T(\acute{x}_{i2}, c)) \tag{5.10}$$

We define: $\Lambda_1 = T(\acute{x}_{i1}, c)$ and $\Lambda_2 = T(\acute{x}_{i2}, c)$, then,

$$\acute{y}_i \quad = S(\Lambda_1, \Lambda_2) \tag{5.11}$$

$$\frac{\partial \acute{y}_i}{\partial \acute{x}_{ij}} = \frac{\partial s}{\partial \Lambda}\frac{\partial \Lambda}{\partial \acute{x}_{ij}} = \acute{S}^1\acute{T}^1, j = 1,2. \tag{5.12}$$

If there exist more state variables, $j = n > 2$, $\acute{x}_i = [\acute{x}_{i1}, \acute{x}_{i2}, \cdots \acute{x}_{in}]$ we can follow in the same manner and write as,

$$\frac{\partial \acute{y}_i}{\partial \acute{x}_{ij}} = \acute{S}^{n-1}\acute{T}^{n-1} \cdots \acute{S}^1\acute{T}^1, j = 1, \cdots, n. \tag{5.13}$$

Hence, to derive the gradient descent based training formulation, the derivative of the error will be,

$$\frac{\partial E}{\partial c_{ij}} = \frac{\partial E}{\partial \underline{y}}\frac{\partial \underline{y}}{\partial \acute{y}_i}\frac{\partial \acute{y}_i}{\partial \acute{x}_{ij}}\frac{\partial \acute{x}_{ij}}{\partial \mu_{ij}}\frac{\partial \mu_{ij}}{\partial c_{ij}} \tag{5.14}$$

$$= e(k)\cdot\left(\frac{d_i-\underline{y}}{\sum_{i=1}^{r}\underline{y}}\right)\cdot\left(\acute{S}^{n-1}\acute{T}^{n-1}\dots\acute{S}^{1}\acute{T}^{1}\right)\cdot exp\left(\frac{-1}{2}\left(\frac{x_{ij-c_{ij}}}{\delta_{ij}}\right)^2\right)\left(\frac{x_{ij}-c_{ij}}{\delta_{ij}^2}\right)$$

$$\frac{\partial E}{\partial \delta_{ij}} = \frac{\partial E}{\partial \underline{y}}\frac{\partial \underline{y}}{\partial \acute{y}_i}\frac{\partial \acute{y}_i}{\partial \acute{x}_{ij}}\frac{\partial \acute{x}i_j}{\partial \mu_{ij}}\frac{\partial \mu_{ij}}{\partial \delta_{ij}} \tag{5.15}$$

$$= e(k)\cdot\left(\frac{d_i-\underline{y}}{\sum_{i=1}^{r}\underline{y}}\right)\cdot\left(\acute{S}^{n-1}\acute{T}^{n-1}\dots\acute{S}^{1}\acute{T}^{1}\right)\cdot exp\left(\frac{-1}{2}\left(\frac{x_{ij-c_{ij}}}{\delta_{ij}}\right)^2\right)\left(\frac{\left(x_{ij}-c_{ij}\right)^2}{\delta_{ij}^3}\right)$$

$$\frac{\partial E}{\partial d_i} = \frac{\partial E}{\partial \underline{y}}\frac{\partial \underline{y}}{\partial d_i} \tag{5.16}$$

$$= e(k)\cdot\left(\frac{\underline{y}_i}{\sum_{i=1}^{r}\underline{y}}\right) \tag{5.17}$$

We want to stress that during the fuzzy adaptation process of the present approach, the membership functions represent linguistic terms of fuzzy model interferences, which are transparent and comprehensible to the system operator. This aspect, which lacks in the earlier works using matrix based relational fuzzy models is one of the strengths of fuzzy modelling.

Table 5.2. The proposed algorithm for rule based fuzzy model identification

Concept: the set of input-output data measurements of the system is available and it is desired to identify the smooth fuzzy model for the system. *Initialization Phase:* 1- Membership function selection: Choose a membership function for fuzzification of the input variables. The implemented Guassian membership function. 2- Rule selection: Select r fuzzy rules and compose the fuzzy model using these r rules. Number of rules can be determined heuristically by the designer according to the complexity of the system. 3- Consequent calculation: Choose a smooth fuzzy composition to realize the inference mechanism. This stage makes the functional expansion of the input variables, according to the structure of the employed smooth fuzzy s-norm and t-norm. 4- Model Output: Make the defuzzification of the variables to convert the fuzzy results into the crisp results. *Parameter Learning Phase:* Choose a desired value of accuracy ϵ .

Actually, the blind performance index used at the matrix based relational fuzzy modelling or artificial neural networks based tuning of the membership functions causes in semantically meaningless linguistic terms at the model interfaces, which we could address effectively. In the following, we will illustrate the properties of the proposed algorithm.

Proposition 5.1: The error function constructed based on the Equation (5.2) is a smooth function.

Proof: The interference mechanism makes the functional expansion of the fuzzified input variables using the different polynomical basic functions, which are all smooth. Hence, the output function of the fuzzy model is a smooth function, and therefore, the obtained error function is a smooth function.

Proposition 5.2: The derivative of the error function constructed based on the Equation (5.2) is a smooth function.

Proof: The interference mechanism makes the functional expansion of the fuzzified input variables using the different polynomical basic functions, which all have smooth derivatives [29-30]. Hence, the output function of the fuzzy model has a smooth derivative, and therefore, the obtained error function has a smooth derivative.

Proposition 5.3: The rate of convergence of the parameter learning phase in Table 5.1 to the optimal solution is quadratic.

Proof: **S**ince the derivative of the error function is smooth almost everywhere, the second derivate of the error function is continuous. Hence, when the initial point of the algorithm is sufficiently close to the optimal point and the derivative function is not zero, parameter learning phase of the algorithm will converge quadratic [30].

*Remark 5.1***:** The algorithm in Table 5.1 will converge only if the assumptions in the proof of Proposition 5.3 are satisfied. The most common difficulty is to choose a proper initial point of search in the basin of convergence of the algorithm. The suggested remedy is to run the algorithm from the several random initial points.

5.3. Self-learning of the Smooth Fuzzy Models

Until now, we have developed the algorithm to make a model from the system's input output data. However, for the time varying systems, after making up the initial model of the system, the system parameters changes and the basic model will not remain useful. Therefore, after that the initial fuzzy model comes available, a modification in the abovementioned algorithm can be useful to improve the system performance in an adaptive self-learning scheme. We make this improvement as Table 5.2.

The overall scheme of the self-learning algorithm is shown in Figure 5.1. In the next section, we demonstrate the application of the algorithm in a practical example of chemical processes.

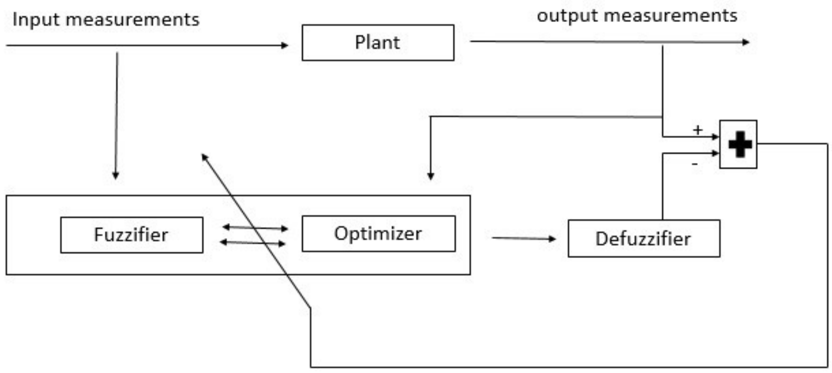


Figure 5.1: Scheme of the proposed self-learning algorithm

Table 5.2. Self Learning Algorithm for the fuzzy model

Concept: Assume that the basic model is available and we want to improve it based on the new measurements of the system. Initialization: Choose a proper ϵ and the simulation horizon. Put $k = 1$. Main Steps: 1- let $k \rightarrow k + 1$. 2- Use the fuzzy model and the system new measurement data to produce the prediction $\hat{y}(k)$. Let, $e(k) = \hat{y}(k) - y(k)$. 3- If $\lvert e(k) \rvert > \epsilon$, then update the parameters of the fuzzy model based on optimization method described above in Section 4, Else return to step (1). 4- End if the simulation horizon terminates; Else return to step (1).

5.4. Case Studies

We have chosen two highly nonlinear systems for examination of the proposed modelling approach. The first system is an example of chaotic time series. We have added parametric uncertainty to demonstrate the effectiveness of the proposed method to the classical modelling scheme.

The second example is about modelling of a continuous stirred tank reactor (CSTR) [24-28]. Different fuzzy models are tested and compared in the uncertain working conditions.

Example 5.1: Application for prediction of Mackey-Glass chaotic time series

In this study, we have employed Mackey-Glass chaotic time series assess the performance prediction of the proposed smooth fuzzy model. Chaos can be inspected commonly in different fields of nonlinear dynamics and can be represented in different forms including by the time series.

The widely assumption on the chaotic time series is that they are nonlinear by nature and extremely sensitive to the initial condition. Therefore, it is a practical technique to evaluate the accuracy of different types of nonlinear models based on their performance in prediction of the chaotic time series.

We have employed the Mackey-Glass time series as,

$$\acute{x} = \frac{ax(t-\tau)}{1+x^{C}(t-\tau)} - bx(t), \tag{5.18}$$

The following parameters are assumed: a=0.2; b=0.1; C=10; with the initial conditions x0=1.2 and τ= 17s. Four different fuzzy models have been trained to predict accurately the generated time series as shown in Fig. 5.2. The error convergence can be seen in Fig. 5.3.

The differences between the sequences derived from the min-max fuzzy model converges more slowly, but note that the range of errors in all the fuzzy compositions is very narrow.

We do not place much emphasis on the min-max error convergence comparison, because the fuzzy min-max model is not differentiable to be solved softly with the gradient descent we applied to the other compositions. Nevertheless, Figure 5.4 does show that the smooth fuzzy models provide better performance with quicker convergence rather than non-smooth compositions.

To study the disturbance- rejection performance of the different fuzzy models, we have evaluated the models through simulation with the parametric change in the chaotic system set to b=0.15.

Figure 5.5 compares the data employed for training to the data employed for validation and prediction.

Figure 5.6 illustrates and compares the disturbance rejection performance of the different fuzzy models. It can be seen that the smooth models have better performances.

To make the model prediction more realistic, the system parameter is randomly varied as disturbances. In Fig. 5.7 the dynamic of the models validation in the noisy environment have been shown and compared. The performance of the model for a validation data set demonstrated that the smooth fuzzy models have a strong disturbance rejection capability rather than classical product-sum compositions and min-max compositions. The noise has been considered as b=0.1+ 0.05*r, where r is assumed to be random signal at every iteration.

To derive a measure of the model accuracy numerically, the employed performance function accounts for the error in the prediction as follows, $F(t) = e(t) \times e(t)$.

It can be seen from Fig. 5.3 to Fig. 5.7 that smooth fuzzy models and the classical product-sum fuzzy model yield compatible results, but the smooth fuzzy models are more robust to the parametric changes and noises and arrive at a better solution in the presence of uncertainties. However, they require slightly more computational efforts than the product-sum fuzzy models. With this type of nonlinear optimization problem, it is difficult to say in general what type of fuzzy compositions scheme works best.

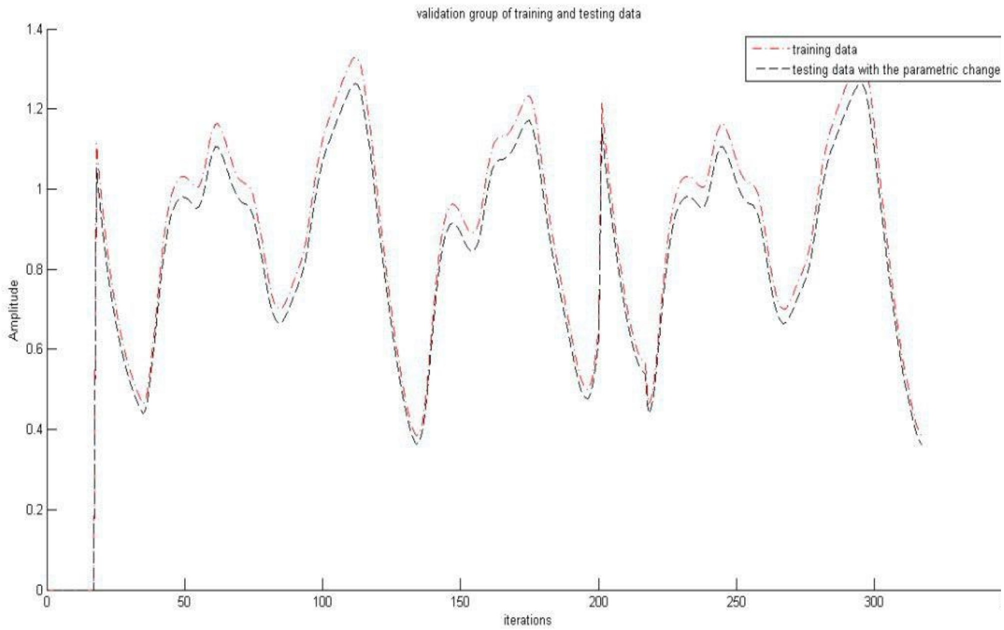


Figure 5.2. Comparison of training data and the validation data

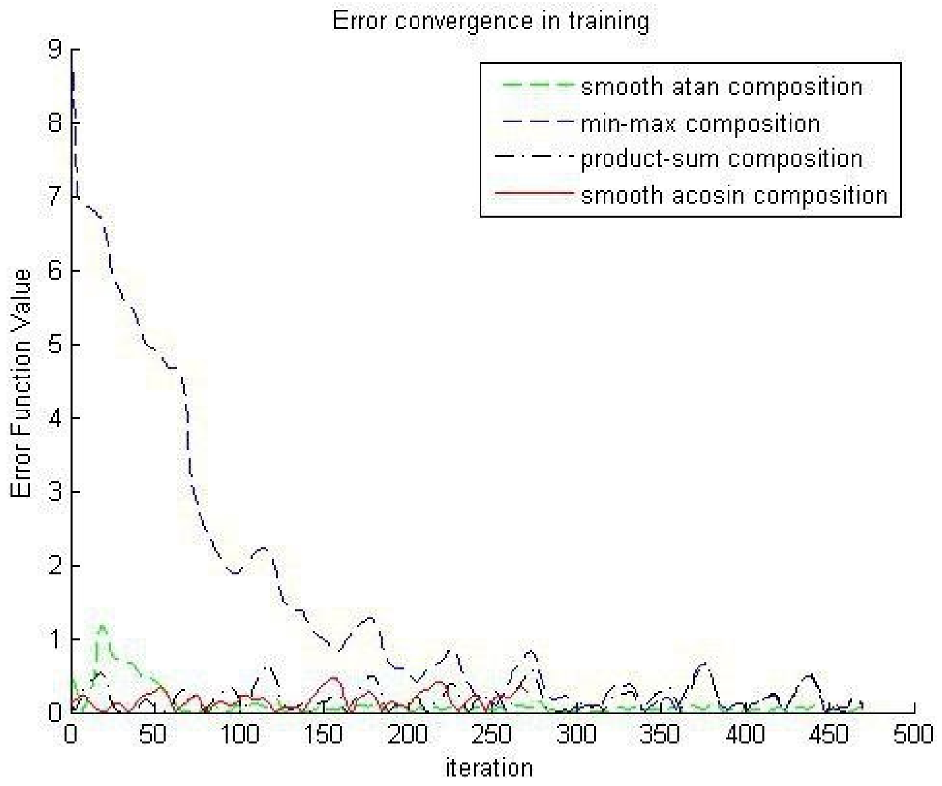


Figure 5.3. Comparison of error convergence for different fuzzy compositions

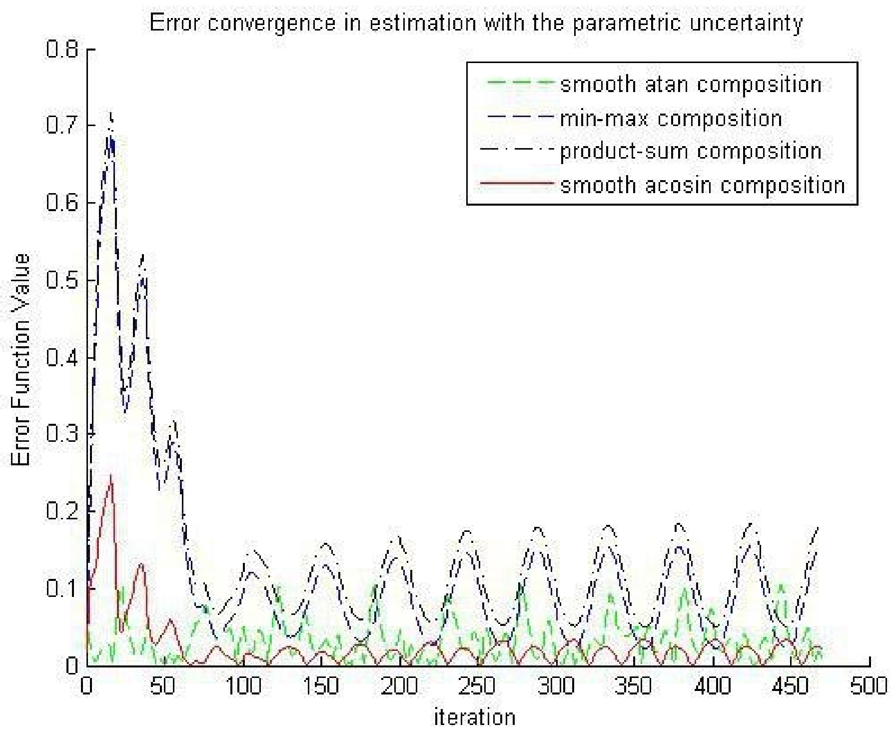


Figure 5.4. Comparison of the performance of the proposed modeling scheme rather than the classical fuzzy scheme in presence of parametric change

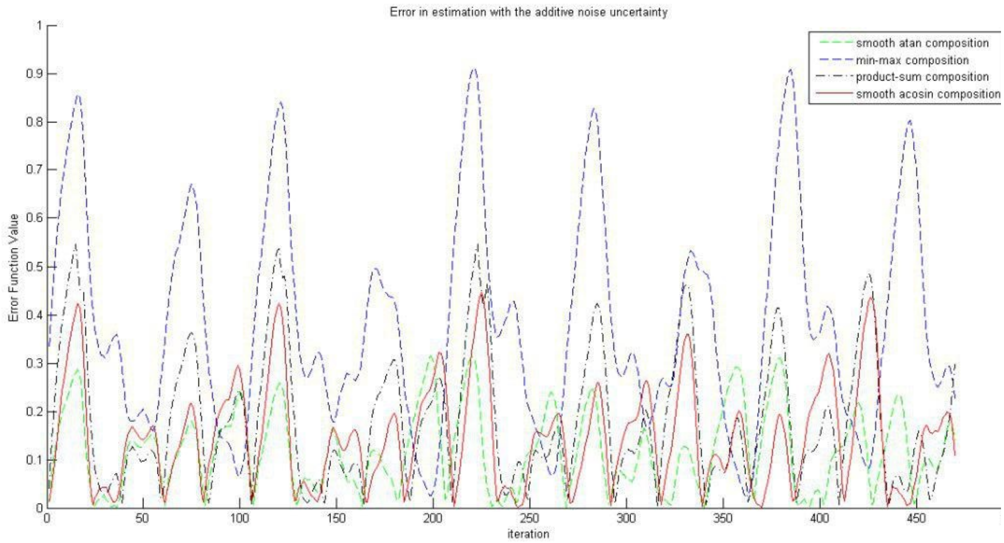


Figure 5.5. Comparison of the performance of the proposed modeling scheme rather than the classical fuzzy scheme in noisy environment

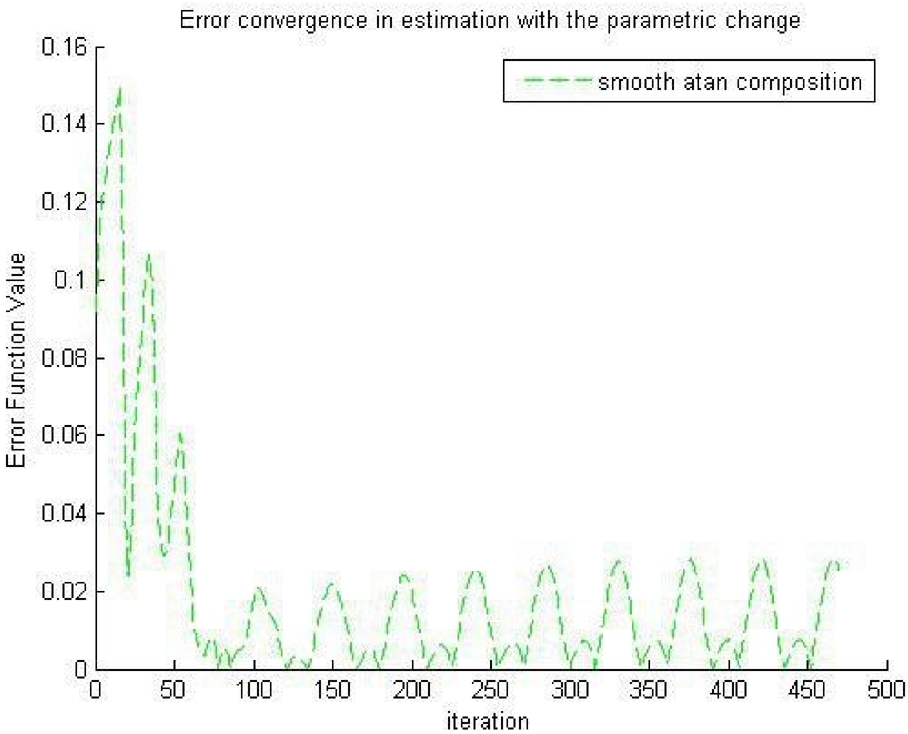


Figure 5.6. Magnified view to the performance of the “atan” fuzzy smooth model in the presence of parametric change in the system

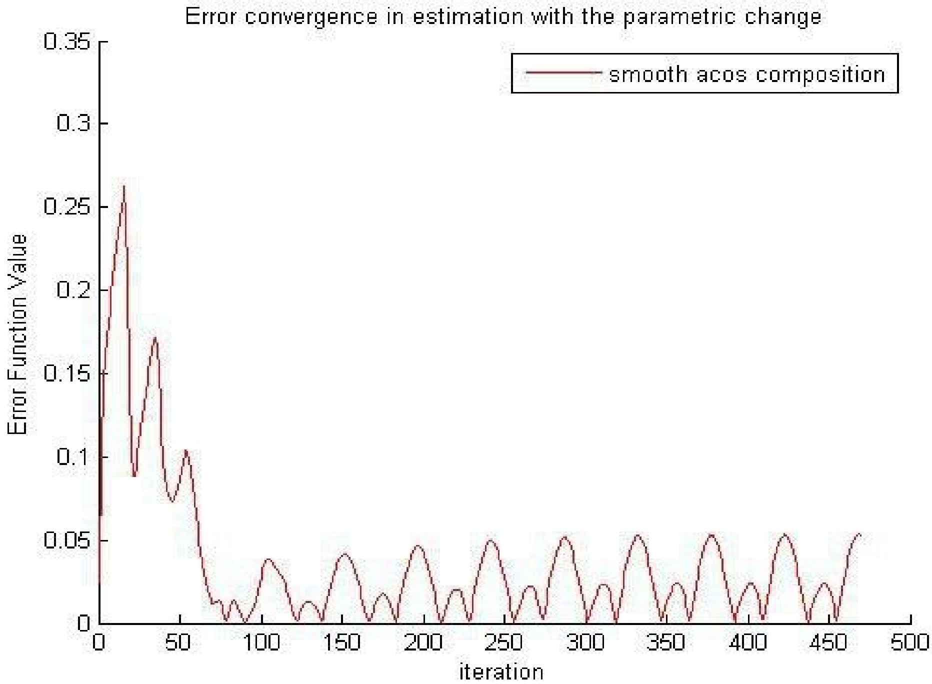


Figure 5.7. Magnified view to the performance of the “acos” fuzzy smooth model in the presence of parametric change in the system

Example 5.2. Evaluation of the Proposed Smooth Fuzzy model with a chemical process

We have considered the dynamic of a highly nonlinear continuous-stirred tank reactor (CSTR) process, as a second benchmark example. This reactor is used commonly in chemical process engineering. Hence, this is a proper modeling problem for test of the algorithm and contrast between different fuzzy compositions. In the reactor, an irreversible, exothermic reaction takes place inside a constant volume to generate a compound A with concentration $C_a(t)$ with the temperature of the mixture $T(t)$. The contents are cooling down through stream of a single coolant with the rate of flow $q_c(t)$. The following equations describe the process model [24, 28]:

$$\frac{dC_a(t)}{dt} = \frac{q}{V}\left(C_{a0} - C_a(t)\right) - k_0 C_a(t) \times \exp\left(\frac{-E}{RT(t)}\right) \qquad (5.19)$$

$$\frac{dT(t)}{Dt} = \frac{q(t)}{V}\left(T_0 - T(t)\right) - k_1 C_a(t) \times \exp\left(\frac{-E}{RT(t)}\right) \qquad (5.20)$$

$$+k_2 q_c(t)\left(1 - \exp\left(-\frac{k_3}{q_c(t)}\right)\right)\left(T_{c0} - T(t)\right)$$

Where the concentration of inlet feed C_{a0} , the rate of process flow q , and the temperatures of the inlet and coolant, respectively, as T_0 and T_{c0}, all are considered as the constant values.

Likewise, $k_0, \frac{E}{R}, V, k_1, k_2$and $k_3$are constants. The nominal values of the process parameters appear in Table 5.4.

$$k_1 = -\frac{\Delta H k_0}{\rho C_p}, k_2 = \frac{\rho_c C_{pc}}{\rho C_p V}, k_3 = \frac{h_a}{\rho_c C_{pc}} \quad (5.21)$$

Considering the product concentration $C_a = 0.1 \mathrm{mol/l}$, the nominal conditions will be,

$$T = 438.5K, q_c = 103.411\ l/min \quad (5.22)$$

The objective in the chemical process is handling the concentration of A,as $C_A(t)$ by proper adjustment of the rate of the coolant flow $q_c(t)$. In the process of fuzzy model making, initially, we have run the mentioned model to derive enough input and output data for the model training. Then, the structure of the considered model is:

$$\hat{C}_a(k+1) = f(\hat{C}_a(k), \hat{C}_a(k-1), \hat{C}_a(k-2), q_c(k-1)) \quad (5.23)$$

The fuzzy model has 3 Gaussian membership functions with the number of rules as $3 \times 3 \times 3 \times 3 = 81$.

Table 5.4. Specification of the CSTR

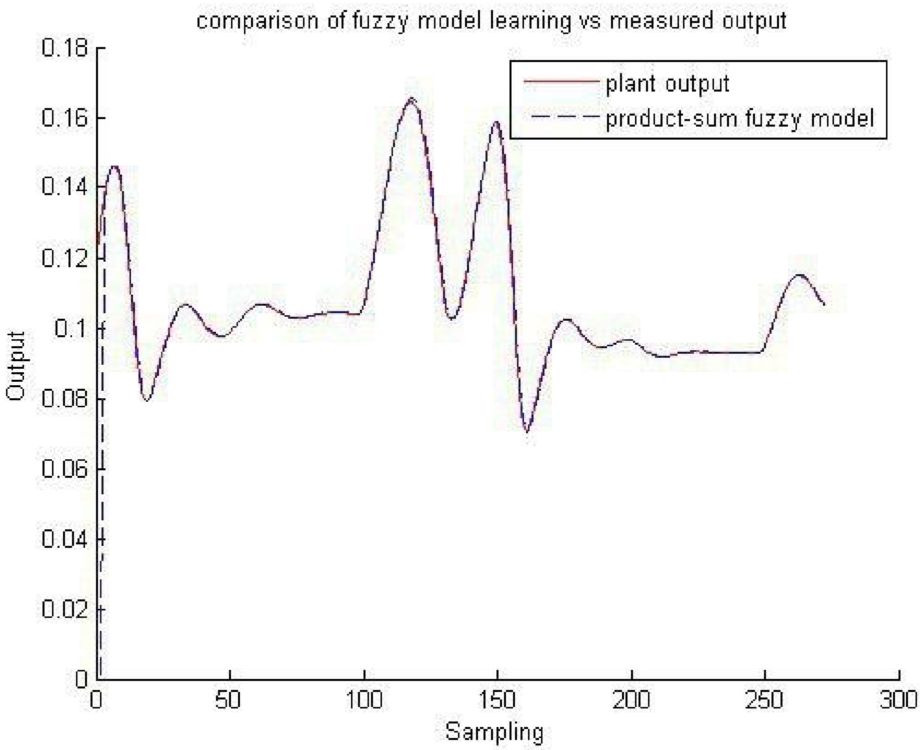
Parameter	Description	Nominal Value
q	Process flow-rate	$100\ l/min$
V	Reactor volume	$100\ l$
k_0	Reaction rate constant	$7.2 \times 10^{10}\ \mathrm{min}^{-1}$
E/R	Activation energy	$10^4 K$
T_0	Feed temperature	$350\ K$
T_{c0}	Inlet coolant temperature	$350\ K$
ΔH	Heat of reaction	$-2 \times 10^5 cal\ /\ mol$
C_p, C_{pc}	Specific heats	$1\ cal\ /\ g/K$
ρ, ρ_c	Liquid densities	$10^3 g/l$

h_a	Heat transfer coefficient	$7 \times 10^5 cal/min/K$
C_{a0}	Inlet feed concentration	$1\ mol/l$

The performance of model for the validation data is depicted in Fig. 5.8. Four different fuzzy compositions are compared: two smooth compositions (based on “atan” and “acos” functions), and two classical fuzzy models using min-max compositions and product-sum compositions.

We have conducted several simulations to inspect how the different set points affect the performance of the system and how different fuzzy structures will track the nonlinear dynamic. Figure 8 shows the open-loop responses upon the different set points when coolant flow rate $q_c(t)$ was changed from 103 l/min to 105, to 110, to 100, to 99 and then to 110. All the developed fuzzy models can reflect the process dynamic behavior almost perfect. The Figure 8 shows the validation error on simulation and the quality of the model is very good.

Fig. 5.9 and Fig. 5.10 demonstrate the disturbance rejection capability of the different fuzzy models. The variation of the coolant temperature T_{c0} are added as the disturbance to the system.



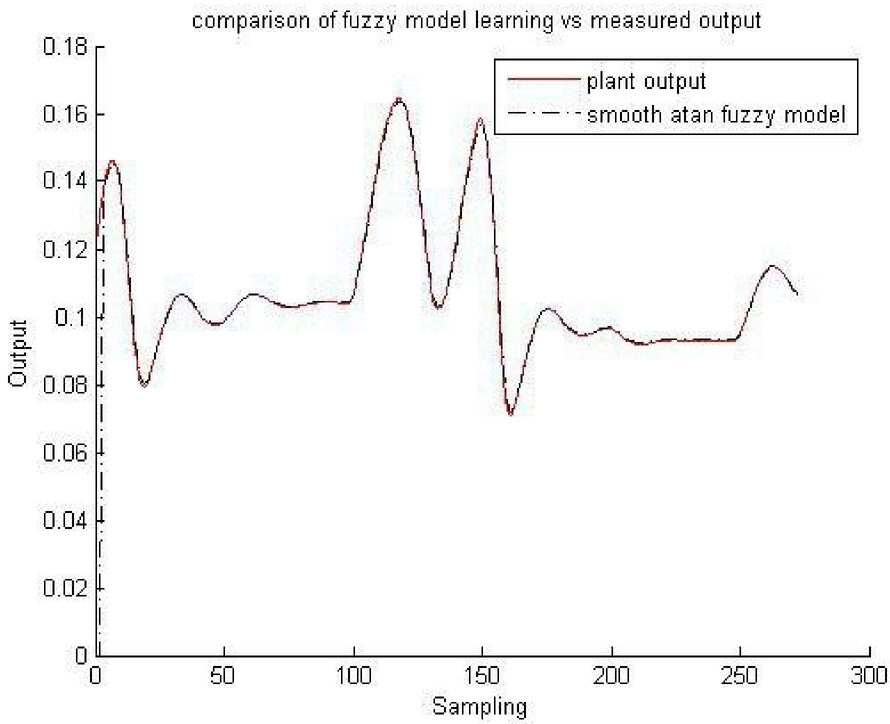


Figure 5.8. The quality of smoothing for the smooth fuzzy model and the classical fuzzy model

(up) Classical model (below) smooth model

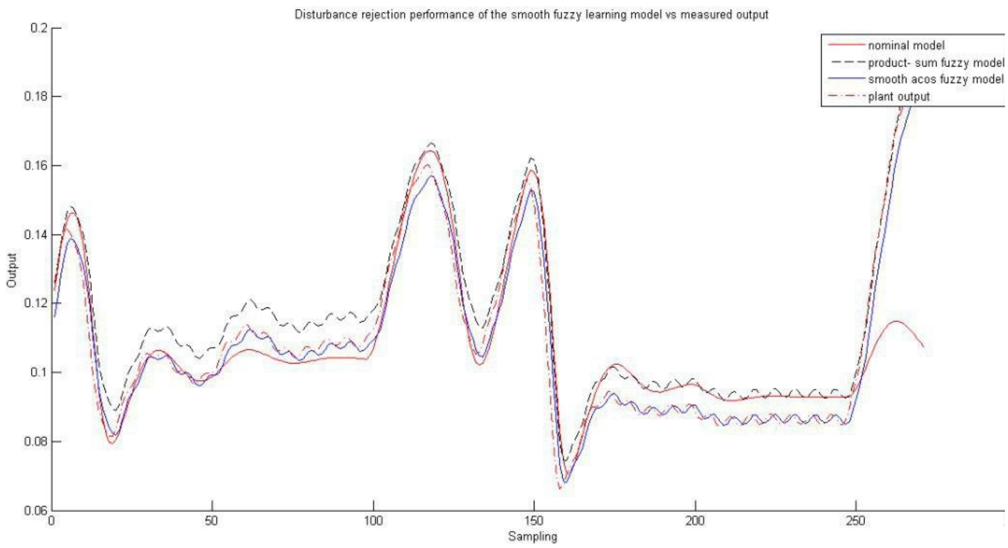


Figure 5.9. Disturbance rejection performance of the proposed smooth fuzzy modeling scheme compared to the classical fuzzy model

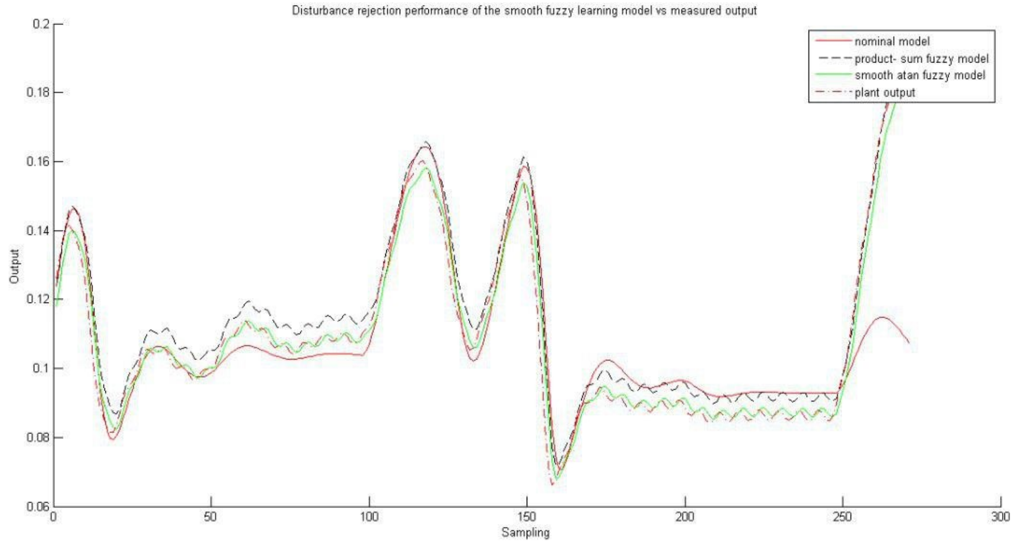


Figure 5.10. Disturbance rejection performance of the proposed smooth fuzzy modeling scheme compared to the classical fuzzy model

The coolant temperature is manipulated as $T_0 = 350 + 5 * \sin(k)$. The dynamic response in the figure shows that the smooth fuzzy models have a strong disturbance rejection capability.

As it can be seen, employing the smooth compositions leads to system prediction with the lower errors. Considering that almost in all the industrial processes on which real time process algorithms are implemented, it is desired to show their smooth dynamics, while the possibility of abrupt changes and parameter variations due to the system faults, aging etc. in the industrial plants are not unneglectable, the importance of making up smooth fuzzy models and the promising applications in the system´s model making and prediction become more transparent.

The key features and main results of developing the presented modelling scheme in the application to CSTR can be briefly summarized as follows,

- The accuracy of modeling with smooth fuzzy compositions are highly better than the classical fuzzy models, which is clear from the comparison of the simulations.
- The transparency of the IF-THEN smooth fuzzy models is much better than the matrix of relational fuzzy models. Hence, the interpretation of the linguistic variable can be useful for better modeling and the subsequent control purpose during the operator interaction.
- The smooth fuzzy model is differentiable and hence, derivative-based iterative optimization algorithms can be applied for better connectivist identification- control approaches.

- The If-THEN smooth model structure has the potential for the theoretical analysis on the robustness and stability properties rather than matrix based relational smooth fuzzy model.
- The smooth compositions bring about higher speed of convergence, based on Proposition 5.3, which results in higher capacity and faster tracking of the parameter changes and changes in the simulations. Our expectation is that, the model based control would damp better uncertainties and variations.
- When the model can track the changes precisely, in the applications of the chemical processes, in particular CSTR, the smooth fuzzy modeling framework makes the model adaptive upon the measurement on a smooth surface of parameters and it enables the calculation of derivative of error surface and fast removal of the local uncertainties.

Bearing this point in mind, we will work for the implementation of the proposed algorithm in the processes that it is required to make up a fast simultaneous measurement and control scheme. The connectivist approach for the measurement based modeling and model based control will lower the down-time production and provide a feasible solution to the challenge of precise and high level of accuracy in the validation and calibration phases, with the minimal level of being underscored by the parameter variations, perturbations and noises. This potentially would give the dynamical systems, possibility of working at higher speeds up to video rate and also utilization for the examination of live processes.

Conclusions

The overall achievement of the chapter is twofold. From theoretical side, one seeks to extend the operational range of applications of smooth fuzzy compositions to make up fuzzy IF-Then models, which comprises lower computational complexities in comparison to the earlier works on the relational fuzzy models, and then, to contribute to the state of smooth fuzzy self-learning algorithm for modeling task of the time variant structures.

We have proposed a novel optimization- based method for fuzzy smooth model construction and compared its performance to the classical fuzzy models. Four different compositions for extracting fuzzy models in the presence of uncertainty have been investigated. The advantages, benefits and limitations of the proposed methods have been investigated though simulation on the benchmark examples. We have investigated the case of parametric uncertainty and a comparison of the speed of convergence has been carried out.

The simulation results validate the smooth fuzzy model superior performance through the comparison and contrast to the classical implementation, on equal conditions, for the chaotic time series model and a CSTR system.

This makes the proposed smooth modeling approach an appealing solution for designing different adaptive identification – controller schemes, especially for the learning of the fast systems, which are supposed to work at the vide rate [23], and for the mutli-target tracking, which considers the multi-objective optimization problems [31, 32].

Look back over the past, with its changing empires that rose and fell, and you can foresee the future too.

Chapter 6

Smooth Fuzzy Model Identification and Model Predictive Control for Dynamic Systems

It has been shown in the previous chapters, that the smooth fuzzy compositions can bring very interesting feature to the fuzzy models ; 1) by structural and smoothness properties they can die out the uncertainties and parameter variation of the systems and environmental noises very fast, 2) they can model the derivate of the plant structure as well as the plant dynamics, 3) they can encompass two (or more) different states of a discontinuous or a switched system and present it by a single smooth fuzzy model, etc. However, the available control schemes rely on the relational matrix fuzzy model, which, from the computational point of view it is not easy to design a controller with them and there exists lack of algorithms for utilization of such interesting fuzzy compositions in the advanced control strategies, due to the large size of the matrixes and the difficulty in the interpretation of the matrixes as the fuzzy model of the large scale plant.

In this chapter, we address this issue and propose a systematic design methodology of combined identification and MPC control through the smooth fuzzy compositions. We will develop a gradient based approach to convert the MPC cost function to an incremental iterative controller design problem to come up with a very fast and simple controller algorithm. The connectivist identification- MPC approach has been designed and tested for the controller with the long-range horizons, in the presence of noises and disturbances.

6.1. Introduction

Model Predictive Control (MPC) is one of the methods that have been considered largely for the purpose of fuzzy logic model based control of nonlinear processes. It can run the complex nonlinear dynamics toward the desired point employing the system data combined with prior knowledge [25, 26, 33, 34]. This control strategy is based on on-line optimization algorithms and can employ the long-range predictive horizons to secure the stability and optimality of the unstable processes. MPC has been employed for the fuzzy logic systems with smooth compositions in [14], where the authors have attempted to make the one step ahead model predictive control of the nonlinear process.

This chapter aims to make the structure of fuzzy model through a harmonious selection of components which simplifies the fuzzy smooth structure, and its subsequent control system. We have extended the earlier works to long-range horizon MPC. Using the long-range horizon, one can predict the impact of the current process input to the future process output, to handle the uncertainty in the system and the model mismatches during the closed loop control performance. Therefore, the proposed algorithm can be employed for the multi-variable systems, to run the system back to the feasibility region in the cases of failure in the actuators, to stabilize the non-minimum phase and dead-time systems.

The other novelty of the findings in this chapter compared to the contribution of AmirAskari and Menhaj [14] is that we made it possible to use any kind of membership function in the modeling of the process for the subsequent control utilization. In the earlier work in [14], they have used the triangular membership function for formulation of the problem. Therefore, although the model has been identified through smooth fuzzy compositions, nevertheless, they then solved the optimization problem using the non-derivative based methods. But, we propose a systemic iterative algorithm without need of solving the optimization problem in every step, which widens the area of application of the algorithm for the industrial applications without the computational power.

This chapter improve the general form of the fuzzy smooth models both theoretically (for mathematical analysis) and practically (for numerical implementation). Hence, our goal here is to formulate the general case for employment of the fuzzy smooth components in the MPC control application.

The chapter is organized as follows. First we present the structure of fuzzy models for dynamic systems which comprises a review on the different fuzzy compositions. Then, the identification problem is addressed and we present an identification scheme employing the smooth fuzzy compositions. In the next section, we employ the model constructed through the proposed identification algorithm for the purpose of model predictive control of the systems in the long horizons. We have provided examples for the proposed uniform identification- control design procedure to show the usefulness of the methods. We will end the chapter with the conclusions.

6.2. Problem Definition

Consider a MIMP system with m inputs $u \in U \subset R^m$ and p outputs $y \in Y \subset R^p$.

$$y(k+1) = f\big(\xi(k), u(k)\big), \tag{6.1}$$

The input vector $u(k) \in R^m$contains the input variables and the regression vector $\xi(k)$ includes the current and lagged inputs and outputs,

$$\xi(k) = \left[\acute{y_1}, \acute{y_2}, \cdots, \acute{y_p}, \acute{u_1}, \cdots, \acute{u_m}\right]^T \tag{6.2}$$

$$\acute{y_I} = \left[y_i(k), y_i(k-1), \cdots, y_i\left(k - n_{yi}\right)\right], i = 1, \cdots, p \tag{6.3}$$

$$\acute{u_J} = \left[u_j(k), u_j(k-1), \cdots, u_j\left(k - n_{uj}\right)\right], j = 1, \cdots, m \tag{6.4}$$

where n_{yi}, and n_{uj}specifiy the number of delayed for the ith output and jth input, respctively. We can define a fuzzy inference for this system as,

R_{li}:if $\xi_1 \in \Omega^{li,1}$ and $\cdots$ and $\xi_p \in \Omega^{li,p}$ and (6.5)

$\xi_{p+1} \in \Omega^{li,p+1}$ and $\cdots$ and $\xi_{p+m} \in \Omega^{li,p+m}$

then $Y_{li}(k+1) = \theta_l\big(\xi(k), u(k)\big), l = 1, \cdots, r.$

where Ω^{li} are the associated interval of existance of the fuzzy set, $\xi_1$is the first element of the vector ξ, and θ_l is the linguistic consequent parameters of the lth fuzzy rule, $\theta = [0,1]^{rp}$ and r is the number of the rules for the system. The output is evaluated from the predicted output corresponding to each rule via the center of gravity method,

$$y_i(k+1) = \frac{\sum_{l=1}^{r} \beta_{li}\theta_{li}}{\sum_{l=1}^{r} \beta_{li}} \tag{6.6}$$

Based on the definition, β is the degree of membership function for the antecedent (states + input) variables as follows,

$$\beta_i: \underbrace{U \times \cdots \times U}_{\text{m times}} \times \underbrace{Y \times \cdots \times Y}_{\text{p+1 times}} \to [0,1]^l \tag{6.7}$$

where the symbol $\times$ represents the Cartesian product in the fuzzy sets. It can be calculated through the s-t composition where s and t are some t-conorm and t-norm, respectively.

Employing different t-norm and s-norm from the above list introduced in chapter 3 to make different compositions, can give rise to a different level of accuracy in modeling of the dynamical systems upon the context, which has been studied in the previous chapter. From them, the smooth fuzzy compositions can make the fuzzy model such that the output is a deferential function of the input variables. Hence, the different schemes of gradient based methods can be used later for the adaptive tuning of the fuzzy model parameters to time varying plant parameters and the uncertainties of the plant. We want to employ this idea for rule-based fuzzy model identification and long-range control horizon model predictive control, to be described in the sequent.

6.3. Generation of Smooth Fuzzy Model

In the process of system identification, we want to train the fuzzy model to capture the functioning of the real plant. We can view this process as an application of the optimization methods to the fuzzy model, very similar to the process of training neural networks, where the least square optimization problem is solved. At every sampling time, we consider a target value $t_i(k)$ for the system´s output $y_i(k)$ and correspondingly, define the overall performance index Ξ of the model as

$$J = \frac{1}{2}\Xi(t-y)^2 \tag{6.8}$$

The parameters of the fuzzy model can be tuned through solving the minimization problem of the performance index. It leads us to have a general method of modifying the fuzzy model at every sampling time k. The goal is to use the performance index to find the optimal shape of the membership functions. Therefore, the variables to find will be the center and the width of the fuzzy membership functions. To simplify the procedure, we consider the normal membership functions with the gradient based variables update algorithm,

$$\rho_{ld}(k+1) = \rho_{ld}(k) - \alpha_\rho \frac{\partial J(k)}{\partial \rho_{ld}} \tag{6.9}$$

$$\theta_{li}(k+1) = \theta_{li}(k) - \alpha_b \frac{\partial J(k)}{\partial \theta_{li}} \tag{6.10}$$

where $\rho = [c_{ld}, \delta_{ld}]$are parameters of the normal membership functions, α_ρ and α_b are the step lengths in the gradient based optimization and $l = 1 \cdots, r, d = 1, \cdots, m+p$ are the number of the system rules and the system inputs, respectively. The error derivatives are straightforward and we study the identification process in more details in appendix 6.1.

6.4. MPC for Smooth Fuzzy Model

In this section we intend to employ the smooth fuzzy model developed in the last section to construct a uniform on-line identification- MPC control framework for the nonlinear processes. In order to facilitate the explanation of the algorithm development, we consider a single-input single output dynamics; however, we emphasize that the results are readily extendable to the multi-input multi-output processes.

We consider the following cost function for the model predictive control purpose,

$$J = \frac{1}{2T}\sum_{t=1}^{T}[e^2(k+t) + \lambda u^2(k+t-1)] \tag{6.11}$$

where the tracking error is defined as

$$e(k+t) \coloneqq r(k+t) - y(k+t) \tag{6.12}$$

$r(k+t)$ is the reference and $y(t+k)$ is the output of the plant both at $(k+t)$ th sampling time instant. We choose $\lambda \geq 0$ as the penalty factor and T as the control horizon. Based on the minimization of the cost function J we derive a sequence of the optimal increase in input signal $\Delta u(k), \cdots, \Delta u(k+T-1)$, however, just the first increase signal is applied to the system. At the next time instant $k+1$, the whole process will be repeated.

To derive the control law, we consider the simple case, where the input signal of the process is comprised of two membership functions as,

$$u(k) = \beta_1 u_1(k) + \beta_2 u_2(k) \tag{6.13}$$

where $\beta_1(k) = \frac{\mu_1}{\mu_1+\mu_2}, \beta_2(k) = \frac{\mu_2}{\mu_1+\mu_2}$. The input signal at the next time step will be,

$$u(k+1) = u(k) + \alpha \Delta u(k) \tag{6.14}$$

or in other formulation,

$$u(k+1) = [\beta_1, \beta_2]\begin{bmatrix} u_1(k) \\ u_2(k) \end{bmatrix} u(k) + \alpha[\beta_1, \beta_2]\begin{bmatrix} \Delta u_1(k) \\ \Delta u_2(k) \end{bmatrix}$$

The incremental input signals $\Delta u_1(k)$ and $\Delta u_2(k)$ are given as,

$$\Delta u_\varrho(k) = \frac{-\partial J}{\partial u_\varrho(k)}, \varrho = 1,2. \tag{6.15}$$

with the length step $\alpha(0 < \alpha \leq 1)$.

Based on the definition, we have,

$$\frac{\partial J}{\partial u_\varrho(k)} = \frac{1}{T}\sum_{t=1}^{T}\left[-\big(r(k+t) - y(k+t)\big)\frac{\partial y(k+t)}{\partial u_\varrho(k)} + \lambda u(k+t-1)\frac{\partial u(k+t-1)}{\partial u_\varrho(k)}\right] \tag{6.16}$$

We assume free $u(k)$ and $(k+t) = u(k), t = 1,2,\cdots,N-1,$. When one considers, $i,j = 1, y = \frac{\sum_{l=1}^{r}\beta_l\theta_l}{\sum_{l=1}^{r}\beta_l}$ and $Y_l(k+1) = \theta_l\big(\xi(k), u(k)\big), l = 1,\cdots,r,$the state vector and inputs become,

$$\xi(k) = [\acute{y}, \acute{u}]^T \tag{6.17}$$

$$\acute{y} = \big[y(k), y(k-1),\cdots, y\big(k-n_y\big)\big] \tag{6.18}$$

$$\acute{u} = [u(k), u(k-1),\cdots, u(k-n_u)] \tag{6.19}$$

The increment of the input signal can be obtained by taking the derivatives,

$$\frac{\partial y(k+t)}{\partial u_\varrho(k)} = \frac{\sum_{l=1}^{r}\theta_l\left(\frac{\partial\beta_l(k+t)}{\partial u_\varrho(k)}\right) - y(k+t)\sum_{l=1}^{r}\frac{\partial\beta_l(k+t)}{\partial u_\varrho(k)}}{\sum_{l=1}^{r}\beta_l(k+t)}, \varrho = 1,2 \tag{6.20}$$

In the Equation (6.20), the value of the derivative $\frac{\partial\beta_l(k+t)}{\partial u_\varrho(k)}$ can be computed after the model and inference structure selection. For the sake of illustration, we consider the following model structure defined by the smooth fuzzy composition,

$$\beta_l(k+t) = S\Big(T\big(R, U(k+t-1), Y(k+t-1)\big)\Big) \tag{6.21}$$

where U and Y are fuzzy values in [0,1].

For input prediction horizon with $t = 1$, (i.e. $\frac{\partial y(k+1)}{\partial u_\varrho(k)}$), the only term depending on $u(k)$ is $U(k)$

Therefore,

$$\frac{\partial\beta_l(k+1)}{\partial u_\varrho(k)} = \frac{\partial S}{\partial T(\cdot,\cdot)}\frac{\partial T(\cdot,\cdot)}{\partial U(k)}\frac{\partial U(k)}{\partial u_\varrho(k)} \tag{6.22}$$

where

$$\frac{\partial U(k)}{\partial u_\varrho(k)} = \frac{\partial f(\acute{u}(k), \rho_u)}{\partial u_\varrho(k)}$$

and $f(\cdot,\cdot)$ is the membership function with the parameters $\rho_u = [c_u, \delta_u]$, c_u is the membership function center and δ_u is the membership function width, obtained in the identification phase.

For the input prediction horizon with $t = 2$, the terms depending on $u(k)$ is $\beta(k+1)$, and

$$\beta_l(k+2) = S\Big(T\big(R, U(k+1), Y(k+1)\big)\Big) \tag{6.23}$$

Therefore,

$$\frac{\partial\beta_l(k+2)}{\partial u_\varrho(k)} = \frac{\partial S}{\partial T(\cdot,\cdot)}\left[\frac{\partial T(\cdot,\cdot)}{\partial U(k+1)}\frac{\partial U(k+1)}{\partial u_\varrho(k)} + \frac{\partial T(\cdot,\cdot)}{\partial Y(k+1)}\frac{\partial Y(k+1)}{\partial u_\varrho(k)}\right], \varrho = 1,2$$

and

$$\frac{\partial Y(k+1)}{\partial u_\varrho(k)} = \frac{\partial Y(k+1)}{\partial y(k+1)}\frac{\partial y(k+1)}{\partial u_\varrho(k)}, \varrho = 1,2$$

where

$$\frac{\partial Y(k+1)}{\partial y(k+1)} = \frac{\partial f_l(\acute{y}(k+1),\rho_y)}{\partial y(k+1)},$$

and $\frac{\partial y(k+1)}{\partial u_\varrho \ (k)}$ is calculated above in (6.20).

For $i = 3$, the only terms depending on $u(k)$ is $\beta(k+2)$, hence,

$$\frac{\partial \beta_l(k+3)}{\partial u_\varrho(k)} = \frac{\partial S}{\partial T(\cdot,\cdot)}\left[\frac{\partial T(\cdot,\cdot)}{\partial \beta_l(k+2)}\frac{\partial \beta_l(k+2)}{\partial u_\varrho(k)}\right], \varrho = 1,2 \tag{6.24}$$

which is calculated above. For $i > 3$, $\beta_l(k+i-1)$ is the only term depending on $u(k)$ which can be calculated recursively.

Remark 6.1: We can extend the control design procedure and the identification process readily for other definitions of the membership function involvement or to the systems with multi inputs, multi outputs.

Lemma 6.1: Provided that there exists a feasible solution for the control problem in (6.11). Then the system dynamics will converge to track the reference signal as $k \to \infty$.

Proof : Assume that we are at the time k and implement the optimal input $u(k) = u_0^*$ that runs the system to the state $x(k+1)$.

$$J_{x(k)} = \min_{\mathrm{u(k)}} \frac{1}{2T}\sum_{t=1}^{T}[e^2(k+t) + \lambda u^2(k+t-1)]$$
(18)

$$J_{x(k)} = \min_{\mathrm{u(k)}} \frac{1}{2T}[e^2(k+1) + \lambda u^2(k)] + J_{x(k)}(k+1) \tag{6.25}$$

At this time, we can determine the associated optimal control input to the system over the horizon 1 to N+1,

$$J_{x(k+1)} = \min_{\mathrm{u(k+1)}} \frac{1}{2T}\sum_{t=1}^{T}[e^2(k+t+1) + \lambda u^2(k+t)] \tag{6.26}$$

However, we can employ the previous sequence of optimal moves followed by zero as well : $u(k+1) = u_0^*$. As this sequence of input is not optimal hence,

$$J_{x(k+1)} \le J_{x(k)} - \min_{\mathrm{u(k)}} \frac{1}{2T}[e^2(k+t) + \lambda u^2(k+t-1)] \tag{6.27}$$

As the value of minimization is positive for $(e,u) \neq (0,0)$, hence the sequence of the optimal costs is stricktly decreasing for all $(e,u) \neq (0,0)$., i.e. $J(k+2) \le J(k+1) \le J(k)$. From the other hand, by the definition in equation (7), we have $0 \le J$. It means that the the sequence of the cost functions $J(k), J(k+1), J(k+2)$ are converging to zero and $e \to 0$ as well.

Corollary 6.1: The feasibility of the control input and state variable implies that the MPC controller will run the state trajectory to zero.

Proof: It can be proved by change of parameters from Lemma 1, considering that,

$$J(\cdot) > 0, (e,u) \neq (0,0), J(e=0, u=0) = 0. \tag{6.28}$$

For the unstable systems, the question will be how to determine the interval N or at least an upper bound such that the system enters the positive invariant set. Several algorithms for the proper selection of the control horizon N have been introduced in the literature.

Lemma 6.2: The obtained control law is continuous and smooth.

Proof: Since the control law u is obtained by the derivation and linear combination of some smooth and continuous functions, the control input is continuous and smooth.

Lemma 6.3: The cost function $J(\cdot)$ is convex, continuous and smooth.

Proof: Since the cost functionu is obtained by the derivation and linear combination of some cosine smooth and continuous functions, it is continuous and smooth. The convexity of the cost function can be proved easily from the equation (6.11).

Corollary 6.2: The control function is the optimal control sequence and the system trajectory is the corresponding optimal trajectory.

Proof: The corollary can be concluded from the convexity property of the cost function $J(\cdot)$ in Lemma 6.3.

The overal proceedure of the connectivist smooth fuzzy identification and MPC control scheme is portrayed in the Figure 6.1.

Remark 6.2: As it is shown in Figure 6.1, we could join the learning capacities of the adaptive modelling scheme to the iterative method of controller design to reach a uniform framework with the parallel processing features.

Remark 6.3: At each step of the adaptation for the fuzzy model, the membership functions linguistically express terms for the model interference that are understandable to a human. This aspect, which has been forgotten in the earlier works using relational fuzzy models [14], is one of the strengths of fuzzy modelling. Actually, the blind performance index used at the relational fuzzy modelling or artificial neural networks based tuning of the membership functions causes in semantically meaningless linguistic terms at the model interfaces, which we could address effectively.

Remark 6.4: In the present work, we have developed a systematic incremental controller using the smoothness and continuity properties of the model structure, to employ the online membership function calibration of the model with the least on-line computational burdens. This is while, the MPC design in typically based on the minimization of a non-convex quadratic performance index.

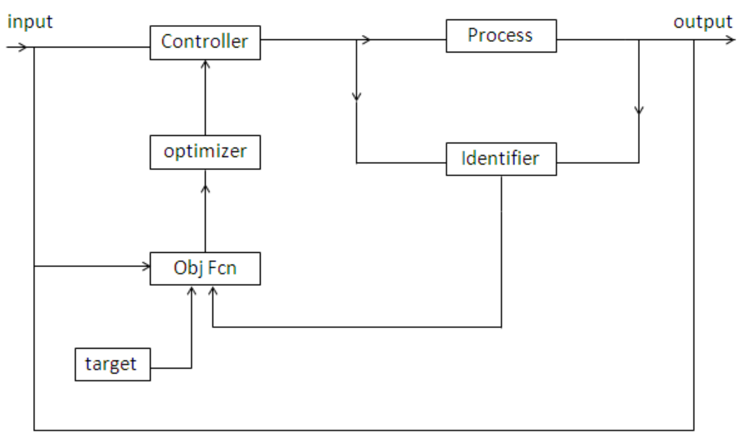


Figure 6.1: The overall scheme of the presented connectivist

Lemma 6.4: The rate of convergence of the control function is quadratic.

Proof: Based on the Lemma 3, the derivative of the control function is smooth almost everywhere, and it´s second derivate is continuous. Hence, when the initial point of the control signal and the system states are sufficiently close to the optimal points and the derivative function is not zero, the optimization algorithm will converge quadratic.

Corollary 6.3: The smooth fuzzy MPC control function will converge faster than the classical fuzzy MPC and to a more stable solution.

Proof: Considering the quadratic rate of convergence for the control function in smooth fuzzy models and the linear rate of convergence of the classical fuzzy model, the corollary can be concluded straight from the Lemma 6.4.

We show the effectiveness of the proposed uniform smooth fuzzy modeling in Section 2 on a time series example and the present connectivist identification-control approach on a nonlinear non minimum phase example below.

6.5. Illustrative Example

In this section, we intend to illustrate the effectiveness of the proposed approach through an example based on [33, 34] for smooth fuzzy IF-THEN model identification – MPC control of a non-minimum phase system. We also study the role of extending control horizon on the overall performance of the controlled system. Consider the following discrete time nonlinear system,

$$y(k+1) = -u(k) + 1.2u(k-1) + 1.4exp\left(-y^2(k)\right) - 0.6y(k-1) \qquad (6.28)$$

The open-loop response shown in Fig. 6.2 indicates that the process is indeed highly nonlinear.

Initially we have modeled the system through the proposed smooth fuzzy modeling scheme. Then, we controlled the system in different control horizons T. We have taken $\alpha = 0.5$ and $\lambda = 0$ for this purpose.

We have run several simulations to examine how the change of the control horizon and the selection of fuzzy composition impact on the effectiveness and quality of the controller. We have examined the control performance for different time horizons.

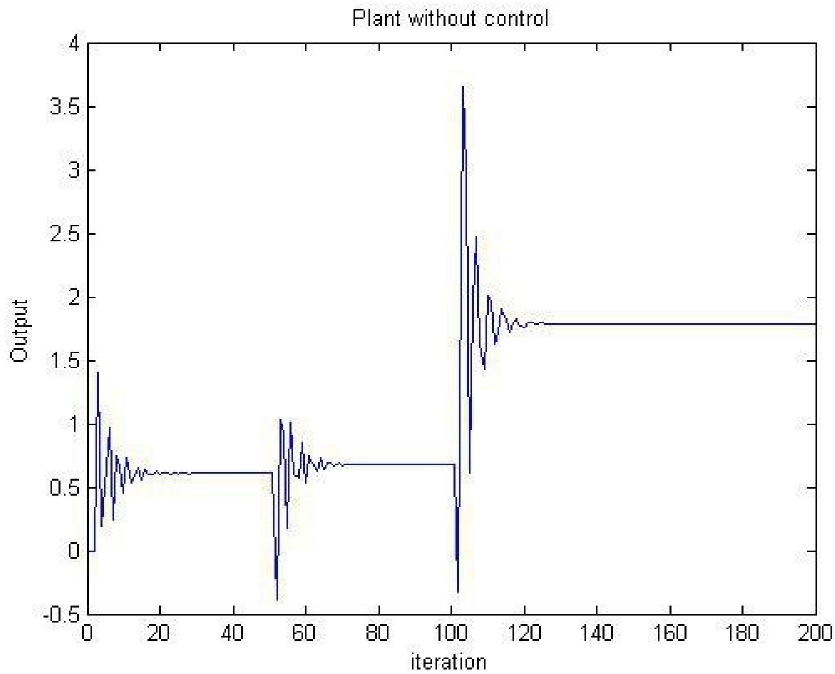


Figure 6.2. open loop of the plant

Fuzzy controller design: We have made several simulations to inspect the quality and effectiveness of the smooth fuzzy controller. In course of the simulations, we changed the set point of the plant in a train of pulses. The dynamic response of the system and system input are depicted in the same figure. We have considered several different set points to properly see the performance of the controller to the system. Fig. 6.3 demonstrates the systems´ responses in the closed-loop for three control horizons using four different fuzzy compositions: two "atan" and "acos" smooth fuzzy controllers and two classical min-max and product-sum fuzzy controllers. Apparently, the control dynamics with all four compositions are good.

Disturbance rejection performance: To make the control problem more realistic, the value of parameters is randomly varied as disturbances.

Both classical fuzzy structure and smooth fuzzy structure are used in the comparative study of the performance of the MPC controller in Fig. (6.3) - (6.6). The dashed lines demonstrate the response of the control action where the smooth fuzzy MPC is applied and the dotted lines depict the response to the control action from classical fuzzy MPC. With a contrast of the obtained results, one can conclude that the smooth fuzzy controller outperformances the classical MPC technique.

In simulation, different types of disturbances and noisy environment also have been considered to affect the system.

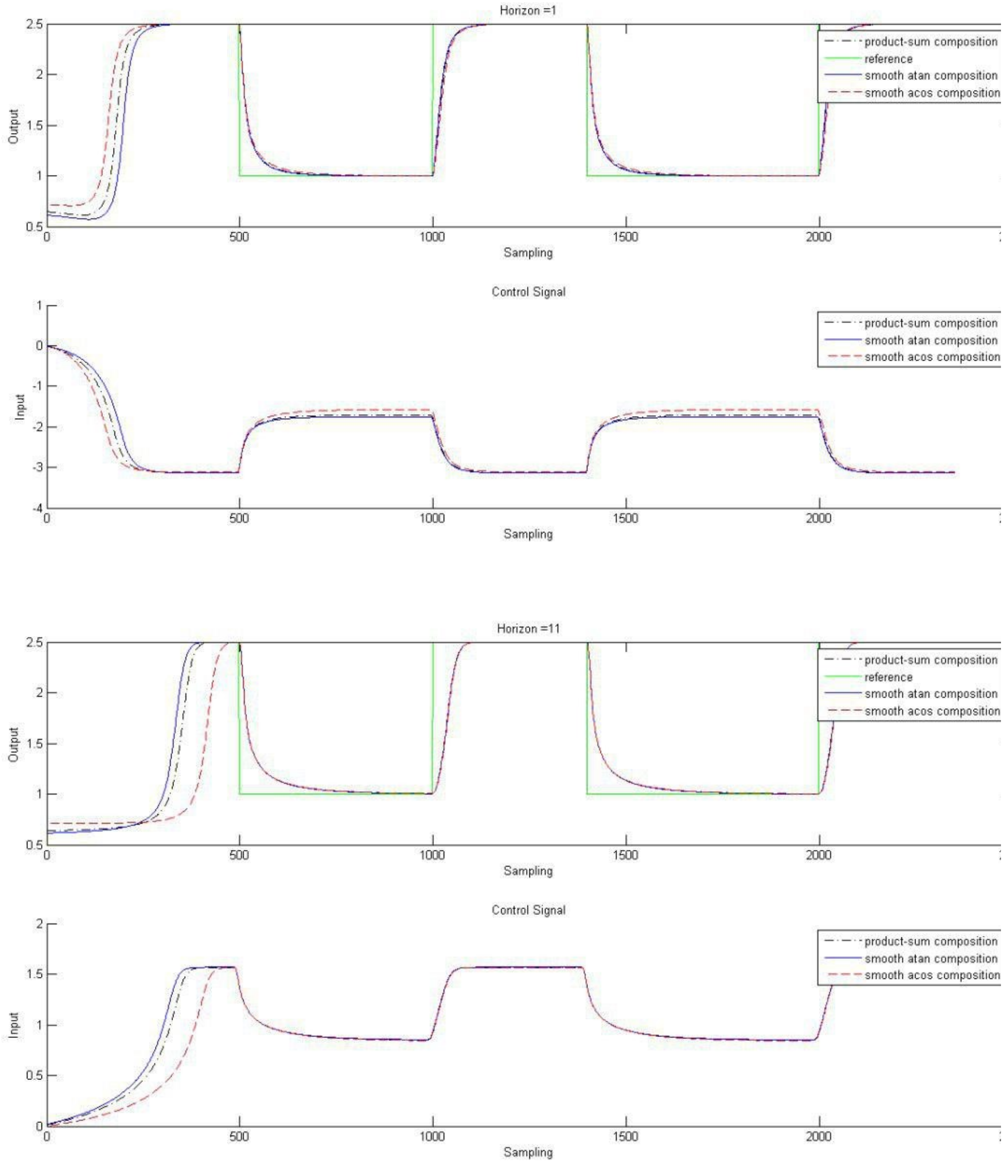


Figure 3. Comparison of three compositions

(top) short term control horizon H=1, (bottom) long term horizon H=11.

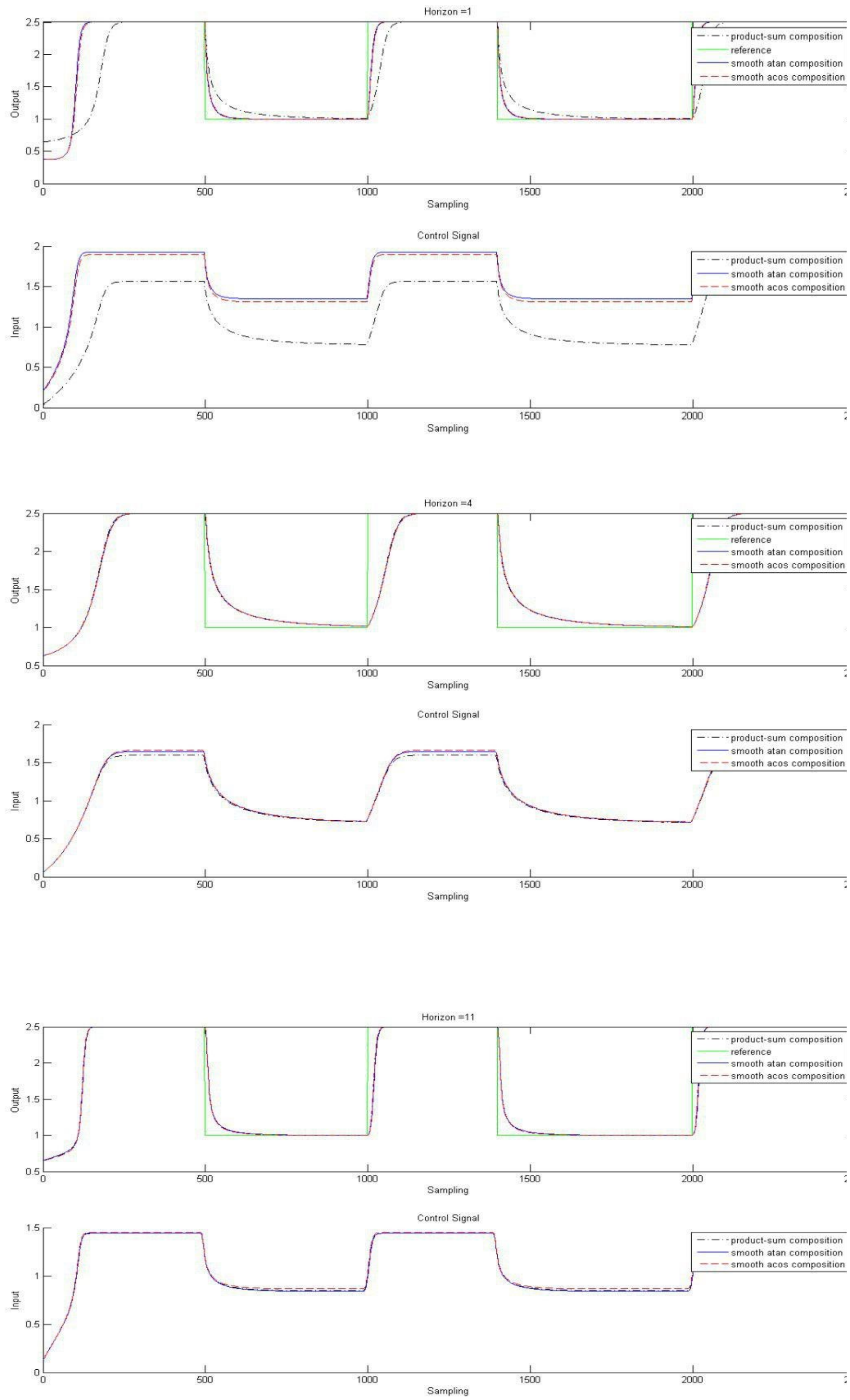


Figure 6.4. Performance of the proposed MPC scheme with three fuzzy compositions in noisy environment. (top) short term control horizon H=1, (middle) medium term horizon H=4, (bottom) long term horizon H=11.

Additive noise: To study the impact of the smooth model prediction on the control performance, we have considered noisy environment which is added to the system. Obviously, this causes degradation in the normal performance of the controllers. A focus on Fig. 6.4. demonstrated the impact of the smooth fuzzy MPC for the noisy conditions, where it is proved to be more robust, rather than the classical scheme. We have considered the followings as the additive noise to the system.

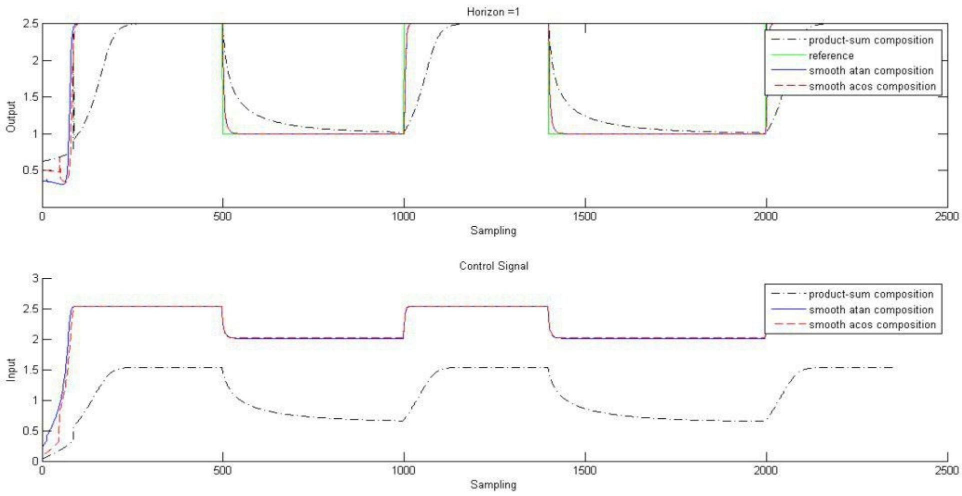
$$y(k+1) = -u(k) + 1.2u(k-1) + 1.4exp\left(-y^2(k)\right) - 0.6y(k-1) + 0.05 * \mathrm{R} \qquad (6.29)$$

where R is a random noise signal.

Parametric change: We also have studied how the change in parameters (usually not measurable) can impact the controller performances. Obviously, this leads to degradation in the normal performance of the controllers. From the other hand, a focus on Fig. 6.5. demonstrates that the proposed smooth scheme outperforms the classical controller in the disturbance rejection. We have considered as the parametric change to the system.

$$y(k+1) = -u(k) + 1.2u(k-1) + (1.4 + 0.08 * R)exp\left(-y^2(k)\right) - 0.6y(k-1) \qquad (6.30)$$

with R as defined above.



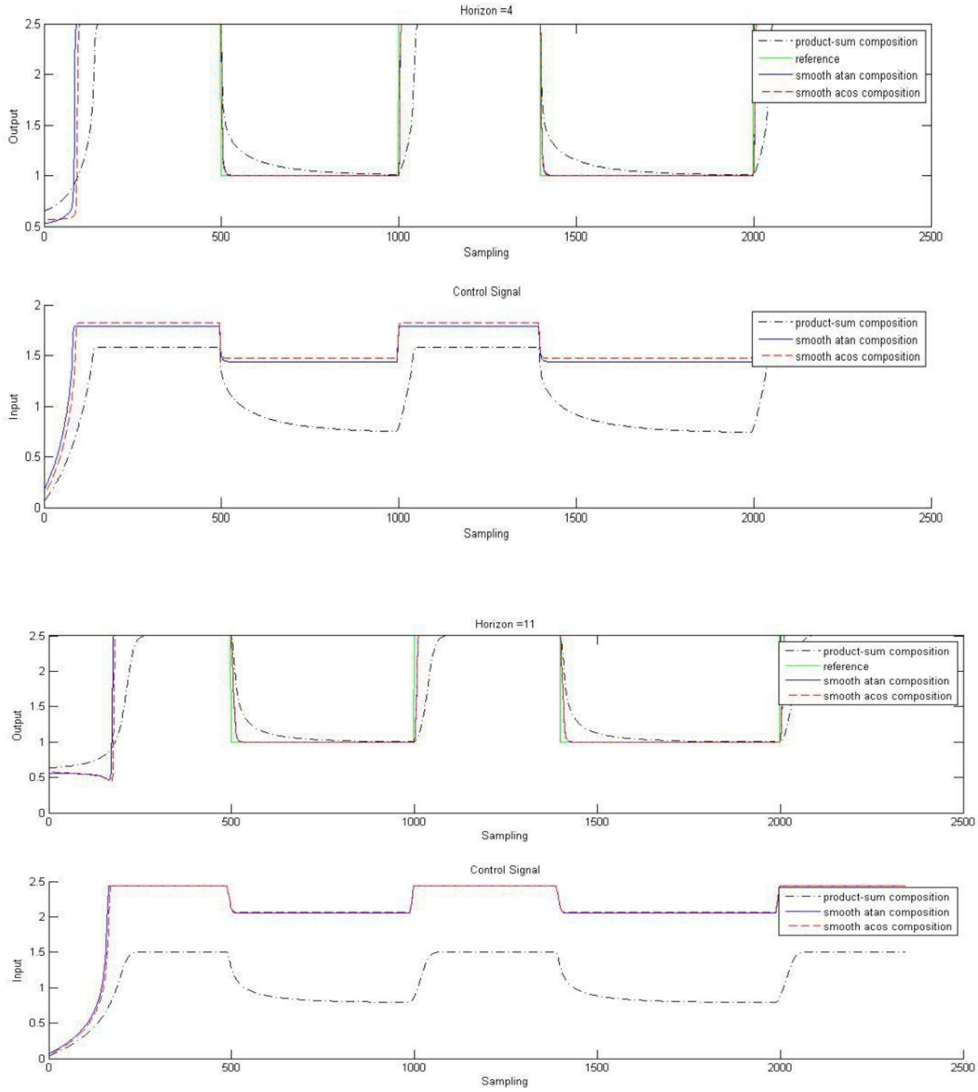
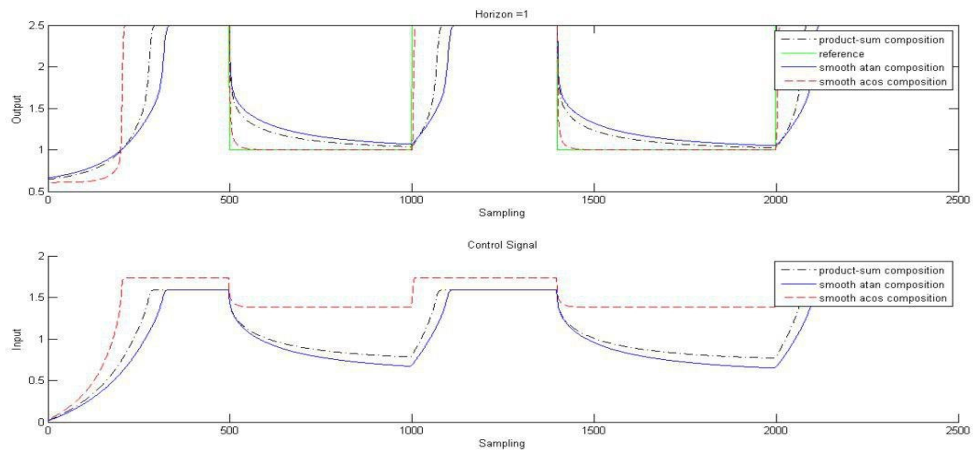


Figure 6.5. Comparison of three compositions with change in the parameters of the plant. (top) short term control horizon H=1, (middle) medium term horizon H=4, (bottom) long term horizon H=11.



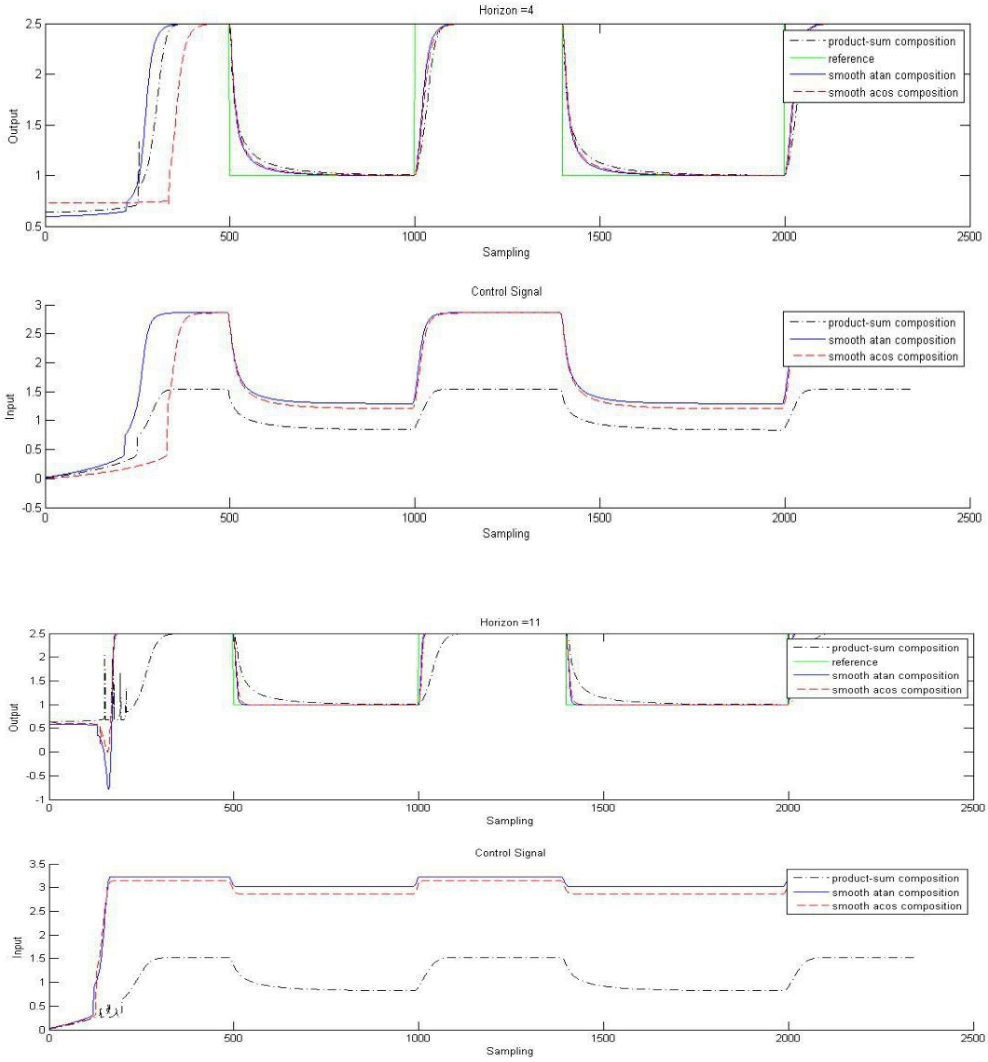


Figure 6.6. Comparison of three compositions with additive time-varying disturbance. (top) short term control horizon H=1, (middle) medium term horizon H=4, (bottom) long term horizon H=11.

<u>Additive time-varying disturbance:</u> We also have studied how time-varying disturbance (usually not measurable) can impact the controller performance. This leads to degradation in the normal performance of the controllers, too. It is demonstrated in Fig. 6.6 that the smooth fuzzy controllers are more robust. We have included the parametric changes to the system. The controller managed to achieve desired reference trajectory under constant disturbance,

$$y(k+1) = -u(k) + 1.2u(k-1) + 1.4exp\left(-y^2(k)\right) - 0.6y(k-1) * 0.05 * \sin(\mathrm{k}) \qquad (6.31)$$

The study of dynamic response shows that the fuzzy smooth predictive controllers have strong disturbance rejection capabilities, in all the figures, in comparison to the classical fuzzy systems.

As it can be seen, employment of the smooth compositions in the fuzzy implementation, has reduced the level of overshoot and made the performance of the controller much smoother, rather than the classical fuzzy modeling and control scheme. Besides, the comparison of the results of study the same initial conditions of different smooth versus classical fuzzy models shows that the smooth representation causes a faster response in the transient state of the system. Hence, it highlights the importance of making up smooth fuzzy models and smooth fuzzy controller and the promising applications they may have in the system modeling and prediction become more transparent.

In all the figures, it can be distinguished that both employed smooth “atan” and “acos” compositions provide satisfactory performance and more stable performance with superiority of the latter classical fuzzy controllers. More analytical study on the robustness and stability of the control system is beyond the scope of the current chapter. The interested reader can consult the reference [29] – [31].

The inspection of results, we obtained so far, could reveal that

i) Converse to the earlier contributions on smooth fuzzy compositions where the whole structure of the system was controlled based on the single relational matrix of operation, we are able to control and run every state and output of the system separately. Hence, it will be more easy to handle the industrial multivariable system here, while manipulating the relational matrix of the industrial plants requires much more computational requirement and lacks the model transparency and interpretability.

ii) As it is depicted in Figures 6.3-6.6, we can vary the control horizon through the presented approach. Hence, in the cases of unstable system, non-minimum phases and/ or delayed systems, it will be possible to increase the control horizon large enough such that the system enters the basin of attraction. While, this is not possible through the one-step ahead prediction strategy of the earlier work of AmirAskari and Menhaj [14].

iii) For the systems with dead time, we can increase the control horizon long enough to check the effect of every input to the system. Hence, it makes the algorithm very applied for the real industrial plants. This is while, it is impossible to do the same for the approach based on the one-step ahead prediction MPC control [14].

iv) The system has been identified based on the fuzzy IF-Then approach, and hence, the designer can employ the experience of the operator for better functioning of the system and suitable manipulation of the control inputs in certain areas of the state space. However, for the system identified by relational matrix model of the strategy presented in the previous contributions [14] it was not possible to do translate the experience of the operator into the blocks of the relational matrix properly, and are not essentially favorable from the viewpoint of model transparency (in view of knowledge extraction and knowledge embedment) and interpretability.

v) The controller input can be computed incrementally, while the previous works needed a to solve a MIP problem, for every iteration from the scratch. It lowers very much the computational burden [29, 30].

vi) Although we have used the normal membership function, however, the present algorithm can be adapted easily for every definition of the membership functions, while the algorithm presented before in [14] is constructed based on the certain definition of the membership function. It worth to note that still normal membership function is more practical and widely used rather than the triangle membership function, used by AmirAskari and Menhaj.

vii) As a matter of fact, the smooth fuzzy components in its general form are very capable modeling structures, that could show robust properties to noise and switched changes. Hence, we tried to establish a standardized formulation for the general form that of smooth fuzzy structure which not only have a good modeling and control capability but also can be employed for all the generic forms of smooth fuzzy compositions defined in section 2.1.

Conclusions

Several interesting properties of smooth fuzzy compositions have been cited and proved in the literature [12-13], and robustness advantage of smooth fuzzy models has been reported in almost all the contributions in the field [14]. However, they cannot be employed for the practical cases and industrial systems until an easy and industrial implementable algorithm appears [15, 25, 26]. This chapter is a response to this requirement and we formulated a scheme for identification and long horizon MPC control of the smooth fuzzy models in its general form.

Hence, the overall achievement of the chapter has been twofold. One seeks to contribute to the state of smooth fuzzy controller design by proposing a general and systematic expert free design methodology, and to extend the operation range of smooth IF-THEN fuzzy models for time variant systems. We have presented an approach for the combined identification and MPC control of the systems through the smooth fuzzy composition. A gradient based optimization approach has been developed to convert the MPC cost function into an incremental controller design problem which results in a very fast and simple controller, in comparison to the other fuzzy MPC approaches that solve the problem through the hessian and gradient approximation.

Different simulations on a non- minimum phase unstable system have been provided to demonstrate the benefits and the limitations of the schemes in the presence of disturbance and noises. Four fuzzy compositions for extracting fuzzy MPC controllers to track the assumed trajectory using the fuzzy model have been contrasted. According to the test results, we can say that the overall smooth fuzzy modeling- control scheme is very much suitable for the adaptive control of time changing and noisy systems. One can conclude from the theoretical studies and the simulations that the smooth fuzzy controller is the best choice to respond in time to the disturbances, with the lowest overshoot during the changing tracking reference.

Appendix 6.1

In order to drive error derivatives, we study the identification process in more details. To begin with, we write the gradient descent method formula and define the vectors as follows,

$$\frac{\partial J}{\partial \rho_{ld}} = \frac{\partial J}{\partial y}\frac{\partial y}{\partial y_i}\frac{\partial y_i}{\partial \acute{y}_{li}}\frac{\partial \acute{y}_{li}}{\partial \xi_{ld}}\frac{\partial \xi_{ld}}{\partial \beta_{ld}}\frac{\partial \beta_{ld}}{\partial \rho_{ld}} \tag{6.32}$$

But to complete the formulation we need to take partial derivative of each variable separately.

1- We define the fuzzy variables $\{\acute{\xi}_1, \acute{\xi}_2, \cdots, \acute{\xi}_r\}$ at every time instant as,
$\acute{\xi}_l = [\xi_{l1}, \xi_{l2} \cdots, \xi_{l,m+p}] = [\beta_{l1}(\xi_1), \beta_{l2}(\xi_2), \cdots, \beta_{l,m+p}(\xi_{m+p})], l = 1 \cdots, r$

and $\acute{\xi} \;\; = [\acute{\xi}_l]_{l=1}^{r}$, where $\beta(\cdot)$, as stated above is value of the membership function for the fuzzy set. In general, this function can be written as,

$\beta_{ld}(\cdot) = exp\left(\frac{-1}{2}\left(\frac{\xi_{ld}-c_{ld}}{\delta_{ld}}\right)^2\right).$

(6.33)

Therefore, for making up the gradient descent method formula, $\frac{\partial \xi_{ld}}{\partial \rho_{ld}}$ can be written as,

$$\frac{\partial \xi_{ld}}{\partial c_{ld}} = exp\left(\frac{-1}{2}\left(\frac{\xi_{ld}-c_{ld}}{\delta_{ld}}\right)^2\right)\left(\frac{\xi_{ld}-c_{ld}}{\delta_{ij}^2}\right) \tag{6.34}$$

$$\frac{\partial \xi_{ld}}{\partial \delta_{ld}} = exp\left(\frac{-1}{2}\left(\frac{\xi_{ld}-c_{ld}}{\delta_{ld}}\right)^2\right)\left(\frac{(\xi_{ld}-c_{ld})^2}{\delta_{ld}}\right) \tag{6.35}$$

2- Based on the compositional rule inference, we can say that estimation of the output, according to our notation is,

$$\acute{y}_{li} \quad = s - norm\left(t - norm\left(\acute{\xi}_l, R_l\big(\acute{\xi}, y_i\big)\right)\right)$$

for all $l = 1, \cdots, r$.Let's abbreviate $S: s - norm$ and $T: t - norm$ in the following.

To facilitate the explanation of the procedure of taking the derivation of $\frac{\partial \acute{y}_{li}}{\partial \xi_{ld}}$, we assume a simple system and put $\acute{\xi}_l = [\xi_{l1}, \xi_{l2}]$ and $c = R\big(\acute{\xi}, y_i\big)$. Then, based on the properties of t-norms, we have,

$$\acute{y}_{li} \quad = S\left(T(T(\xi_{l1}, \xi_{l2}), c)\right) = S\big(T(\xi_{l1}, c), T(\xi_{l2}, c)\big)$$

We define: $\Lambda_1 = T(\xi_{l1}, c)$ and $\Lambda_2 = T(\xi_{l2}, c)$, then,

$$\acute{y}_{li} \quad = S(\Lambda_1, \Lambda_2) \tag{6.36}$$

$$\frac{\partial \acute{y}_{li}}{\partial \xi_{l1}} = \frac{\partial S}{\partial \Lambda_d}\frac{\partial \Lambda_d}{\partial \xi_{l1}} = \acute{S}^1\acute{T}^1, d = 1,2. \tag{6.37}$$

If there exists more state variables in the augmented state vector, $\acute{\xi}_l = \left[\xi_{l1}, \xi_{l2} \cdots, \xi_{l,m+p}\right]$we can continue in the same manner and write as,

$$\frac{\partial \acute{y}_{li}}{\partial \xi_{ld}} = \acute{S}^{m+p-1}\acute{T}^{m+p-1} \cdots \acute{S}^1\acute{T}^1. \tag{6.38}$$

Hence, to derive the gradient descent method formulation, the general formula for the error derivation will be,

$$\frac{\partial J}{\partial c_{ld}} = \frac{\partial J}{\partial y}\frac{\partial y}{\partial y_i}\frac{\partial y_i}{\partial \xi_{ld}}\frac{\partial \xi_{ld}}{\partial \beta_{ld}}\frac{\partial \beta_{ld}}{\partial c_{ld}} \tag{6.39}$$

$$= e(k) \cdot \left(\frac{\theta_{li} - y_i}{\sum_{i=1}^{r} \beta_{li}}\right) \cdot \left(\acute{S}^{m+p-1}\acute{T}^{m+p-1} \cdots \acute{S}^1\acute{T}^1\right) \cdot exp\left(\frac{-1}{2}\left(\frac{\xi_{ld} - c_{ld}}{\delta_{ld}}\right)^2\right)\left(\frac{\xi_{ld} - c_{ld}}{\delta_{ld}^2}\right)$$

$$\frac{\partial J}{\partial \delta_{ld}} = \frac{\partial J}{\partial y}\frac{\partial y}{\partial y_i}\frac{\partial y_i}{\partial \xi_{ld}}\frac{\partial \xi_{ld}}{\partial \beta_{ld}}\frac{\partial \beta_{ld}}{\partial \delta_{ld}}$$

(6.40)

$$= e(k) \cdot \left(\frac{\theta_{li} - y_i}{\sum_{i=1}^{r} \beta_{li}}\right) \cdot \left(\acute{S}^{m+p-1}\acute{T}^{m+p-1} \cdots \acute{S}^{1}\acute{T}^{1}\right) \cdot exp\left(\frac{-1}{2}\left(\frac{\xi_{ld} - c_{ld}}{\delta_{ld}}\right)^2\right) \left(\frac{(\xi_{ld} - c_{ld})^2}{\delta_{ld}^3}\right)$$

$$\frac{\partial J}{\partial \theta l_i} = \frac{\partial J}{\partial y}\frac{\partial y}{\partial y_i}\frac{\partial y_i}{\partial \theta_{li}}$$

(6.41)

$$= e(k) \cdot \left(\frac{\theta_{li} - \beta_{li}}{\sum_{i=1}^{r} \beta_{li}}\right)$$

Begin at the beginning, the King said, very gravely, and go on till you come to the end: then stop.'

Conclusions

7.1. Summary of The Thesis

The complexity of the industrial processes always demands for more sophisticated modeling and control systems to uphold the quality of products, make the process safer and more stable, employing the available hardware and experience and knowledge of the operator. Artificial learning methods and fuzzy logic based algorithms for a long time have been the most attractive responses to such challenges, and despite all the achievements and popularity, there are still open questions in the field.

One of the open questions from the design point of view always has been that which membership function (either Gaussian or Trapezoid or other else) would work better regarding the simplicity of calculations and the continuity of the models and smoothness. Parallel to that, which other parameter of the fuzzy model can be manipulated properly to reduce the number of trial and error and the required time and effort for a suitable controller design.

In this thesis, we demonstrated that upon the application of the smooth compositions, we will come to a united scheme of the continuity and smoothness, no matter what the membership function has been selected. Besides, we have tried to show how the smooth fuzzy compositions can be incorporated effectively into the nonlinear system modelling, identification and model based predicative control schemes and described the features on the speed of convergence, optimality and robustness the system will gain alongside. Although the initial works have been carried out before, we have emphasized on the theoretical study of the smooth fuzzy models and have tried to transfer a clear interpretation of such smooth models.

In later chapters, we have endeavored to determine the optimal model structure and controller design for the long horizon in model predictive control strategy. Several simulations have demonstrated that the proposed approaches bring remarkable attributes to the classical fuzzy modelling and control schemes, while offering better performance in the presence of noise and uncertainties to the highly nonlinear processes.

The accuracy of the modelling and control methods when be examined using simulations of the well-known benchmarks and the theoretical studies on the convergence and stability of the algorithms have been provided in the body of the thesis. While the numerical approaches have been examined for the algorithms, the emphasis has been on theoretical studies as well to prove the efficiency of the derived algorithms, which enjoy from the smoothness properties of the model at the optimization phase. Hence, all of the developed algorithms could outperform in modeling and control based on the classical fuzzy structures, especially in the presence of disturbance and noise. The results of this work, have been obtained without appealing to the nonlinear NP hard optimization algorithms, while, they have been a core for the application of smooth fuzzy compositions to the relational fuzzy models in the computation of the matrix operations, in the previous works of the earlier researchers.

The thesis concludes that the smooth fuzzy model for system identification and control brings about the following features:

- Can be used with different membership functions, and again brings the same level of smoothness and continuity to the model and to the derivative of model as the Gaussian

- In combination to the predictive control strategy facilitates the consideration of long control horizon whereby could reach better robustness and speed of convergence
- Smooth fuzzy IF-THEN model building through the least squares algorithm will show higher level of convergence in the optimization step and higher disturbance rejection capacity
- Allows dead time compensation in the MPC applications

7.2 Future Directions

We believe that this study represents the initial steps in a direction that appears to be promising in the smooth fuzzy modelling of the complex systems. Even though we endeavored to make a throughout study of the topic from different angels of optimality, stability and adaptiveness during the course of PhD works, it seems inevitable that some goal could not be achieved, given the limited period of time dedicate for termination of this dissertation.

Since, one feature of smooth fuzzy models is the higher speed of convergence, the future works can focus on the development of a detailed error mapping of the smooth fuzzy models for characterization of high speed stages used in the noisy environments for precise measurement and manipulation. When we employ the smooth compositions in the fuzzy models, derivative of the model and error mapping can be obtained analytically.

In fact, the success in robust modeling will empower to predict the experimental results accurately in the face of environmental conditions and parametric variations. Hence, we employed the smooth fuzzy compositions for the model building and self-learning of the models in the practical example of CSTR. Therefore, we suggest the proposers to give priority to the experimental verification of the benefits of the developed algorithm and work on it to meet the industrial needs and take measures for the transfer of it into the industry. Other works can focus on the applications of different control theories to the smooth models to improve the calibration accuracy of systems and decrease the number of interactions between the systems/tools/equipment and changes in the measurement configurations during the manipulation, validation and calibration phases.

Also, it will be possible to deal with the theoretical works and focus on the system constraints on the manipulated variables by handling them through proper application of the penalty functions. Moreover, it would be possible to consider the multi-objective optimization criteria for the nonlinear processes [29]. We believe that the future works also can dedicate to study the analytical robustness conditions in the controller design phase.

In our analysis and transformation, to run the approximation error and its derivative tend to zero we need to increase the number of partitions in the dense grid as well as the fuzzy rules. It means that in the practical applications, we will have growing numbers of fuzzy rules to make use of the smooth approximation properties. Therefore, there is a trade-off between the accuracy of fuzzy model and the modelling complexity. Hence, it is required to think about a method for finding the minimal number of fuzzy rules for a given accuracy of the fuzzy model in the future researches. One suggestion will be to discard the rules which have weak contribution to the output. The interested reader is referred to [23] for such solutions. Also, the application to the multi-agent systems and cooperative control can be followed as the schemes of [35, 36], as the other future paths.

References

[1] Mamdani, E., Assilan, S.: An experiment in linguistic synthesis with a fuzzy logic controller. Int. J. Man-Machine Studies 20(2), pp. 1–13, 1975.

[2] Zadeh L.,, Fuzzy Sets, Information and Control 8, pp. 338-353, 1965.

[3] Takagi, T., and Sugeno, M.: Fuzzy identification of systems and its application to modeling and control. IEEE Trans. Syst. Man Cybern. 15(1), pp. 116–132 1985.

[4] Zadeh LA, Fuzzy logic, Computer 21 (4), pp. 83-93, 1988.

[5] Bobal V., J. Bohm, J. Fessl and J. Machacek. Digital Self Tuning Controllers, Springer. 2005.

[6] Kruse R., C. Borgelt, F. Klawonn, C. Moewes, M. Steinbrecher, and P. Held, Computational Intelligence: A Methodological Introduction, Springer, 2013.

[7] Mendes N., and P. Neto, Indirect adaptive fuzzy control for industrial robots: A solution for contact applications, Expert Systems with Applications 42 (22), pp. 8929-8935, 2015.

[8] Jantzen J., Foundations of fuzzy control. Wiley, 2007.

[9] L. Giupponi, A. I. Perez-Neira, Fuzzy-based spectrum handoff in cognitive radio networks, 3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2008), DOI: 10.1109/CROWNCOM.2008.4562535

[10] Carlsson C., M. Fedrizzi, R. Fuller, Fuzzy logic in management, Springer Science+Business Media New York, DOI: 10.1007/978-1-4419-8977-2

[11] Nguyen H. T., N. R. Prasad, C. L. Walker, E. A. Walker, A first course in fuzzy and neural control, Chapman and Hall/CRC, ISBN 9781584882442 - CAT# C2441, 2002.

[12] Ashtiani A. A. and M. B. Menhaj, Some new smooth fuzzy relational compositions, Journal of Mathematics and Computer Science, 2 (4), pp. 717-722, 2011.

[13] Ashtiani A. A., and M. B. Menhaj, Introducing the fuzzy relational hybrid model as a building block for intelligent modeling of hybrid dynamical systems; IEEE Trans. on Fuzzy Systems, pp. 1971-1983, 2015.

[14] Askari M. A. and M. B. Menhaj, Fuzzy model predictive control based on modified fuzzy relational model, 13th Iranian Conference on Fuzzy Systems(IFSC), DOI: 10.1109/IFSC.2013.6675679, 2013.

[15] Sadjadi E., J. G. Herrero, J. M. Molina, and Z. Hatami On approximation properties of the smooth fuzzy models. International Journal of Fuzzy Systems, 20 (8), pp. 2657–2667, Springer, 2018.

[16] Wu D., and J. M. Mendel, On the continuity of type-1 and interval type-2 fuzzy logic systems, IEEE Transactions on Fuzzy Systems, 19 (1), pp. 179-192, 2011.

[17] Wang L. X. and J. M. Mendel, Fuzzy basis functions, universal approximation, and orthogonal least-squares learning, IEEE Trans. on Neural Networks, 3, pp. 807–813, 1992.

[18] Castro, J. L., Fuzzy logic controllers are universal approximators, IEEE Trans. on Systems, Man, and Cybernetics, 25(4), pp. 629–635, 1995.

[19] Kreinovich V., H. T. Nguyen, and Y. Yam, "Fuzzy systems are universal approximators for a smooth function and its derivatives," International Journal of Intelligent Systems, 15 (6), pp. 565–574, 1999.

[20] Sadjadi E., and A. H Borzabadi, Optimal control of hybrid systems by logic-based benders decomposition, IFAC Proceedings 42 (17), pp. 104-107, Available online 23 April 2016
[21] Borzabadi A., E. Sadjadi, and B. Moshiri, A Numerical scheme for approximate optimal control of nonlinear hybrid systems, International Journal of Innovative Computing, Information and Control, 6 (6), pp. 2715-2724, 2010.

[22] Branicky M. S., Introduction to hybrid systems. In Handbook of Networked and Embedded Control Systems, pp. 91–116, 2005.

[23] J. García, M. A. Patricio, A. Berlanga, J. M. Molina, Fuzzy region assignment for visual tracking, Journal Soft Computing - A Fusion of Foundations, Methodologies and Applications - Special Issue on Evolutionary Optimization and Learning archive, 15 (9), pp. 1845-1864, 2011.
[24] Venkat A. N., P. Vijaysai, R. D. Gudi, Identification of complex nonlinear processes based on fuzzy decomposition of the steady state space, Journal of Process Control, 13 (6), pp. 473-488, 2003.

[25] S. Mollov, R. Babuska, J. Abonyi, H. B. Verbruggen. Effective optimization for fuzzy model predictive control, IEEE Transactions on Fuzzy System s, 2004

[26] Huang Y.L., H.H. Lou, J.P. Gong, T.F. Edgar, Fuzzy model predictive control, IEEE Transactions on Fuzzy Systems, 8(6), pp. 665 - 678, 2000, DOI: 10.1109/91.890326

[27] Aljamaan I., D. Westwick, M. Foley, Non-linear closed-loop identification of CSTR in the presence of a non-stationary disturbances, IFAC-Papers, 48 (28), pp. 1029-1034, 2015.

[28] Kandiah S., Fuzzy model based predictive control of chemical processes: PhD Thesis, University of Sheffield, Sheffield, 1996.

[29] Borzabadi A. H., M. Hasanabadi, E. N. Sadjadi, Approximate pareto optimal solutions of multi objective optimal control problems by evolutionary algorithms, Journal of Control and Optimization in Applied Mathematics, 1 (1), pp 1-19, 2016.

[30] Borzabadi A. H., and E. Sadjadi, Epsilon Alternative solutions of the optimal control problems, Journal of Advanced Research in Scientific Computing, 1(1), pp 48-61, 2009.

[31] Perez O., J. Garcia, A. Berlanga, and J. M. Molina, Evolving parameters of surveillance video systems for non-overfitted learning, in Proceedings of the 7th European Workshop on Evolutionary Computation in Image Analysis and Signal Processing (EvoIASP '05), pp. 386–395, Lausanne, Switzerland, 2005.

[32] Herrero J. G., J. M. Molina, J. A. Besada and J. I. Portillo, A multitarget tracking video system based on fuzzy and neuro-fuzzy techniques, EURASIP Journal on Applied Signal Processing, special issue on Advances in Intelligent Vision Systems: Methods and Applications, 14, pp. 2341–2358, 2005.
[33] Camacho, F.E.; Bordons, A.C., Model predictive control in the process industry, Springer Verlag London Limited, 1995.

[34] Oliveira J. V., and Lemos, J. M., Long-Term predictive adaptive relational control, Fuzzy Sets and Systems, 70, pp 337–357, 1995.

[35] Molina J.M., J. García, F.J. Jiménez, J.R. Casar., Fuzzy reasoning in a multi agent system of surveillance sensors to manage cooperatively the sensor-to-task assignment problem, Applied Artificial Intelligence 18 (8), pp. 673–711, 2004.

[36] Molina J.M., J. Garcia, F.J. Jimenez, J.R. Casar, Cooperative management in a net of intelligent surveillance agent-sensors, International Journal of Intelligent Systems, 18 (3), pp. 279–307, 2003.