

# Interpretability Improvement of Fuzzy Systems: Reducing the Number of Unique Singletons in Zeroth order Takagi-Sugeno Systems

Andri Riid and Ennu Rüstern, *Member, IEEE*

**Abstract**—This paper addresses one specific aspect of complexity reduction/interpretability improvement in fuzzy systems - how to limit the number of unique singletons in 0-th order Takagi-Sugeno (TS) systems, where the common practice is to assign an unique singleton to each rule. While abundance of free parameters makes 0-th order TS systems effective in data-driven identification, it also presents a computational load and an obstacle for interpretability and reliability of fuzzy rules. The developed reduction algorithm that utilizes singleton mapping matrix, subtractive clustering and least squares estimation algorithms, is able to bring the number of unique singletons down to the desired level without substantial accuracy loss.

## I. INTRODUCTION

At this date, interpretability of fuzzy systems has a research history of about 15 years and during this time, many definitions and measures of interpretability as well as different constraints for preserving interpretability have been suggested in literature ([1], [2]). If not anything else, the research has revealed that interpretability is a complex concept with many aspects.

In [3] we have argued that interpretability should not be observed as an isolated property of the system - this is because we usually want to exploit interpretability for the problem at hand. What we generally seek is an overall system *consistency* and it follows naturally that what we understand as interpretability today is the internal consistency of the fuzzy system, its consistency with itself. System interpretation, however, has no practical use if the system is numerically grossly inaccurate or its rules cannot be relied on because they express information that cannot be confirmed otherwise (e.g. by available numerical data). The latter properties (accuracy and reliability) that make up external consistency are therefore extremely important even though they are rarely seen to have anything common with interpretability (accuracy and interpretability are most often seen as conflicting properties) and what we generally aim for is a certain balance between internal and external consistency (better known as interpretability-accuracy tradeoff, however, the latter terminology is insufficient).

The main aspects of interpretability/internal consistency are transparency, linguistic integrity and complexity.

Andri Riid is with the Laboratory of Proactive Technologies, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia (phone: +372 6202 108; email: andri@dcc.ttu.ee).

Ennu Rüstern is with the Department of Computer Control, Tallinn University of Technology, Ehitajate tee 5, 19086 Tallinn, Estonia (phone: +372 6202 104; email: ennu.rustern@dcc.ttu.ee).

Transparency is perhaps most explicitly defined in our own works, e.g. [4], as a property that enables one to understand the influence of each system parameter on the output of the system and as a crucial precondition for interpretability that is forced by imposing certain constraints on the membership functions (MFs) of the system. It is noteworthy that these constraints are often satisfied in applications of fuzzy logic by default, for one reason or another.

For the most part the question of linguistic integrity (especially if system transparency is taken care for) boils down to the proper labeling of MFs. Requirements such as "the ordering of linguistic labels sets should reflect the order of membership values of corresponding fuzzy sets" or "MFs carrying semantically negative labels should not appear in the positive side of the domain" - typical linguistic integrity considerations - can be generally solved by revision and relabeling of fuzzy sets, which generally requires no other skills than common sense.

Complexity of the system is a more universal concept. Considering interpretability, complexity plays assisting role as the systems with less rules and rules with lesser components can be interpreted with less effort. For fuzzy systems, it is understood that the number of variables, the number of MFs and the number of rules should be moderate. Obviously, the issue of computational cost is also involved. In the end, the problem is about how to make the system as simple as possible without jeopardizing its functionality.

In current paper we address one specific aspect of complexity reduction - reduction of the number of output MFs in 0-th order Takagi-Sugeno systems, where these MFs are singletons and each rule is assigned an unique output singleton. While this is a specific class of fuzzy systems, it is also a very popular one (abundance of output parameters makes these systems an extremely useful tool in identification and optimization).

The algorithm that utilizes subtractive clustering [5], least squares estimation and most importantly, a mapping matrix that maps the limited number of singletons to individual rules, is able bring down the number of output singletons to a desired level without substantial accuracy loss. Furthermore, without any further performance loss, the resulting systems can be generally simplified further thanks to extensive sharing of output singletons and redundancy that can be revealed using tools [6] that have become available recently. For reliability-critical applications, appropriately modified version of Nozaki's algorithm [7] is suggested. The proposed

algorithms can be considered as a healthy alternative to genetic algorithms that have the potential to achieve the same goals in fuzzy optimization [8] but are notorious for computational overhead and algorithmical clumsiness.

Despite the fact that current study deals directly with the aspect of internal consistency (i.e. complexity), it does have implications to external consistency - obviously the reduction of the number of free parameters in the system limits its adaptation properties. On the other hand, reduction of system parameters may improve system reliability if the initial system happens to be overdetermined.

## II. PRELIMINARIES

Consider a multi-input single output fuzzy system, consisting of  $R$  rules:

$$\begin{aligned} &\text{IF } x_1 \text{ is } A_{1r} \text{ AND } x_2 \text{ is } A_{2r} \text{ AND } \dots \\ &\dots \text{ AND } x_N \text{ is } A_{Nr} \text{ THEN } y \text{ is } B_r \\ &\text{OR } \dots, \end{aligned} \quad (1)$$

where  $A_{ir}$  denote the linguistic labels of the  $i$ -th input variable associated with the  $r$ -th rule ( $i = 1, \dots, N$ ), and  $B_r$  is the linguistic label of the output variable, associated with the  $r$ -th rule.

Each  $A_{ir}$  has its representation in the numerical domain - the membership function  $\mu_{ir}$  (the same applies to  $B_r$  that is represented by  $\gamma_r$ ). In general case the inference function that computes the fuzzy output  $F(y)$  of the system (1) has the following form

$$F(y) = \bigcup_{r=1}^R \left( \left( \bigcap_{i=1}^N \mu_{ir}(x_i(k)) \right) \cap \gamma_r \right), \quad (2)$$

where  $\bigcup_r^R$  denotes the aggregation operator (corresponds to OR in (1)),  $\cap$  is the implication operator (THEN) and  $\bigcap_i^N$  is the conjunction operator (AND). In order to obtain crisp output  $y(k)$  corresponding to a given  $k$ -th ( $k = 1, \dots, K$ ) input vector  $\mathbf{x}_k = [x_1(k), x_2(k), \dots, x_N(k)]$ , (2) is generally defuzzified with center-of-gravity method

$$y = \frac{\int_Y y F(y) dy}{\int_Y F(y) dy}. \quad (3)$$

If implication and aggregation operators are product and sum respectively (product-sum inference) and  $\gamma_r$  are symmetric triangles (symmetricity of output MFs is one of the transparency requirements, see [4]), given by parameters  $b_r$ ,  $s_r$  (center and width of  $\gamma_r$ ) then (3) reduces to

$$y(k) = \frac{\sum_{r=1}^R \tau_r(k) b_r s_r}{\sum_{r=1}^R \tau_r(k) s_r}, \quad (4)$$

where  $\tau_r(k)$  is the activation degree of the  $r$ -th rule.

$$\tau_r(k) = \prod_{i=1}^N \mu_{ir}(x_i(k)). \quad (5)$$

Furthermore, if  $\gamma_r$  are of equal width ( $\forall r, s_r = \xi$ , where  $\xi$  is an arbitrary positive constant), (4) reduces to

$$y(k) = \frac{\sum_{r=1}^R \tau_r(k) b_r}{\sum_{r=1}^R \tau_r(k)}, \quad (6)$$

which is the inference function for well-known 0-th order Takagi-Sugeno systems, meaning, of course that if the output MFs of the fuzzy system (4) are of equal width, these can be with no loss of generalization reduced to scalars.

Using the notations

$$\Gamma = \begin{bmatrix} \tau_1(1) & \tau_2(1) & \dots & \tau_R(1) \\ \tau_1(2) & \tau_2(2) & \dots & \tau_R(2) \\ \dots & \dots & \dots & \dots \\ \tau_1(K) & \tau_2(K) & \dots & \tau_R(K) \end{bmatrix}, \quad (7)$$

$$\mathbf{s} = [s_1, s_2, \dots, s_R]^T, \mathbf{b} = [b_1, b_2, \dots, b_R]^T, \quad (8)$$

and

$$\mathbf{y} = [y(1), y(2), \dots, y(K)]^T, \quad (9)$$

we can easily see that (4) can be expressed by

$$\text{diag}(\Gamma \cdot \mathbf{s}) \cdot \mathbf{y} = \Gamma \cdot \text{diag}(\mathbf{s}) \cdot \mathbf{b}, \quad (10)$$

where  $\text{diag}()$  denotes the operation which transforms a column vector (its argument) into a diagonal matrix. The output vector  $\mathbf{y}$  can be computed using

$$\mathbf{y} = \text{pinv}(\text{diag}(\Gamma \cdot \mathbf{s})) \cdot \Gamma \cdot \text{diag}(\mathbf{s}) \cdot \mathbf{b}, \quad (11)$$

where  $\text{pinv}()$  is the Moore-Penrose pseudoinverse [9] that is applied for matrix inversion throughout the paper. Note, however, that in (11), the inverted matrix is a diagonal one so its inversion can as well be obtained by replacing each element in the diagonal with its reciprocal ( $\text{pinv}()$  is there for the convenience of notation).

In order to get a similar expression for (6) just replace  $\mathbf{s}$  in (11) by a  $(R \times 1)$  vector of ones ( $\mathbf{e}$ ) by which (11) reduces to

$$\mathbf{y} = \text{pinv}(\text{diag}(\Gamma \cdot \mathbf{e})) \cdot \Gamma \cdot \mathbf{b}. \quad (12)$$

Similarly, if  $\Gamma$ ,  $\mathbf{s}$  and  $\mathbf{y}$  are known, we can use the pseudoinverse to compute a least squares solution to (11) that lacks an exact solution in terms of  $\mathbf{b}$ .

$$\mathbf{b} = \text{pinv}(\Gamma \cdot \text{diag}(\mathbf{s})) \cdot \text{diag}(\Gamma \cdot \mathbf{s}) \cdot \mathbf{y} \quad (13)$$

For (12), the least squares solution is given by

$$\mathbf{b} = \text{pinv}(\Gamma) \cdot \text{diag}(\Gamma \cdot \mathbf{e}) \cdot \mathbf{y} \quad (14)$$

The latter is *de facto* standard identification method for the output parameters (singletons) of 0-th order TS systems.

If  $\Gamma$ ,  $\mathbf{b}$  and  $\mathbf{y}$  are known and  $\mathbf{s}$  is unknown, we face the homogeneous matrix equation

$$(\text{diag}(\mathbf{y}) \cdot \Gamma - \Gamma \cdot \text{diag}(\mathbf{b})) \cdot \mathbf{s} = 0 \quad (15)$$

To find a non-trivial solution to (15), we can apply singular value decomposition [11], SVD for short, i.e. find matrices  $U$ ,  $\Sigma$  and  $V$  so that

$$U \cdot \Sigma \cdot V^T = \text{diag}(\mathbf{y}) \cdot \Gamma - \Gamma \cdot \text{diag}(\mathbf{b}), \quad (16)$$

where  $\Sigma$  is a diagonal matrix and  $U$  and  $V$  are orthogonal matrices. The least-squares solution is given by the column of  $V$ , which corresponds to the smallest diagonal entry of  $\Sigma$ .

For the parameter identification of (4) we can start with (14) and apply (16) and (13) then repeatedly until the solution converges or until all elements of  $\mathbf{s}$  maintain the same sign (to preserve the physical meaning of  $s_r$ ).

Note that in a normal fuzzy system the number of MFs per  $i$ -th variable ( $S_i$ ) and output MFs ( $T$ ) is relatively small - this is also one of the complexity-driven interpretability requirements. MFs of the system are thus shared between the rules (whose number  $R$  is generally much larger). Generally, it is desired that all possible unique combinations of input MFs are represented ( $R = \prod_{i=1}^N S_i$ ). The information about which MF belongs to which rule appears in a separate  $R \times (N + 1)$  dimensional matrix that accommodates the identifiers  $m_{r,i} \in \{1, 2, \dots, S_i\}$  of MFs associated with fuzzy rules (except for the last column that corresponds to the output variable  $m_{r,N+1} \in \{1, 2, \dots, T\}$ ).

However, for 0-th order Takagi-Sugeno systems (and TS systems in general), it is common that the number of unique output MFs (singletons in this case) is equal to  $R$ , i.e. each rule is assigned a unique output MF. This usually derives from the fact that the output singletons  $\mathbf{b}$  of 0-th order TS systems are almost exclusively computed by (14). Our goal in the current paper is to develop an algorithm that keeps the number of output parameters of the system at a reasonable level, without jeopardizing its performance.

### III. THE ALGORITHM

If the number of unique output MFs is smaller than the number of rules ( $T < R$ ), it follows that these MFs must be shared among rules. Let  $\mathbf{b}'$  be a  $T \times 1$  vector of output singletons. The information about which output MFs belongs to which rule can be expressed by a  $R \times T$  mapping matrix  $M$  (that can be considered as a crisp version of the fuzzy relational matrix introduced by Pedrycz [12]), in which each row is an unity vector.

For example, given a 0-th order TS system with  $\Gamma$  that is a  $K \times 6$  matrix and

$$\mathbf{b}' = [b_1, b_2, b_3, b_4]^T \quad (17)$$

then

$$M = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (18)$$

maps  $b_1$  to the second rule,  $b_2$  to the third and sixth rule,  $b_3$  to the first and fifth rule and  $b_4$  to the fourth rule.

It follows then that we can replace  $\mathbf{b}$  with  $M \cdot \mathbf{b}'$  in (12) and  $\mathbf{s}$  and  $\text{diag}(\mathbf{s}) \cdot \mathbf{b}$  with  $M \cdot \mathbf{s}'$  and  $M \cdot \text{diag}(\mathbf{s}') \cdot \mathbf{b}$ , respectively (where  $\mathbf{s}'$  is a  $T \times 1$  vector of spreads of output MFs), in (11) so that the latter becomes

$$\mathbf{y} = \text{pinv}(\text{diag}(\Gamma \cdot M \cdot \mathbf{s}')) \cdot \Gamma \cdot M \cdot \text{diag}(\mathbf{s}') \cdot \mathbf{b}' \quad (19)$$

Assume we have a fully defined 0-th order Takagi-Sugeno system (12). The algorithm that reduces the vocabulary of

the output variable consists of three steps. In first step the initial definition of  $\mathbf{b}'$  is found. It would be natural to use some kind of clustering algorithm for finding the estimate, e.g. subtractive clustering [5] that determines the number of clusters automatically based on pre-specified cluster radius (k-means clustering [13] is used for the same problem in [14] and c-means clustering [15] is suggested in [16] so the choice of the clustering algorithm seems to be a matter of taste). These cluster centers serve only as the prototypes of final parameters. In the next step, the mapping matrix is found (initially a  $R \times T$  zero matrix). For the  $r$ -th rule the  $j$ -th cluster center that is closest to given  $b_r$  is found and the element in  $j$ -th column and  $r$ -th row in  $M$  is assigned the value of one. As soon as all  $R$  rows of  $M$  have been processed, we can proceed to the third step. In the third step output singletons are identified by

$$\mathbf{b}' = \text{pinv}(\Gamma \cdot M) \cdot \text{diag}(\Gamma \cdot \mathbf{e}) \cdot \mathbf{y}, \quad (20)$$

If our aim is to obtain a 0-th order TS system then this completes the algorithm. However, system accuracy can be potentially improved by adopting Mamdani system structure (11). If this is our desire, we can proceed to the fourth step in which  $\mathbf{s}'$  are first estimated with

$$U \cdot \Sigma \cdot V^T = \text{svd}(\text{diag}(\mathbf{y}) \cdot \Gamma \cdot M - \Gamma \cdot M \cdot \text{diag}(\mathbf{b}')), \quad (21)$$

where  $\mathbf{s}'$  is the column of  $V$ , which corresponds to the smallest diagonal entry in  $\Sigma$ , and  $\mathbf{b}'$  are re-calculated using

$$\mathbf{b}' = \text{pinv}(\Gamma \cdot M \cdot \text{diag}(\mathbf{s}')) \cdot \text{diag}(\Gamma \cdot M \cdot \mathbf{s}') \cdot \mathbf{y} \quad (22)$$

(21) and (22) can be repeated until the parameters converge (and all elements of  $\mathbf{s}'$  maintain the same sign).

We also provide a measure for estimating the information loss experienced through parameter reduction. A fuzzy relational matrix

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1T} \\ w_{21} & w_{22} & \dots & w_{2T} \\ \dots & \dots & \dots & \dots \\ w_{R1} & w_{R2} & \dots & w_{RT} \end{bmatrix} \quad (23)$$

that gives the mapping

$$b_r = \sum_{j=1}^T w_{rj} b_j, \quad (24)$$

can also be used for improving interpretability through the reduction of output vocabulary. It can be shown that a 0-th order Takagi-Sugeno system can be replaced by a numerically equivalent relational model with a limited number of output MFs. Nozaki [7] goes further, by showing that for reproducing each rule consequent  $b_r$  it is sufficient to consider just two output MFs from  $B'$  (i.e. there would be two nonzero elements in the  $r$ -th row of  $W$ ),  $b'$  and  $b''$ . The initial system (6) is transformed into a weighted fuzzy system

$$y(k) = \frac{\sum_{r=1}^R \tau_r(k) (b' w'_r + b'' w''_r)}{\sum_{r=1}^R \tau_r(k) (w'_r + w''_r)}, \quad (25)$$

where  $b'$  and  $b''$  are two closest elements from  $\mathbf{b}'$  to  $b_r$  so that  $b' < b_r < b''$  and  $w'_r, w''_r$  are corresponding rule weights. If  $\forall r$ ,

$$w'_r + w''_r = 1 \quad (26)$$

and

$$b'w'_r + b''w''_r = b_r, \quad (27)$$

it is obvious that (25) is equivalent to (12) and the weights  $w'_r, w''_r$  satisfying (26), (27) can be computed from the expressions

$$w'_r = (b_r - b'') / (b' - b'') \quad (28)$$

and

$$w''_r = (b' - b_r) / (b' - b''). \quad (29)$$

If  $W_1$  and  $W_2$  are fuzzy relational matrices containing  $w'$  and  $w''$  respectively so that  $W = W_1 + W_2$ , (25) translates into

$$\mathbf{y} = \text{pinv}(\text{diag}(\Gamma \cdot (W_1 + W_2))) \cdot \Gamma \cdot (W_1 + W_2) \cdot \mathbf{b}', \quad (30)$$

We can reorganize matrices  $W_1$  and  $W_2$  so that  $\forall r$ , from the pair  $w'_r, w''_r$  the one having larger value goes into  $W_1$  and the other one into  $W_2$ .

The ideal situation, of course, is if we have been able to identify such  $b_j$  that the nonzero weights in  $W_1$  are equal to one (and  $W_2$  is a zero matrix), in which case  $W_1 + W_2$  reduces to  $M$ , plus, it is numerically equivalent to the initial system. This, however, is a rare situation that requires original  $\mathbf{b}$  to be highly redundant. One would think that the difference of  $W_2$  from zero could serve as a measure of information loss through reduction. This is true but the measure is highly relative because the value of the weight depends heavily on the value of  $b' - b''$ . To estimate the information loss, we propose therefore an alternative measure (root mean squared error of sorts)

$$J_q = \|\mathbf{b} - M \cdot \mathbf{b}'\| / R, \quad (31)$$

#### IV. EXAMPLES

For the preliminary assessment of the performance of proposed algorithm we use two relatively simple two-input 16-rule 0-th order TS models of pressure in a laboratory fermenter and heat transfer system from [16]. The consequent parameters of these two models,  $b_r$ , are depicted in Figure 1. We can see that in case of the heat transfer process model, these parameters are lumped into four distinct groups, whereas for the pressure model, there is a slightly exponential distribution.

The training data input matrix  $X = [\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_K^T]^T$  is obtained by uniform discretization of the input space (121 samples) and corresponding output vectors  $\mathbf{y}$  are computed by original models.

To obtain the initial positions of output singletons, we use uniform distribution, k-means clustering, c-means clustering and subtractive clustering (two scenarios are considered - one with  $T = 5$ , another with  $T = 9$ ). For illustration, the initial positions of 5 output singletons of the heat transfer model obtained by various techniques are plotted in Figure 2.

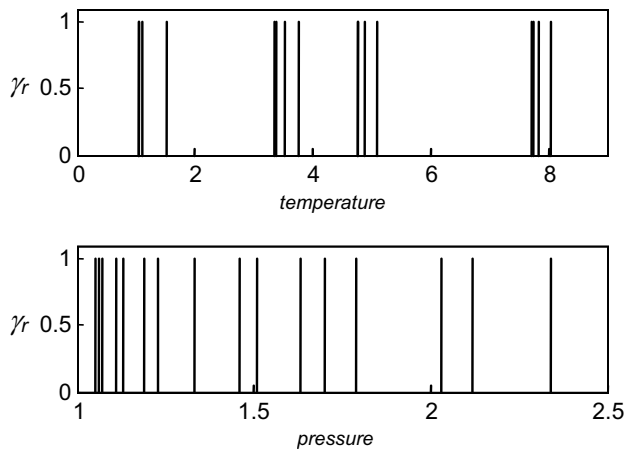


Fig. 1. Consequent parameters of two models from [16]

We can see that the main difference between clustering approaches is where they position one extra cluster center.

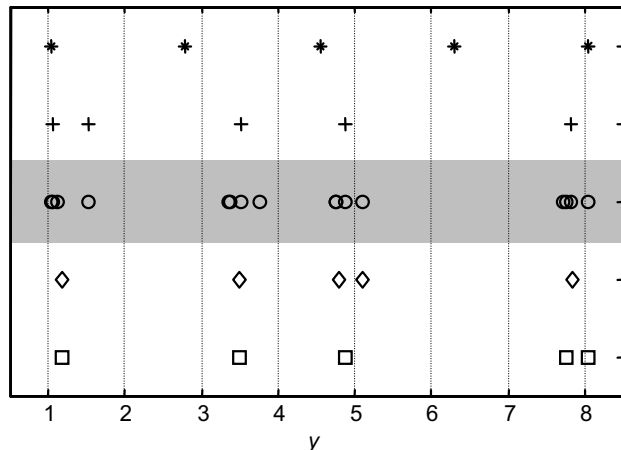


Fig. 2. Original singletons of the heat transfer model (o) vs. the uniformly distributed singleton prototypes (\*) and clusters extracted by k-means clustering (□), c-means clustering (◇) and subtractive clustering (+)

The modeling results in terms of  $\epsilon$  - modeling root mean squared error (rmse) and  $J_q$  (31) are given in Tables I- IV.

As it turns out, the initial setting (positioning of initial singletons) has a significant effect on the modeling result. As a rule, subtractive clustering leads to best results (another useful feature of this clustering approach is that it uses the available data samples as potential cluster centers and in contrast to k- or c-means clustering that are initialized randomly, always results in the same set of cluster centers). On the other hand, uniform distribution of initial singletons is particularly useless if the distribution of original singletons is highly non-linear. For example, for heat transfer process the solution (see Table II) originating from uniform distribution is able to make use of only six of available nine singletons (the remaining three are not connected to any rules through the matrix  $M$ ).

We can also see that the accuracy improvement that stems from utilization of the vector  $s'$  is generally rather small (Figure 3 depicts the output MFs of the pressure model initialized with subtractive clustering before and after the fourth step of the algorithm). In some cases, such solutions do not converge (this is indicated by blanks in the tables).

Some models converge to the same set of parameters, e.g. the solutions originating from k- and c-means clustering in Table III as well as solutions originating from the uniform distribution and subtractive clustering in Table IV.

TABLE I  
COMPARISON OF REDUCED MODELS ( $T = 5$ ) OF HEAT TRANSFER PROCESS IN TERMS OF  $\epsilon$  AND  $J_q$

clustering algorithm		initial	third step	final
uniform	$\epsilon$	0.2941	0.1637	-
	$J_q$	0.1050	0.0726	-
k-means	$\epsilon$	0.1151	0.1097	0.1078
	$J_q$	0.0354	0.0382	0.0365
c-means	$\epsilon$	0.1261	0.1243	-
	$J_q$	0.0362	0.0375	-
subtractive	$\epsilon$	0.0975	0.0957	0.0932
	$J_q$	0.0313	0.0316	0.0363

TABLE II  
COMPARISON OF REDUCED MODELS ( $T = 9$ ) OF HEAT TRANSFER SYSTEM PROCES IN TERMS OF  $\epsilon$  AND  $J_q$

clustering algorithm		initial	third step	final
uniform	$\epsilon$	0.1745	0.0797	-
	$J_q$	0.0601	0.0287	-
k-means	$\epsilon$	0.0493	0.0465	0.0381
	$J_q$	0.0197	0.0207	0.0220
c-means	$\epsilon$	0.0339	0.0310	0.0285
	$J_q$	0.0103	0.0109	0.0106
subtractive	$\epsilon$	0.0291	0.0245	0.0228
	$J_q$	0.0095	0.0090	0.0087

TABLE III  
COMPARISON OF REDUCED MODELS ( $T = 5$ ) OF PRESSURE IN THE FERMENTER IN TERMS OF  $\epsilon$  AND  $J_q$

clustering algorithm		initial	third step	final
uniform	$\epsilon$	0.0688	0.0481	-
	$J_q$	0.0200	0.0195	-
k-means	$\epsilon$	0.0507	0.0488	0.0470
	$J_q$	0.0177	0.0181	0.0178
c-means	$\epsilon$	0.0525	0.0488	0.0470
	$J_q$	0.0179	0.0181	0.0178
subtractive	$\epsilon$	0.0385	0.0336	0.0307
	$J_q$	0.0165	0.0166	0.0162

Another, more demanding case study of the algorithm is performed on a fault diagnosis problem in a wastewater anaerobic digestion process (where organic matter is decomposed into biogas through biological processes taking place in the absence of oxygen). We focus on the acidogenic state of the process that is important to detect properly. The data set that originates from the LBE, a laboratory situated in Narbonne, France, consists of 559 samples coming from a

TABLE IV  
COMPARISON OF REDUCED MODELS ( $T = 9$ ) OF PRESSURE IN THE FERMENTER IN TERMS OF  $\epsilon$  AND  $J_q$

clustering algorithm		initial	third step	final
uniform	$\epsilon$	0.0325	0.0127	0.0120
	$J_q$	0.0111	0.0063	0.0064
k-means	$\epsilon$	0.0388	0.0360	-
	$J_q$	0.0147	0.0155	-
c-means	$\epsilon$	0.0161	0.0143	0.0132
	$J_q$	0.0070	0.0073	0.0072
subtractive	$\epsilon$	0.0186	0.0127	0.0120
	$J_q$	0.0076	0.0063	0.0064

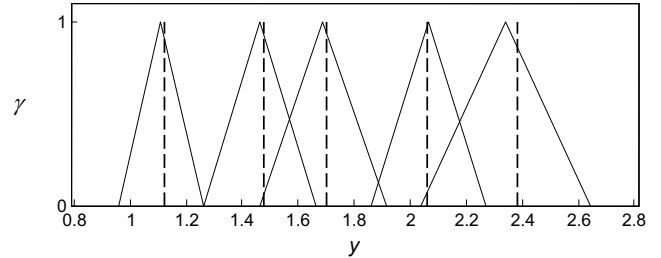


Fig. 3. Singletons (dashed lines) and triangular MFs of the output variable in the identified pressure model

pilot-scale up-flow anaerobic fixed bed reactor with a volume of  $0.984m^3$ . Four input variables -  $pH$  (pH in the reactor),  $vfa$  (volatile fatty acid concentration),  $qIn$  (input flow rate) and  $CH_4$  ( $CH_4$  concentration in biogas) - are considered. The output is a number from 0 to 1, measuring to what extent the actual state can be considered acidogenic.

The original model in [14] consisting of 53 rules and producing rmse of 0.046 is obtained by the application of interpretability preserving modification of orthogonal least squares (OLS). The output singletons of the model are shown in Figure 4. It is noteworthy that quite a few of these are located outside  $[0,1]$ , which can be contributed to the properties of OLS algorithm and is a sign of phenomenon of unreliability of the model. Output vocabulary reduction in [14] by which the number of distinct output singletons reduces to 6 from 51 is based on k-means clustering and includes some additional fiddling so as to get all singletons into  $[0,1]$ . As a consequence, the modeling rmse increases to 0.056.

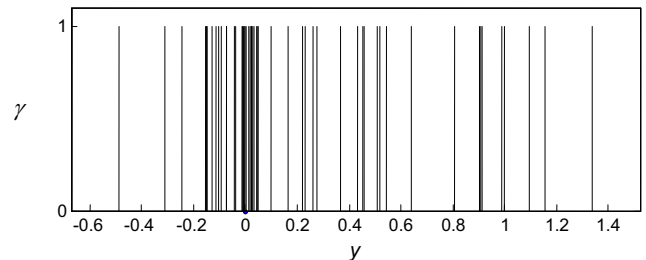


Fig. 4. Original singletons of the acidogenic state model

We apply subtractive clustering (6 clusters) in association

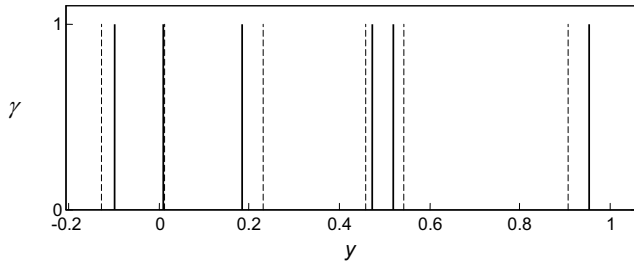


Fig. 5. Singletons after applying the proposed algorithm (dashed lines are the singleton prototypes estimated by subtractive clustering)

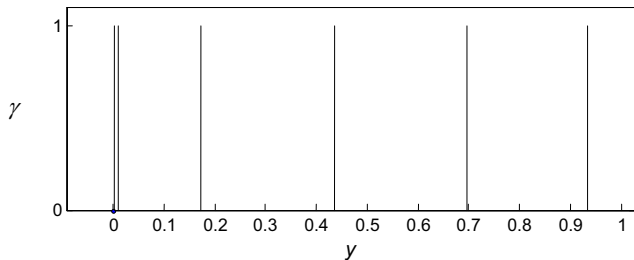


Fig. 6. Singletons after applying modified Nozaki's algorithm

with our algorithm, by which we reach the error value of 0.0501.

Only one of the singletons is outside  $[0,1]$  (see Figure 5). This may not still be acceptable for reliability-critical applications and for such cases we adopt the heuristic approach proposed in [7] that in matrix form appears as

$$\mathbf{b} = \text{pinv}(\text{diag}(\Gamma^T \cdot \mathbf{e})) \cdot \Gamma^T \cdot \mathbf{y}, \quad (32)$$

where  $\mathbf{e}$  is a  $R \times 1$  vector of ones. However, for the identification of  $\mathbf{b}'$  we apply it in the following form

$$\mathbf{b}' = \text{pinv}(\text{diag}(M^T \cdot \Gamma^T \cdot \mathbf{e})) \cdot M^T \cdot \Gamma^T \cdot \mathbf{y}, \quad (33)$$

Note that in (32) as well as in (33),  $\Gamma$  may be exponentiated elementwise with  $m$  so that each element in  $\Gamma$  becomes  $(\tau_r(k))^m$  for achieving desired reliability-accuracy tradeoff. Using  $m = 10$  we are able to obtain  $\epsilon = 0.0598$  with the singletons depicted in Figure 6. For convenience, the results of this modeling experiment are given in Table V.

TABLE V  
REDUCTION OF THE ORIGINAL ACIDOGENIC STATE MODEL FROM [14]  
WITH  $\epsilon = 0.046$

	k-means [14]	proposed algorithm	modified Nozaki
$\epsilon$	0.0560	0.0501	0.0598
$J_q$	0.0166	0.0138	0.0189

## V. CONCLUSIONS

Overparameterization of Takagi-Sugeno systems that makes them so powerful in identification has its setbacks in unnecessary complexity and non-reliability. In current paper we developed several algorithms for the reduction of the number of output MFs so as to minimize information loss and accuracy degradation, depending on the context of the problem. A welcome side effect is observed - reduction of the number of output singletons that introduces extensive sharing of those singletons, makes the system a prey for a recently developed tool for redundancy detection and removal [6] that, when applied, reduces the number of rules in heat transfer model from 16 to 5, the number of rules in the pressure model from 16 to 11 and the number of rules in the acidogenic state model from 53 to 40 without any further performance degradation.

## ACKNOWLEDGMENT

This work was partially supported by Estonian Science Foundation under grant no. 6837. The authors would also like to thank Sebastien Destercke, Serge Guillaume and Brigitte Charnomordic for their cooperation.

## REFERENCES

- [1] J. Casillas, O. Cordon, F. Herrera and L. Magdalena (Eds.), *Interpretability Issues in Fuzzy Modeling*, Springer, Berlin, 2003.
- [2] C. Mencar and A. M. Fanelli, "Interpretability constraints for fuzzy information granulation," *Information Sciences: an International Journal*, vol. 178, no. 24, pp. 4585-4618, 2008.
- [3] A. Riid and E. Rüstern, "Interpretability of Fuzzy Systems and Its Application to Process Control," *Proc. IEEE Int. Conf. Fuzzy Systems*, London, 6 pp., 2007.
- [4] A. Riid and E. Rüstern, "Transparent fuzzy systems and modeling with transparency protection," *Proc. IFAC Symp. on Artificial Intelligence in Real Time Control*, Budapest, pp. 229-234, 2000.
- [5] S. L. Chiu, "Fuzzy Model Identification Based on Cluster Estimation," *Journal of Intelligent and Fuzzy Systems*, vol. 2, pp. 267-278, 1994.
- [6] A. Riid, K. Saastamoinen and E. Rüstern, "Error-free Simplification of Transparent Mamdani Systems," *Proc. IEEE Int. Conf. Intelligent Systems*, Varna, vol. 1, pp. 2.8-2.13, 2008.
- [7] K. Nozaki, H. Ishibuchi and H. Tanaka, "A simple but powerful heuristic method for generating fuzzy rules from numerical data," *Fuzzy Sets and Systems*, vol. 65, pp. 251-270, 1997.
- [8] M. Antonelli, P. Ducange and B. Lazzarini, "Learning Concurrently Granularity, Membership Function Parameters and Rules of Mamdani Fuzzy Rule-based Systems," *Proc. 2009 IFSA World Congress/EUSFLAT Conf.*, Lisbon, pp. 1033-1038, 2009.
- [9] R. Penrose, "A generalized inverse for matrices," *Proc. Cambridge Philosophical Society*, vol. 51, pp. 406-413, 1955.
- [10] G. H. Golub and W. Kahan, "Calculating the singular values and pseudo-inverse of a matrix," *Journal of the Society for Industrial and Applied Mathematics: Series B, Numerical Analysis*, vol. 2, no. 2, pp. 205-224, 1965.
- [11] W. Pedrycz, "Applications of fuzzy relational equations for methods of reasoning in presence of fuzzy data," *Fuzzy Sets and Systems*, vol. 16, pp. 163-175, 1985.
- [12] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Information Theory*, vol. 28, no. 2, pp. 1291-137, 1982.
- [13] S. Destercke, S. Guillaume and B. Charnomordic, "Building an interpretable fuzzy rule base from data using Orthogonal Least Squares - Application to a depollution problem," *Fuzzy Sets and Systems*, vol. 158, no. 18, pp. 2078-2094, 2007.
- [14] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function*, Plenum Press, New York, 1981.
- [15] R. Babuska *Fuzzy Modeling for Control*, Kluwer Academic Publishers, 1998.