Fuzzy identification of Discrete Time Nonlinear Stochastic Systems

Ginalber L. O. Serra
Federal Institute of Education, Science and Technology (IFMA)
Brasil

1. Introduction

System identification is the task of developing or improving a mathematical description of dynamic systems from experimental data (Ljung (1999); Söderström & Stoica (1989)). Depending on the level of a priori insight about the system, this task can be approached in three different ways: *white box modeling*, *black box modeling* and *gray box modeling*. These models can be used for simulation, prediction, fault detection, design of controllers (*model based control*), and so forth. Nonlinear system identification (Aguirre *et al.* (2005); Serra & Bottura (2005); Sjöberg *et al.* (1995); **?**) is becomming an important tool which can be used to improve control performance and achieve robust behavior (Narendra & Parthasarathy (1990); Serra & Bottura (2006a)). Most processes in industry are characterized by nonlinear and time-varying behavior and are not amenable to conventional modeling approaches due to the lack of precise, formal knowledge about it, its strongly nonlinear behavior and high degree of uncertainty. Methods based on fuzzy models are gradually becoming established not only in academic view point but also because they have been recognized as powerful tools in industrial applications, faciliting the effective development of models by combining information from different sources, such as empirical models, heuristics and data (Hellendoorn & Driankov (1997)). In fuzzy models, the relation between variables are based on if-then rules such as IF $<$ *antecedent* $>$ THEN $<$ *consequent* $>$, where antecedent evaluate the model inputs and consequent provide the value of the model output. Takagi and Sugeno, in 1985, developed a new approach in which the key idea was partitioning the input space into fuzzy areas and approximating each area by a linear or a nonlinear model (Takagi & Sugeno (1985)). This structure, so called Takagi-Sugeno (TS) fuzzy model, can be used to approximate a highly nonlinear function of simple structure using a small number of rules. Identification of TS fuzzy model using experimental data is divided into two steps: structure identification and parameter estimation. The former consists of antecedent structure identification and consequent structure identification. The latter consists of antecedent and consequent parameter estimation where the consequent parameters are the coefficients of the linear expressions in the consequent of a fuzzy rule. To be applicable to real world problems, the parameter estimation must be highly efficient. Input and output measurements may be contaminated by noise. For low levels of noise the least squares (LS) method, for example, may produce excellent estimates of the consequent parameters. However, with larger levels of noise, some modifications in this method are required to overcome this inconsistency. Generalized least squares (GLS) method, extended least squares (ELS) method, prediction error (PE) method, are examples of such modifications. A problem with the use of these methods, in a fuzzy modeling context, is that the inclusion of the prediction error past values in the regression vector, which defines the input linguistic variables, increases the complexity of the fuzzy model structure and are inevitably dependent upon the accuracy of the noise model. To obtain consistent parameter estimates in a noisy environment without modeling the noise, the instrumental variable (IV) method can be used. It is known that by choosing proper instrumental variables, it provides a way to obtain consistent estimates with certain optimal properties (Serra & Bottura (2004; 2006b); Söderström & Stoica (1983)). This paper proposes an approach to nonlinear discrete time systems identification based on instrumental variable method and TS fuzzy model. In the proposed approach, which is an extension of the standard linear IV method (Söderström & Stoica (1983)), the chosen instrumental variables, statistically uncorrelated with the noise, are mapped to fuzzy sets, partitioning the input space in subregions to define valid and unbiased estimates of the consequent parameters for the TS fuzzy model in a noisy environment. From this theoretical background, the *fuzzy instrumental variable* (FIV) concept is proposed, and the main statistical characteristics of the FIV algorithm such as consistency and unbias are derived. Simulation results show that the proposed algorithm is relatively insensitive to the noise on the measured input-output data.

This paper is organized as follow: In Section 2, a brief review of the TS fuzzy model formulation is given. In Section 3, the fuzzy NARX structure is introduced. It is used to formulate the proposed approach. In Section 4, the TS fuzzy model consequent parameters estimate problem in a noisy environment is studied. From this analysis, three Lemmas and one Theorem are proposed to show the consistency and unbias of the parameters estimates in a noisy environment with the proposed approach. The fuzzy instrumental variable concept is also proposed and considerations about how the FIV should be chosen are given. In Section 5, off-line and on-line schemes of the fuzzy instrumental variable algorithm are derived. Simulation results showing the efficiency of the FIV approach in a noisy environment are given in Section 6. Finally, the closing remarks are given in Section 7.

2. Takagi-Sugeno Fuzzy Model

The TS fuzzy inference system is composed by a set of IF-THEN rules which partitions the input space, so-called *universe of discourse*, into fuzzy regions described by the rule antecedents in which consequent functions are valid. The consequent of each rule i is a functional expression $y_i = f_i(x)$ (King (1999); Papadakis & Theocaris (2002)). The i-th TS fuzzy rule has the following form:

$$R^{i|i=1,2,\ldots,l}: \text{ IF } x_1 \text{ is } F_1^i \text{ AND } \cdots \text{ AND } x_q \text{ is } F_q^i \text{ THEN } y_i = f_i(\mathbf{x}) \tag{1}$$

where l is the number of rules. The vector $\mathbf{x} \in \Re^q$ contains the antecedent linguistic variables, which has its own universe of discourse partitioned into fuzzy regions by the fuzzy sets representing the linguistic terms. The variable x_j belongs to a fuzzy set F_j^i with a truth value given by a membership function $\mu_{F_j}^i : \Re \rightarrow [0,1]$. The truth value h_i for the complete rule i is computed using the aggregation operator, or t-norm, AND, denoted by $\otimes : [0,1] \times [0,1] \rightarrow [0,1]$,

$$h_i(x) = \mu_1^i(x_1) \otimes \mu_2^i(x_2) \otimes \ldots \mu_q^i(x_q) \tag{2}$$

Among the different t-norms available, in this work the algebraic product will be used, and

$$h_i(\mathbf{x}) = \prod_{j=1}^{q} \mu_j^i(x_j) \tag{3}$$

The degree of activation for rule i is then normalized as

$$\gamma_i(\mathbf{x}) = \frac{h_i(\mathbf{x})}{\sum_{r=1}^{l} h_r(\mathbf{x})} \tag{4}$$

This normalization implies that

$$\sum_{i=1}^{l} \gamma_i(\mathbf{x}) = 1 \tag{5}$$

The response of the TS fuzzy model is a weighted sum of the consequent functions, i.e., a convex combination of the local functions (models) f_i,

$$y = \sum_{i=1}^{l} \gamma_i(\mathbf{x}) f_i(\mathbf{x}) \tag{6}$$

which can be seen as a linear parameter varying (LPV) system. In this sense, a TS fuzzy model can be considered as a mapping from the antecedent (input) space to a convex region (politope) in the space of the local submodels defined by the consequent parameters, as shown in Fig. 1 (Bergsten (2001)). This property simplifies the analysis of TS fuzzy models in a robust

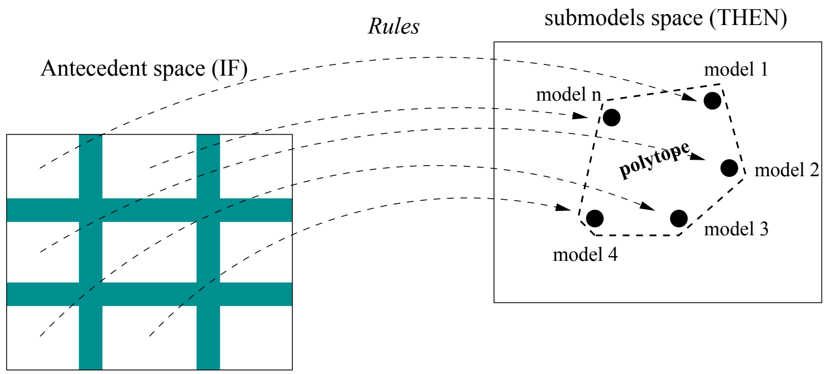


Fig. 1. Mapping to local submodels space.

linear system framework for identification, controllers design with desired closed loop characteristics and stability analysis (Johansen *et al.* (2000); Kadmiry & Driankov (2004); Tanaka *et al.* (1998); Tong & Li (2002)).

3. Fuzzy Structure Model

The nonlinear input-output representation is often used for building TS fuzzy models from data, where the regression vector is represented by a finite number of past inputs and outputs of the system. In this work, the nonlinear autoregressive with exogenous input (NARX) structure model is used. This model is applied in most nonlinear identification methods such as neural networks, radial basis functions, cerebellar model articulation controller (CMAC), and also fuzzy logic (Brown & Harris (1994)). The NARX model establishes a relation between the collection of past scalar input-output data and the predicted output

$$y(k+1) = F[y(k), \ldots, y(k-n_y+1), u(k), \ldots, u(k-n_u+1)] \tag{7}$$

where k denotes discrete time samples, n_y and n_u are integers related to the system's order. In terms of rules, the model is given by

$$\begin{aligned} R^i:\ \text{IF } y(k) \text{ is } F_1^i \text{ AND } \cdots \text{ AND } y(k-n_y+1) \text{ is } F_{n_y}^i \\ \text{AND } u(k) \text{ is } G_1^i \text{ AND } \cdots \text{ AND } u(k-n_u+1) \text{ is } G_{n_u}^i \\ \text{THEN } \hat{y}_i(k+1) = \sum_{j=1}^{n_y} a_{i,j} y(k-j+1) + \sum_{j=1}^{n_u} b_{i,j} u(k-j+1) + c_i \end{aligned} \tag{8}$$

where $a_{i,j}$, $b_{i,j}$ and c_i are the consequent parameters to be determined. The inference formula of the TS fuzzy model is a straightforward extension of (6) and is given by

$$y(k+1) = \frac{\sum_{i=1}^{l} h_i(\mathbf{x}) \hat{y}_i(k+1)}{\sum_{i=1}^{l} h_i(\mathbf{x})} \tag{9}$$

or

$$y(k+1) = \sum_{i=1}^{l} \gamma_i(\mathbf{x}) \hat{y}_i(k+1) \tag{10}$$

with

$$\mathbf{x} = [y(k), \ldots, y(k-n_y+1), u(k), \ldots, u(k-n_u+1)] \tag{11}$$

and $h_i(\mathbf{x})$ is given as (3). This NARX model represents multiple input and single output (MISO) systems directly and multiple input and multiple output (MIMO) systems in a decomposed form as a set of coupled MISO models.

4. Consequent Parameters Estimate

The inference formula of the TS fuzzy model in (10) can be expressed as

$$\begin{aligned} y(k+1) = \gamma_1(\mathbf{x}_k)[a_{1,1}y(k) + \ldots + a_{1,ny}y(k-n_y+1) \\ +b_{1,1}u(k) + \ldots + b_{1,nu}u(k-n_u+1) + c_1] + \gamma_2(\mathbf{x}_k)[a_{2,1}y(k) \\ + \ldots + a_{2,ny}y(k-n_y+1) + b_{2,1}u(k) + \ldots + b_{2,nu}u(k-n_u+1) \\ +c_2] + \ldots + \gamma_l(\mathbf{x}_k)[a_{l,1}y(k) + \ \ldots \ + a_{l,ny}y(k-n_y \\ +1) + b_{l,1}u(k) + \ldots + b_{l,nu}u(k-n_u+1) + c_l] \end{aligned} \tag{12}$$

which is linear in the consequent parameters: $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$. For a set of N input-output data pairs $\{(\mathbf{x}_k, y_k)|i = 1, 2, \ldots, N\}$ available, the following vetorial form is obtained

$$\mathbf{Y} = [\psi_1\mathbf{X}, \psi_2\mathbf{X}, \ldots, \psi_l\mathbf{X}]\theta + \Xi \tag{13}$$

where $\psi_i = diag(\gamma_i(\mathbf{x}_k)) \in \Re^{N\times N}$, $\mathbf{X} = [\mathbf{y}_k, \ldots, \mathbf{y}_{k-ny+1}, \mathbf{u}_k, \ldots, \mathbf{u}_{k-nu+1}, \mathbf{1}] \in \Re^{N\times(n_y+n_u+1)}$, $\mathbf{Y} \in \Re^{N\times 1}$, $\Xi \in \Re^{N\times 1}$ and $\theta \in \Re^{l(n_y+n_u+1)\times 1}$ are the normalized membership degree matrix of (4), the data matrix, the output vector, the approximation error vector and the estimated parameters vector, respectively. If the unknown parameters associated variables are *exactly known* quantities, then the least squares method can be used efficiently. However, in practice, and in the present context, the elements of $\mathbf{X}$ are no exactly known quantities so that its value can be expressed as

$$y_k = \chi_k^T \theta + \eta_k \tag{14}$$

where, at the k-th sampling instant, $\chi_k^T = [\gamma_k^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]$ is the vector of the data with error in variables, $\mathbf{x}_k = [y_{k-1}, \ldots, y_{k-n_y}, u_{k-1}, \ldots, u_{k-n_u}, 1]^T$ is the vector of the data with exactly known quantities, e.g., free noise input-output data, ξ_k is a vector of noise associated with the observation of $\mathbf{x}_k$, and η_k is a disturbance noise.
The normal equations are formulated as

$$[\sum_{j=1}^{k} \chi_j \chi_j^T]\hat{\theta}_k = \sum_{j=1}^{k} \chi_j y_j \tag{15}$$

and multiplying by $\frac{1}{k}$ gives

$$\{\frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j + \xi_j), \ldots, \gamma_j^l(\mathbf{x}_j + \xi_j)][\gamma_j^1(\mathbf{x}_j + \xi_j), \ldots,$$

$$\gamma_j^l(\mathbf{x}_j + \xi_j)]^T\}\hat{\theta}_k = \frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j + \xi_j), \ldots, \gamma_j^l(\mathbf{x}_j + \xi_j)]y_j$$

Noting that $y_j = \chi_j^T \theta + \eta_j$,

$$\{\frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j + \xi_j), \ldots, \gamma_j^l(\mathbf{x}_j + \xi_j)][\gamma_j^1(\mathbf{x}_j + \xi_j), \ldots,$$

$$\gamma_j^l(\mathbf{x}_j + \xi_j)]^T\}\hat{\theta}_k = \frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j + \xi_j), \ldots, \gamma_j^l(\mathbf{x}_j + \xi_j)]$$

$$[\gamma_j^1(\mathbf{x}_j + \xi_j), \ldots, \gamma_j^l(\mathbf{x}_j + \xi_j)]^T\theta + \frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j + \xi_j), \ldots,$$

$$\gamma_j^l(\mathbf{x}_j + \xi_j)]\eta_j \tag{16}$$

and

$$\tilde{\theta}_k = \{\frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j + \xi_j), \ldots, \gamma_j^l(\mathbf{x}_j + \xi_j)][\gamma_j^1(\mathbf{x}_j + \xi_j), \ldots,$$

$$\gamma_j^l(\mathbf{x}_j + \xi_j)]^T\}^{-1}\frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j + \xi_j), \ldots, \gamma_j^l(\mathbf{x}_j + \xi_j)]\eta_j \tag{17}$$

where $\tilde{\theta}_k = \hat{\theta}_k - \theta$ is the parameter error. Taking the probability in the limit as $k \to \infty$,

$$\text{p.lim}\ \tilde{\theta}_k = \text{p.lim}\ \{\frac{1}{k}\mathbf{C}_k^{-1}\frac{1}{k}\mathbf{b}_k\} \tag{18}$$

with

$$\mathbf{C}_k = \sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)][\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]^T$$

$$\mathbf{b}_k = \sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]\eta_j$$

Applying Slutsky's theorem and assuming that the elements of $\frac{1}{k}\mathbf{C}_k$ and $\frac{1}{k}\mathbf{b}_k$ converge in probability, we have

$$\text{p.lim}\ \tilde{\theta}_k = \text{p.lim}\ \frac{1}{k}\mathbf{C}_k^{-1}\ \text{p.lim}\ \frac{1}{k}\mathbf{b}_k \tag{19}$$

Thus,

$$\text{p.lim}\ \frac{1}{k}\mathbf{C}_k = \text{p.lim}\ \frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]$$
$$[\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]^T$$

$$\text{p.lim}\ \frac{1}{k}\mathbf{C}_k = \text{p.lim}\ \frac{1}{k}\sum_{j=1}^{k}(\gamma_j^1)^2(\mathbf{x}_j+\xi_j)(\mathbf{x}_j+\xi_j)^T +$$
$$\ldots + \text{p.lim}\ \frac{1}{k}\sum_{j=1}^{k}(\gamma_j^l)^2(\mathbf{x}_j+\xi_j)(\mathbf{x}_j+\xi_j)^T$$

Assuming $\mathbf{x}_j$ and ξ_j statistically independent,

$$\text{p.lim}\ \frac{1}{k}\mathbf{C}_k = \text{p.lim}\ \frac{1}{k}\sum_{j=1}^{k}(\gamma_j^1)^2[\mathbf{x}_j\mathbf{x}_j^T + \xi_j\xi_j^T] + \ldots$$
$$+\text{p.lim}\ \frac{1}{k}\sum_{j=1}^{k}(\gamma_j^l)^2[\mathbf{x}_j\mathbf{x}_j^T + \xi_j\xi_j^T]$$

$$\text{p.lim}\ \frac{1}{k}\mathbf{C}_k = \text{p.lim}\ \frac{1}{k}\sum_{j=1}^{k}\mathbf{x}_j\mathbf{x}_j^T[(\gamma_j^1)^2+\ldots+(\gamma_j^l)^2]$$
$$+\text{p.lim}\ \frac{1}{k}\sum_{j=1}^{k}\xi_j\xi_j^T[(\gamma_j^1)^2+\ldots+(\gamma_j^l)^2] \tag{20}$$

with $\sum_{i=1}^{l}\gamma_j^i = 1$. Hence, the asymptotic analysis of the TS fuzzy model consequent parameters estimation is based in a weighted sum of the fuzzy covariance matrices of $\mathbf{x}$ and ξ. Similarly,

$$\text{p.lim}\ \frac{1}{k}\mathbf{b}_k = \text{p.lim}\ \frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]\eta_j$$

$$\text{p.lim}\,\frac{1}{k}\mathbf{b}_k = \text{p.lim}\,\frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1\xi_j\eta_j,\ldots,\gamma_j^l\xi_j\eta_j] \tag{21}$$

Substituting from (20) and (21) in (19), results

$$\text{p.lim}\,\tilde{\theta}_k = \{\text{p.lim}\,\frac{1}{k}\sum_{j=1}^{k}\mathbf{x}_j\mathbf{x}_j^T[(\gamma_j^1)^2+\ldots+(\gamma_j^l)^2]+$$
$$\text{p.lim}\,\frac{1}{k}\sum_{j=1}^{k}\xi_j\xi_j^T[(\gamma_j^1)^2+\ldots+(\gamma_j^l)^2]\}^{-1}\text{p.lim}\,\frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1\xi_j\eta_j,$$
$$\ldots,\gamma_j^l\xi_j\eta_j] \tag{22}$$

with $\sum_{i=1}^{l}\gamma_j^i = 1$. For the case of only one rule ($l = 1$), the analysis is simplified to the linear one, with $\gamma_j^i\,|_{j=1,\ldots,k}^{i=1} = 1$. Thus, this analysis, which is a contribution of this article, is an extension of the standard linear one, from which can result several studies for fuzzy filtering and modeling in a noisy environment, fuzzy signal enhancement in communication channel, and so forth. Provided that the input u_k continues to excite the process and, at the same time, the coefficients in the submodels from the consequent are not all zero, then the output y_k will exist for all k observation intervals. As a result, the fuzzy covariance matrix $\sum_{j=1}^{k}\mathbf{x}_j\mathbf{x}_j^T[(\gamma_j^1)^2 + \ldots + (\gamma_j^l)^2]$ will also be non-singular and its inverse will exist. Thus, the only way in which the asymptotic error can be zero is for $\xi_j\eta_j$ identically zero. But, in general, ξ_j and η_j are correlated, the asymptotic error will not be zero and the least squares estimates will be asymptotically biased to an extent determined by the relative ratio of noise to signal variances. In other words, least squares method is not appropriate to estimate the TS fuzzy model consequent parameters in a noisy environment because the estimates will be inconsistent and the bias error will remain no matter how much data can be used in the estimation.

4.1 Fuzzy instrumental variable (FIV)

To overcome this bias error and inconsistence problem, generating a vector of variables which are independent of the noise inputs and correlated with data vetor $\mathbf{x}_j$ from the system is required. If this is possible, then the choice of this vector becomes effective to remove the asymptotic bias from the consequent parameters estimates. The fuzzy least squares estimates is given by:

$$\{\frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)][\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,$$
$$\gamma_j^l(\mathbf{x}_j+\xi_j)]^T\}\hat{\theta}_k = \frac{1}{k}\sum_{j=1}^{k}[\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,$$
$$\gamma_j^l(\mathbf{x}_j+\xi_j)]\{[\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]^T\theta+\eta_j\}$$

Using a new fuzzy vector of variables of the form $[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]$, the last equation can be placed as

$$\{\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j][\gamma_j^1(\mathbf{x}_j+\xi_j), \ldots, \gamma_j^l(\mathbf{x}_j+\xi_j)]^T\}\hat{\theta}_k =$$
$$\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]\{[\gamma_j^1(\mathbf{x}_j+\xi_j), \ldots, \gamma_j^l(\mathbf{x}_j+\xi_j)]^T\theta +$$
$$\eta_j\} \tag{23}$$

where $\mathbf{z}_j$ is a vector with the order of $\mathbf{x}_j$, associated to the dynamic behavior of the system, and $\beta_j^i \mid_{j=1,\ldots,k}^{i=1,\ldots,l}$ is the normalized degree of activation, as in (4), associated to $\mathbf{z}_j$. For convergence analysis of the estimates, with the inclusion of this new fuzzy vector, the following is proposed:

Lemma 1 *Consider* $\mathbf{z}_j$ *a vector with the order of* $\mathbf{x}_j$*, associated to dynamic behavior of the system and independent of the noise input* ξ_j*; and* $\beta_j^i \mid_{j=1,\ldots,k}^{i=1,\ldots,l}$ *is the normalized degree of activation, a variable defined as in (4) associated to* $\mathbf{z}_j$*. Then, at the limit*

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]\xi_j^T = \mathbf{0} \tag{24}$$

Proof: Developing the left side of (24), results

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]\xi_j^T = \lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j\xi_j^T, \ldots,$$
$$\beta_j^l \mathbf{z}_j\xi_j^T]$$

As $\beta_j^i \mid_{j=1,\ldots,k}^{i=1,\ldots,l}$ is a scalar, and, by definition, the chosen variables are independent of the noise inputs, the inner product between $\mathbf{z}_j$ and ξ_j will be zero. Thus, taking the limit, results

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j\xi_j^T, \ldots, \beta_j^l \mathbf{z}_j\xi_j^T] = \mathbf{0}$$

□

Lemma 2 *Under the same conditions as Lemma 1 and* $\mathbf{z}_j$ *independent of the disturbance noise* η_j*, then, at the limit*

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]\eta_j = \mathbf{0} \tag{25}$$

Proof: Developing the left side of (25), results

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]\eta_j = \lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j\eta_j, \ldots,$$
$$\beta_j^l \mathbf{z}_j\eta_j]$$

Because the chosen variables are independent of the disturbance noise, the product between $\mathbf{z}_j$ and η_j will be zero in the limit. Hence,

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1\mathbf{z}_j\eta_j,\ldots,\beta_j^l\mathbf{z}_j\eta_j]=\mathbf{0}$$

□

Lemma 3 *Under the same conditions as Lemma 1, according to (23), at the limit*

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1\mathbf{z}_j,\ldots,\beta_j^l\mathbf{z}_j][\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]^T=\mathbf{C}_{\mathbf{zx}}\neq 0 \tag{26}$$

Proof: Developing the left side of (26), results

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1\mathbf{z}_j,\ldots,\beta_j^l\mathbf{z}_j][\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]^T=$$
$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1\gamma_j^1\mathbf{z}_j(\mathbf{x}_j+\xi_j)^T+\ldots+\beta_j^l\gamma_j^l\mathbf{z}_j(\mathbf{x}_j+\xi_j)^T]$$

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1\mathbf{z}_j,\ldots,\beta_j^l\mathbf{z}_j][\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]^T=$$
$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1\gamma_j^1(\mathbf{z}_j\mathbf{x}_j^T+\mathbf{z}_j\xi_j^T)+\ldots+\beta_j^l\gamma_j^l(\mathbf{z}_j\mathbf{x}_j^T+\mathbf{z}_j\xi_j^T)]$$

From the **Lemma 1**, this expression is simplifyied as

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1\mathbf{z}_j,\ldots,\beta_j^l\mathbf{z}_j][\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]^T=$$
$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1\gamma_j^1\mathbf{z}_j\mathbf{x}_j^T+\ldots+\beta_j^l\gamma_j^l\mathbf{z}_j\mathbf{x}_j^T]$$

Due to correlation between $\mathbf{z}_j$ and $\mathbf{x}_j$, this fuzzy covariance matrix has the following property:

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1\gamma_j^1\mathbf{z}_j\mathbf{x}_j^T+\ldots+\beta_j^l\gamma_j^l\mathbf{z}_j\mathbf{x}_j^T]\neq 0 \tag{27}$$

and

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1\gamma_j^1\mathbf{z}_j\mathbf{x}_j^T+\ldots+\beta_j^l\gamma_j^l\mathbf{z}_j\mathbf{x}_j^T]=\mathbf{C}_{\mathbf{zx}}\neq\mathbf{0}$$

□

Theorem 1 *Under suitable conditions outlined from Lemma 1 to 3, the estimation of the parameter vector θ for the model in (12) is strongly consistent, i.e, at the limit*

$$\text{p.lim}\ \tilde{\theta} = 0 \tag{28}$$

Proof: From the new fuzzy vector of variables of the form $[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]$, the fuzzy least square estimation can be modifyied as follow:

$$\{\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j][\gamma_j^1(\mathbf{x}_j+\xi_j), \ldots, \gamma_j^l(\mathbf{x}_j+\xi_j)]^T\}\hat{\theta}_k =$$

$$\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]\{[\gamma_j^1(\mathbf{x}_j+\xi_j), \ldots, \gamma_j^l(\mathbf{x}_j+\xi_j)]^T\theta+\eta_j\}$$

which can be expressed in the form

$$\{\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j][\gamma_j^1(\mathbf{x}_j+\xi_j), \ldots, \gamma_j^l(\mathbf{x}_j+\xi_j)]^T\}(\hat{\theta}_k-\theta) =$$

$$\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]\eta_j$$

and

$$\{\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j][\gamma_j^1(\mathbf{x}_j+\xi_j), \ldots, \gamma_j^l(\mathbf{x}_j+\xi_j)]^T\}\tilde{\theta} =$$

$$\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]\eta_j$$

Taking the probability in the limit as $k \to \infty$, and applying the Slutsky's theorem, we have

$$\text{p.lim}\ \tilde{\theta}_k = \{\text{p.lim}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j][\gamma_j^1(\mathbf{x}_j+\xi_j), \ldots,$$

$$\gamma_j^l(\mathbf{x}_j+\xi_j)]^T\}^{-1}\{\text{p.lim}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]\eta_j\}$$

According to **Lemma 1** and **Lemma 3**, results

$$\text{p.lim}\ \tilde{\theta}_k = \{\text{p.lim}\mathbf{C}_{\mathbf{zx}}\}^{-1}\{\text{p.lim}\frac{1}{k}\sum_{j=1}^{k}[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]\eta_j\}$$

where the fuzzy covariance matrix $\mathbf{C}_{\mathbf{zx}}$ is non-singular and, as a consequence, the inverse exist. From the **Lemma 2**, we have

$$\text{p.lim}\ \tilde{\theta}_k = \{\text{p.lim}\mathbf{C}_{\mathbf{zx}}\}^{-1}\mathbf{0}$$

Thus, the limit value of the parameter error, in probability, is

$$\text{p.lim}\,\tilde{\theta} = 0 \tag{29}$$

and the estimates are asymptotically unbiased, as required. □

As a consequence of this analysis, the definition of the vector $[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]$ as the *fuzzy instrumental variable vector* or simply the *fuzzy instrumental variable* (FIV) is proposed. Clearly, with the use of the FIV vector in the form suggested, becomes possible to eliminate the asymptotic bias while preserving the existence of a solution. However, the statistical efficiency of the solution is dependent on the degree of correlation between $[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]$ and $[\gamma_j^1 \mathbf{x}_j, \ldots, \gamma_j^l \mathbf{x}_j]$. In particular, the lowest variance estimates obtained from this approach occur only when $\mathbf{z}_j = \mathbf{x}_j$ and $\beta_j^i \mid_{j=1,\ldots,k}^{i=1,\ldots,l} = \gamma_j^i \mid_{j=1,\ldots,k}^{i=1,\ldots,l}$, i.e., when the $\mathbf{z}_j$ are equal to the dynamic system "free noise" variables, which are unavailable in practice. According to situation, several fuzzy instrumental variables can be chosen. An effective choice of FIV would be the one based on the delayed input sequence

$$\mathbf{z}_j = [u_{k-\tau}, \ldots, u_{k-\tau-n}, u_k, \ldots, u_{k-n}]^T$$

where τ is chosen so that the elements of the fuzzy covariance matrix $\mathbf{C_{zx}}$ are maximized. In this case, the input signal is considered persistently exciting, e.g., it continuously perturbs or excites the system. Another FIV would be the one based on the delayed input-output sequence

$$\mathbf{z}_j = [y_{k-1-dl}, \cdots, y_{k-n_y-dl}, u_{k-1-dl}, \cdots, u_{k-n_u-dl}]^T$$

where dl is the applied delay. Other FIV could be the one based in the input-output from a "fuzzy auxiliar model" with the same structure of the one used to identify the nonlinear dynamic system. Thus,

$$\mathbf{z}_j = [\hat{y}_{k-1}, \cdots, \hat{y}_{k-n_y}, u_{k-1}, \cdots, u_{k-n_u}]^T$$

where $\hat{y}_k$ is the output of the fuzzy auxiliar model, and u_k is the input of the dynamic system. The inference formula of this fuzzy auxiliar model is given by

$$\begin{gathered}
\hat{y}(k+1) = \beta_1(\mathbf{z}_k)[\alpha_{1,1}\hat{y}(k) + \ldots + \alpha_{1,ny}\hat{y}(k-n_y+1) + \\
\rho_{1,1}u(k) + \ldots + \rho_{1,nu}u(k-n_u+1) + \delta_1] + + \beta_2(\mathbf{z}_k)[\alpha_{2,1}\hat{y}(k) \\
+ \ldots + \alpha_{2,ny}\hat{y}(k-n_y+1) + \rho_{2,1}u(k) + \ldots + \rho_{2,nu}u(k- \\
n_u+1) + \delta_2] + \ldots + \beta_l(\mathbf{z}_k)[\alpha_{l,1}\hat{y}(k) + \ldots + \alpha_{l,ny}\hat{y}(k- \\
n_y+1) + \rho_{l,1}u(k) + \ldots + \rho_{l,nu}u(k-n_u+1) + \delta_l]
\end{gathered}$$

which is also linear in the consequent parameters: α, ρ and δ. The closer these parameters are to the actual, but unknown, system parameters ($\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$) as in (12), more correlated $\mathbf{z}_k$ and $\mathbf{x}_k$ will be, and the obtained FIV estimates closer to the optimum.

5. FIV Algorithm

The FIV approach is a simple and attractive technique because it does not require the noise modeling to yield consistent, asymptotically unbiased consequent parameters estimates.

5.1 Off-line scheme

The FIV normal equations are formulated as

$$\sum_{j=1}^{k}[\beta_j^1\mathbf{z}_j,\ldots,\beta_j^l\mathbf{z}_j][\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]^T\hat{\theta}_k-\sum_{j=1}^{k}[\beta_j^1\mathbf{z}_j,\ldots,\beta_j^l\mathbf{z}_j]y_j=0 \tag{30}$$

or, with $\zeta_j=[\beta_j^1\mathbf{z}_j,\ldots,\beta_j^l\mathbf{z}_j]$,

$$[\sum_{j=1}^{k}\zeta_j\chi_j^T]\hat{\theta}_k-\sum_{j=1}^{k}\zeta_j y_j=0 \tag{31}$$

so that the FIV estimate is obtained as

$$\hat{\theta}_k=\{\sum_{j=1}^{k}[\beta_j^1\mathbf{z}_j,\ldots,\beta_j^l\mathbf{z}_j][\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]^T\}^{-1}\sum_{j=1}^{k}[\beta_j^1\mathbf{z}_j,\ldots,\beta_j^l\mathbf{z}_j]y_j \tag{32}$$

and, in vectorial form, the interest problem may be placed as

$$\hat{\theta}=(\Gamma^T\Sigma)^{-1}\Gamma^T\mathbf{Y} \tag{33}$$

where $\Gamma^T\in\Re^{l(n_y+n_u+1)\times N}$ is the fuzzy extended instrumental variable matrix with rows given by ζ_j, $\Sigma\in\Re^{N\times l(n_y+n_u+1)}$ is the fuzzy extended data matrix with rows given by χ_j and $\mathbf{Y}\in\Re^{N\times 1}$ is the output vector and $\hat{\theta}\in\Re^{l(n_y+n_u+1)\times 1}$ is the parameters vector. The models can be obtained by the following two approaches:

- *Global approach* : In this approach all linear consequent parameters are estimated simultaneously, minimizing the criterion:

$$\hat{\theta}=\ \arg\min\ \|\ \Gamma^T\Sigma\theta-\Gamma^T\mathbf{Y}\ \|_2^2 \tag{34}$$

- *Local approach* : In this approach the consequent parameters are estimated for each rule i, and hence independently of each other, minimizing a set of weighted local criteria $(i=1,2,\ldots,l)$:

$$\hat{\theta}_i=\ \arg\min\ \|\ \mathbf{Z}^T\Psi_i\mathbf{X}\theta_i-\mathbf{Z}^T\Psi_i\mathbf{Y}\ \|_2^2 \tag{35}$$

  where $\mathbf{Z}^T$ has rows given by $\mathbf{z}_j$ and Ψ_i is the normalized membership degree diagonal matrix according to $\mathbf{z}_j$.

5.2 On-line scheme

An on line FIV scheme can be obtained by utilizing the recursive solution to the FIV equations and then updating the fuzzy auxiliar model continuously on the basis of these recursive consequent parameters estimates. The FIV estimate in (32) can take the form

$$\hat{\theta}_k=\mathbf{P}_k\mathbf{b}_k \tag{36}$$

where

$$\mathbf{P}_k=\sum_{j=1}^{k}[\beta_j^1\mathbf{z}_j,\ldots,\beta_j^l\mathbf{z}_j][\gamma_j^1(\mathbf{x}_j+\xi_j),\ldots,\gamma_j^l(\mathbf{x}_j+\xi_j)]^T\}^{-1}$$

and

$$\mathbf{b}_k = \sum_{j=1}^{k} [\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j] y_j$$

which can be expressed as

$$\mathbf{P}_k^{-1} = \mathbf{P}_{k-1}^{-1} + [\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k][\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \tag{37}$$

and

$$\mathbf{b}_k = \mathbf{b}_{k-1} + [\beta_k^1 \mathbf{z}_k, \ldots, \beta_k^l \mathbf{z}_k] y_k \tag{38}$$

respectively. Pre-multiplying (37) by $\mathbf{P}_k$ and post-multiplying by $\mathbf{P}_{k-1}$ gives

$$\mathbf{P}_{k-1} = \mathbf{P}_k + \mathbf{P}_k[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k][\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1} \tag{39}$$

then firstly post-multiplying (39) by the FIV vector $[\beta_j^1 \mathbf{z}_j, \ldots, \beta_j^l \mathbf{z}_j]$, and after that, post-multiplying by $\{1 + [\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1}[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k]\}^{-1}$ $[\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1}$, results

$$\begin{aligned}
\mathbf{P}_{k-1}[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k] = \mathbf{P}_k[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k] + \\
\mathbf{P}_k[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k][\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \\
\mathbf{P}_{k-1}[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k]\mathbf{P}_{k-1}[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k] = \\
\mathbf{P}_k[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k]\{1 + [\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \\
\mathbf{P}_{k-1}[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k]\}
\end{aligned} \tag{40}$$

Then, post-multiplying by $\{1 + [\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1}[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k]\}^{-1}$ $[\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1}$, we obtain

$$\begin{aligned}
\mathbf{P}_{k-1}[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k]\{1 + [\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \\
\mathbf{P}_{k-1}[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k]\}^{-1}[\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \\
\mathbf{P}_{k-1} = \mathbf{P}_k[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k][\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T \\
\mathbf{P}_{k-1}
\end{aligned} \tag{41}$$

Substituting (39) in (41), we have

$$\begin{aligned}
\mathbf{P}_k = \mathbf{P}_{k-1} - \mathbf{P}_{k-1}[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k]\{1 + [\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \\
\gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1}[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k]\}^{-1}[\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \\
\gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1}
\end{aligned} \tag{42}$$

Substituting (42) and (38) in (36), the recursive consequent parameters estimates will be:

$$\begin{aligned}
\hat{\theta}_k = \{\mathbf{P}_{k-1} - \mathbf{P}_{k-1}[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k]\{1 + [\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \\
\gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1}[\beta_k^1 \mathbf{z}_k, \ldots, \beta_j^l \mathbf{z}_k]\}^{-1}[\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \\
\gamma_k^l(\mathbf{x}_k + \xi_k)]^T \mathbf{P}_{k-1}\}\{\mathbf{b}_{k-1} + [\beta_k^1 \mathbf{z}_k, \ldots, \beta_k^l \mathbf{z}_k] y_k\}
\end{aligned}$$

so that finally,

$$\hat{\theta}_k = \hat{\theta}_{k-1} - \mathbf{K}_k\{[\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T\hat{\theta}_{k-1} - y_k\} \tag{43}$$

where

$$\mathbf{K}_k = \mathbf{P}_{k-1}[\beta_k^1\mathbf{z}_k, \ldots, \beta_k^l\mathbf{z}_k]\{1 + [\gamma_j^1(\mathbf{x}_k + \xi_k), \ldots, \gamma_k^l(\mathbf{x}_k + \xi_k)]^T\mathbf{P}_{k-1}[\beta_k^1\mathbf{z}_k, \ldots, \beta_j^l\mathbf{z}_k]\}^{-1} \tag{44}$$

The equations (42)-(44) compose the FIV recursive estimation formula, and are implemented to determine unbiased estimates for the TS fuzzy model consequent parameters in a noisy environment.

6. COMPUTATIONAL RESULTS

In the sequel, two examples will be presented to demonstrate the effectiveness and applicability of the proposed algorithm in a noisy environment. Practical application of this method can be seen in (**?**), where was performed the identification of an aluminium beam, a complex nonlinear time varying plant whose study provides a great background for active vibration control applications in mechanical structures of aircrafts and/or aerospace vehicles.

6.1 Polynomial function approximation

Consider a nonlinear function defined by

$$u_k = u_k^i + \nu_k \tag{45}$$

$$y_k^i = 1 - 2u_k + u_k^2 \tag{46}$$

$$y_k = y_k^i + c_k - 0.25c_{k-1} \tag{47}$$

In Fig. 2 are shown the true system ($u_k^i \in [0, 2]$,y_k^i) and the noisy (u_k,y_k) input-output observations with measurements corrupted by normal noise conditions of $\sigma_c = \sigma_\nu = 0.2$. The results for the TS fuzzy models obtained by applying the proposed FIV algorithm as well as the LS estimation to tune the consequent parameters are shown in Fig. 3. It can be seen, clearly, that the curves for the polynomial function and for the proposed FIV based identification almost cover each other. The fuzzy c-means clustering algorithm was used to criate the antecedent membership functions of the TS fuzzy models, which are shown in Fig. 4. The FIV was based on the filtered output from a "fuzzy auxiliar model" with the same structure of the TS fuzzy model used to identify the nonlinear function. The clusters centers of the membership functions for the LS and FIV estimations were $\mathbf{c} = [-0.0983, 0.2404, 0.6909, 1.1611]^T$ and $\mathbf{c} = [0.1022, 0.4075, 0.7830, 1.1906]^T$, respectively. The TS fuzzy models have the following structure:

$$R^i: \text{ IF } y_k \text{ is } F_i \text{ THEN } \hat{y}_k = a_0 + a_1u_k + a_2u_k^2$$

where $i = 1, 2, \ldots, 4$. For the FIV approach, the "fuzzy auxiliar model" has the following structure:

$$R^i: \text{ IF } y_{filt} \text{ is } F_i \text{ THEN } y_{filt} = a_0 + a_1u_k + a_2y_{filt}^2$$

where y_{filt} is the filtered output, based on the consequent parameters LS estimation, and used to criate the membership functions, as shown in Fig. 4, as well as the instrumental variable matrix. The resulting TS fuzzy models based on the LS estimation are:

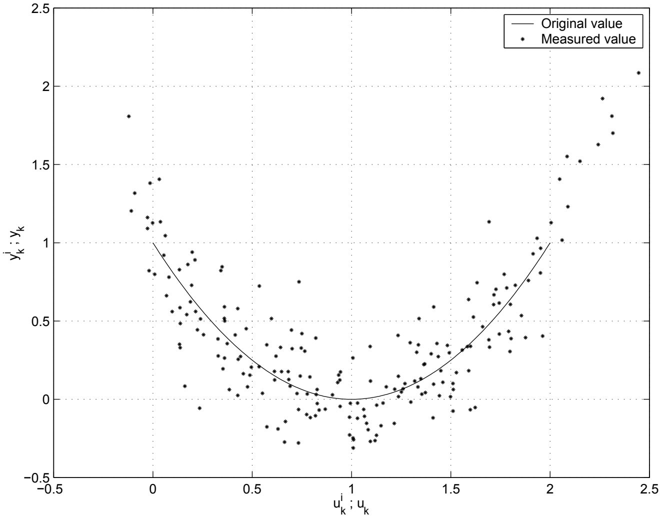


Fig. 2. Polynomial function with error in variables.

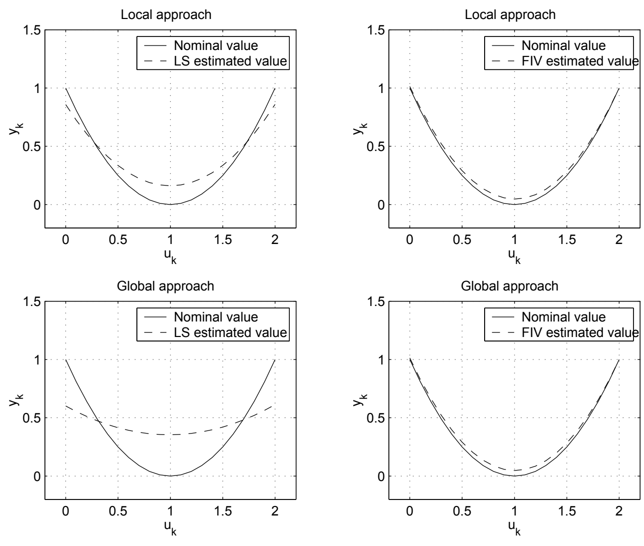


Fig. 3. Approximation of the polynomial function.

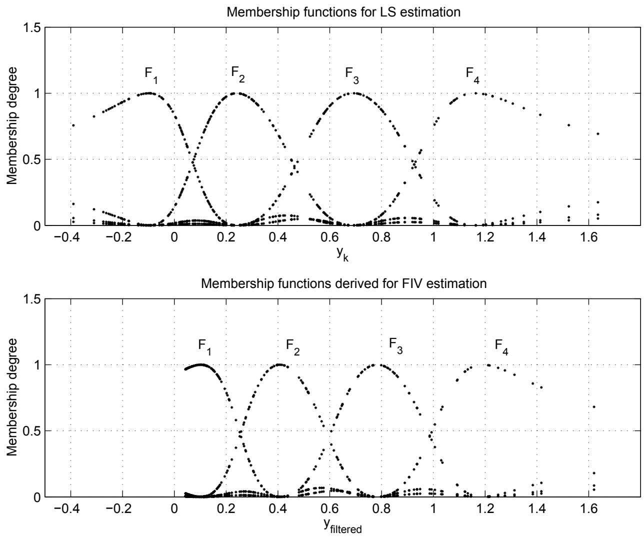


Fig. 4. Antecedent membership functions.

Local approach:

$$R^1: \text{ IF } y_k \text{ is } F_1 \text{ THEN } \hat{y}_k = 0.7074 - 1.7120u_k + 0.8717u_k^2$$
$$R^2: \text{ IF } y_k \text{ is } F_2 \text{ THEN } \hat{y}_k = 0.7466 - 1.2077u_k + 0.5872u_k^2$$
$$R^3: \text{ IF } y_k \text{ is } F_3 \text{ THEN } \hat{y}_k = 0.8938 - 1.1831u_k + 0.5935u_k^2$$
$$R^4: \text{ IF } y_k \text{ is } F_4 \text{ THEN } \hat{y}_k = 1.0853 - 1.4776u_k + 0.7397u_k^2$$

Global approach:

$$R^1: \text{ IF } y_k \text{ is } F_1 \text{ THEN } \hat{y}_k = 0.0621 - 0.4630u_k + 0.2272u_k^2$$
$$R^2: \text{ IF } y_k \text{ is } F_2 \text{ THEN } \hat{y}_k = 0.3729 - 0.3068u_k + 0.1534u_k^2$$
$$R^3: \text{ IF } y_k \text{ is } F_3 \text{ THEN } \hat{y}_k = 0.7769 - 0.3790u_k + 0.1891u_k^2$$
$$R^4: \text{ IF } y_k \text{ is } F_4 \text{ THEN } \hat{y}_k = 1.1933 - 0.8500u_k + 0.4410u_k^2$$

According to Fig. 3, the obtained TS fuzzy models based on LS estimation are very poor and they were not able to aproximate the original nonlinear function data. It shows the influency of noise on the regressors of the data matrix, as explained in section 4, making the consequent parameters estimation biased and inconsistent. On the other hand, the resulting TS fuzzy models based on the FIV estimation are of the form:
Local approach:

$$R^1: \text{ IF } y_k \text{ is } F_1 \text{ THEN } \hat{y}_k = 1.0130 - 1.9302u_k + 0.9614u_k^2$$
$$R^2: \text{ IF } y_k \text{ is } F_2 \text{ THEN } \hat{y}_k = 1.0142 - 1.9308u_k + 0.9618u_k^2$$
$$R^3: \text{ IF } y_k \text{ is } F_3 \text{ THEN } \hat{y}_k = 1.0126 - 1.9177u_k + 0.9555u_k^2$$
$$R^4: \text{ IF } y_k \text{ is } F_4 \text{ THEN } \hat{y}_k = 1.0123 - 1.9156u_k + 0.9539u_k^2$$

Global approach:

$$R^1: \text{ IF } y_k \text{ is } F_1 \text{ THEN } \hat{y}_k = 1.0147 - 1.9310u_k + 0.9613u_k^2$$
$$R^2: \text{ IF } y_k \text{ is } F_2 \text{ THEN } \hat{y}_k = 1.0129 - 1.9196u_k + 0.9570u_k^2$$
$$R^3: \text{ IF } y_k \text{ is } F_3 \text{ THEN } \hat{y}_k = 1.0125 - 1.9099u_k + 0.9508u_k^2$$
$$R^4: \text{ IF } y_k \text{ is } F_4 \text{ THEN } \hat{y}_k = 1.0141 - 1.9361u_k + 0.9644u_k^2$$

In this application, to ilustrate the parametric convergence property, the consequent functions have the same structure of the polynomial function. It can be seen that the consequent parameters of the obtained TS fuzzy models based on FIV estimation are close to the nonlinear function parameters in (45)-(47), which shows the robustness of the proposed FIV method in a noisy environment as well as the capability of the identified TS fuzzy models for approximation and generalization of any nonlinear function with error in variables. Two criteria, widely used in analysis of experimental data and fuzzy modeling, can be applied to evaluate the fitness of the obtained TS fuzzy models : Variance Accounted For (VAF)

$$\mathbf{VAF}(\%) = 100 \times \left[1 - \frac{var(\mathbf{Y} - \hat{\mathbf{Y}})}{var(\mathbf{Y})}\right] \tag{48}$$

where $\mathbf{Y}$ is the nominal output of the plant, $\hat{\mathbf{Y}}$ is the output of the TS fuzzy model and *var* means signal variance, and Mean Square Error (MSE)

$$\mathbf{MSE} = \frac{1}{N}\sum_{k=1}^{N}(y_k - \hat{y}_k)^2 \tag{49}$$

where y_k is the nominal output of the plant, $\hat{y}_k$ is the output of the TS fuzzy model and N is the number of points. The obtained TS fuzzy models based on LS estimation presented performance with VAF and MSE of 74.4050% and 0.0226 for the local approach and of 6.0702% and 0.0943 for the global approach, respectively. The obtained TS fuzzy models based on FIV estimation presented performance with VAF and MSE of 99.5874% and 0.0012 for the local approach and of 99.5730% and 0.0013 for the global approach, respectively. The chosen fuzzy instrumental variables satisfied the Lemmas 1-3 as well as the Theorem 1, in section 4.1 and, as a consequence, the proposed algorithm becomes more robust to the noise.

6.2 On-line identification of a second-order nonlinear dynamic system

The plant to be identified consists on a second order highly nonlinear discrete-time system

$$\begin{aligned} u_k &= u_k^i + \nu_k \\ x_{k+1} &= \frac{x_k x_{k-1}(x_k + 2.5)}{1 + x_k^2 + x_{k-1}^2} + u(k) \\ y_{k+1} &= x_{k+1} + c_k - 0.5c_{k-1} \end{aligned} \tag{50}$$

which is, without noise, a benchmark problem in neural and fuzzy modeling (Narendra & Parthasarathy (1990); Papadakis & Theocaris (2002)), where $x(k)$ is the plant output and $u_k^i = 1.5\sin(\frac{2\pi k}{25})$ is the applied input. In this case ν_k and c_k are white noise with zero mean and variance $\sigma_\nu^2 = \sigma_c^2 = 0.1$ meaning that the noise level applied to outputs takes values between

0 and ±20% from its nominal values, which is an acceptable practical percentage of noise. The rule base, for the TS fuzzy model, is of the form:

$$R^i: \text{ IF } y_k \text{ is } F^i_{1,2} \text{ AND } y_{k-1} \text{ is } G^i_{1,2} \text{ THEN}$$
$$\hat{y}_{k+1} = a_{i,1}y_k + a_{i,2}y_{k-1} + b_{i,1}u_k + c_i \tag{51}$$

where $F^i_{1,2}|^{i=1,2,\ldots,l}$ are gaussian fuzzy sets. For the FIV approach, the "fuzzy auxiliar model" has the following structure:

$$R^i: \text{ IF } y^{filt}_k \text{ is } F^i_{1,2} \text{ AND } y^{filt}_{k-1} \text{ is } G^i_{1,2} \text{ THEN}$$
$$\hat{y}^{filt}_{k+1} = a_{i,1}y^{filt}_k + a_{i,2}y^{filt}_{k-1} + b_{i,1}u_k + c_i \tag{52}$$

where $\hat{y}^{filt}$ is the filtered output, based on the consequent parameters LS estimation, and used to criate the membership functions as well as the fuzzy instrumental variable matrix. The number of rules is 4 for the TS fuzzy model, the antecedent parameters are obtained by the ECM method proposed in (Kasabov & Song (2002)). An experimental data set of 500 points is created from (50). The linguistic variables partitions obtained by the ECM method are shown in Fig. 5. The TS fuzzy model consequent parameters recursive estimate result is shown in

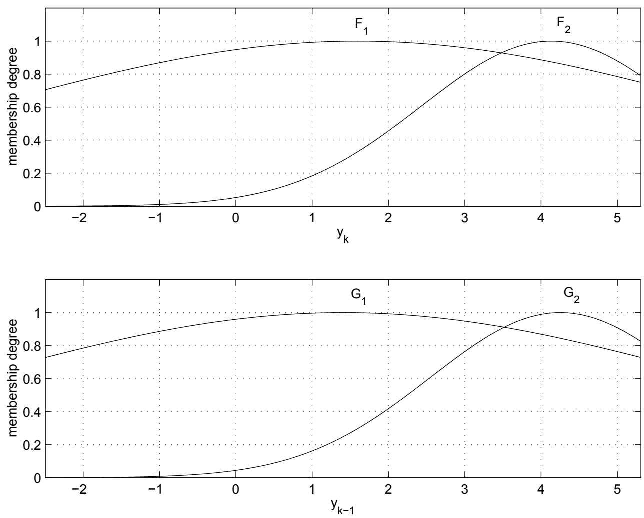


Fig. 5. Antecedent membership functions.

Fig. 6. The coefficient of determination, widely used in analysis of experimental data for time-series modeling, can be applied to evaluate the fitness of the obtained TS fuzzy models:

$$R^2_T = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)}{\sum_{i=1}^{T} y_i^2} \tag{53}$$

where y_i is the nominal output of the plant, $\hat{y}_i$ is the output of the TS fuzzy model and R_T is simply a normalized measure of the degree of explanation of the data. For its experiment the

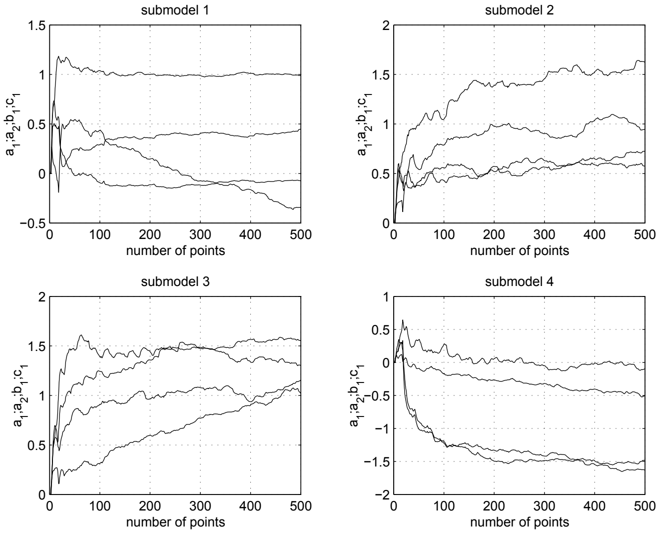


Fig. 6. Recursive consequent parameters estimate.

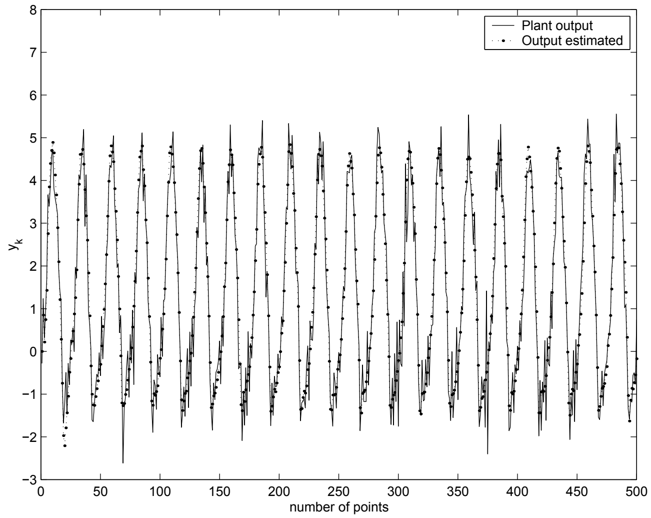


Fig. 7. Fuzzy instrumental variable output tracking.

coefficient of determination is 0.9771.
According to Fig. 6, it can be seen that the algorithm is sensitive to the nonlinear plant behavior, the parameters estimates are consistent and converge rapidly. As expected, the proposed method provides unbiased and suficiently accurate estimates of the consequent parameters and, as a consequence, high speed of convergence of the TS fuzzy model to the nonlinear plant behavior in a noisy environment. These characteristics are very important in adaptive

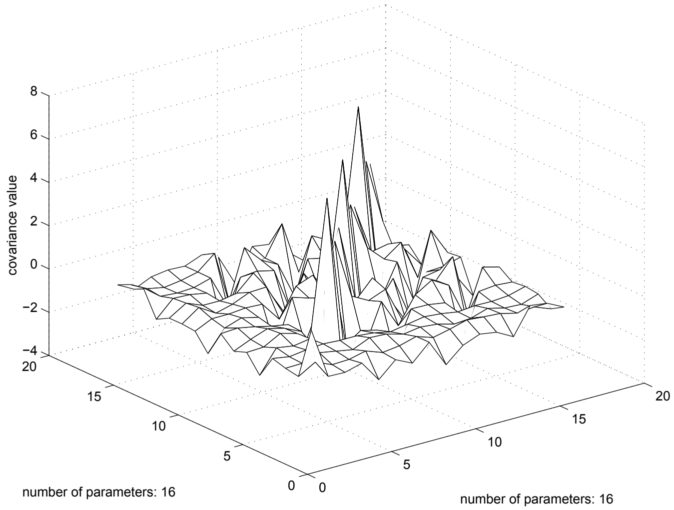


Fig. 8. Fuzzy covariance matrix P_k.

control design applications. The tracking of the nonlinear plant output is shown in Fig. 7. Figure 8 shows the fuzzy covariance matrix $\mathbf{P}_k$ of the recursive parameters estimatesfor the last point. It can be seen that the parametric uncertainty is close to zero and the higher values at this 3-D plot represent the principal diagonal entries, which determine the non-sigular property of this matrix due to fuzzy instrumental variable approach during the estimation process.

7. Conclusions

The concept of fuzzy instrumental variable and an approach for fuzzy identification of nonlinear discrete time systems were proposed. Convergence conditions for identification in a noisy environment in a fuzzy context were studied. Simulation results for off-line and on-line schemes evidence the good quality of this fuzzy instrumental variable approach for identification and function approximation with observation errors in input and output data.

8. References

Aguirre, L.A.; Coelho, M.C.S. & Correa, M.V. (2005). On the interpretation and practice of dynamical differences between Hammerstein and Wiener models, *IEE Proceedings of Control Theory and Applications*, Vol. 152, No. 4, Jul 2005, 349–356, ISSN 1350-2379.

Bergsten, P. (2001). *Observers and controllers for Takagi-Sugeno fuzzy systems*, Thesis, Örebro University.

Brown, M. & Harris, C. (1994). *Neurofuzzy adaptive modelling and control*, Prentice Hall.

Hellendoorn, H. & Driankov, D. (1997). *Fuzzy model identification: selected approaches*, Springer-Verlag.

Johansen, T.A.; Shorten, R. & Murray-Smith, R. (2000). On the interpretation and identification of dynamic Takagi-Sugeno fuzzy models, *IEEE Transactions on Fuzzy Systems*, Vol. 8, No. 3, Jun 2000, 297–313, ISSN 1063-6706.

Kadmiry, B. & Driankov, D. (2004). A fuzzy gain-scheduler for the attitude control of an unmanned helicopter, *IEEE Transactions on Fuzzy Systems*, Vol. 12, No. 3, Aug 2004, 502–515, ISSN 1063-6706.

Kasabov, N.K. & Song, Q. (2002). DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction, *IEEE Transactions on Fuzzy Systems*, Vol. 10, No. 2, Apr 2002, 144–154, ISSN 1063-6706.

King, R.E. (1999). *Computational intelligence in control engineering*, Marcel Dekker.

Ljung, L. (1999). *System Identification: Theory for the user*, Prentice Hall.

Narendra, K.S. & Parthasarathy, K. (1990). Identification and control of dynamical systems using neural networks, *IEEE Transactions on Neural Networks*, Vol. 1, No. 1, Mar 1990, 4–27, ISSN 1045-9227.

Papadakis, S.E. & Theocaris, J.B. (2002). A GA-based fuzzy modeling approach for generating TSK models, *Fuzzy Sets and Systems*, Vol. 131, No. 2, Oct 2002, 121–152, ISSN 0165-0114.

Serra, G.L.O. & Bottura, C.P. (2004). An algorithm for fuzzy identification of nonlinear discrete-time systems, *Proceedings of* 43^{rd} *IEEE Conference on Decision and Control*, Vol. 5, pp.5521–5526, ISBN 0-7803-8682-5, Bahamas, Dec. 2004, Nassau.

Serra, G. L. O. & Bottura, C. P. (2005). Fuzzy instrumental variable concept and identification algorithm, *Proceedings of* 14^{th} *IEEE Conference on Fuzzy Systems*, pp.1062–1067, ISBN 0-7803-9159-4, NV, May 2005, Reno.

Serra, G. L. O. & Bottura, C.P. (2006a). Multiobjective evolution based fuzzy PI controller design for nonlinear systems, *Engineering Applications of Artificial Intelligence*, Vol. 19, No. 2, Mar-2006, 157–167.

Serra, G. L. O. & Bottura, C. P. (2006b). An IV-QR algorithm for neuro-fuzzy multivariable on-line identification, *IEEE Transactions on Fuzzy Systems*, Vol. 15; No. 2, Apr-2007, 200–210, ISSN 1063-6706.

Sjöberg, J.; Zhang, Q.; Ljung, L.; Benveniste, A.; Delyon, B.; Glorennec, P.; Hjalmarsson, H. & Juditsky, A. (1995). Nonlinear black-box modeling in system identification : an unified overview, *Automatica: Special issue on trends in system identification*, Vol. 31, No. 12, Dec-1995, 1691–1724, ISSN 0005-1098.

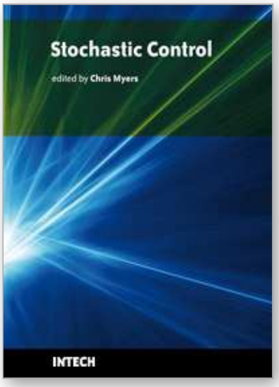
Söderström, T. and Stoica, P. (1989). *System identification*, Prentice Hall.

Söderström, T. and Stoica, P. (1983). *Instrumental variable methods for system identification*, Springer.

Takagi, T. & Sugeno, M. (1985). Fuzzy identification of systems and its aplication to modeling and control, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 15, No. 1, 116–132, ISSN: 1083-4419.

Tanaka, K.; Ikeda, T. & Wang, H. (1998). Fuzzy regulators and fuzzy observers: relaxed stability conditions and LMI-based designs, *IEEE Transactions on Fuzzy Systems*, Vol. 6, No. 2, May-1998, 250–265, ISSN 1063-6706.

Tong, S. & Li, H. (2002). Observer-based robust fuzzy control of nonlinear systems with parametric uncertainties. *Fuzzy Sets and Systems*, Vol. 131, No. 2, Oct-2002, 165–184, ISSN 0165-0114.



Stochastic Control
Edited by Chris Myers

ISBN 978-953-307-121-3
Hard cover, 650 pages
Publisher Sciyo
Published online 17, August, 2010
Published in print edition August, 2010

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821