

Possibilistic identification of reliable finite impulse response models

Dominik Hose* and Michael Hanss

Institute of Engineering and Computational Mechanics, University of Stuttgart, Germany

*Corresponding author: dominik.hose[at]itm.uni-stuttgart.de

Abstract

In this contribution, a novel approach to system identification based on concepts from possibility theory is presented. It can be employed to infer a fuzzy-valued finite-impulse response from input/output data yielding reliable predictions about the behavior of the identified dynamical system. Furthermore, a procedure for the analysis of the system properties in a possibilistic context is provided.

Keywords: Possibilistic Regression Analysis, System Identification, Finite Impulse Response, Fuzzy-Valued Modeling, Robust Control.

1 Introduction

System identification is an important task in a variety of disciplines where a full white-box model of the system under consideration is not easily obtained. One of the most simple approaches for black-box modeling, i.e. when no prior information about the system is available, is to consider the time-discrete finite-impulse response of the system [1]. This is often the first step in any system identification procedure and requires a regression analysis to determine the correlation between the inputs and outputs of the system.

In classical statistics, the aim is to find an estimator for the finite-impulse response, discriminating between what is considered the actual signal and what is assumed to be noise. The standard estimator for the finite-impulse response is the solution of an ordinary least-squares problem [2]. While providing a reasonable fit depending on the number of parameters and the training signal that is being employed, this estimation technique is not concerned with the necessary parameter variations to cover the whole output training signal. The residuals are attributed to noise even though the respective output may still contain system-relevant signals.

To overcome these limitations, the assumed model is considered to be imperfect and not able to account for all of the system dynamics by using just one vector of parameters. Consequently, a (fuzzy) set of parameter vectors is identified that is able to provide reliable bounds for the system output and enable e.g. the design of robust controllers.

To the authors' knowledge, currently, few approaches to identifying fuzzy-valued parameters of dynamical systems exist. However, fuzzy regression analysis has recently received much attention. The multitude of proposed fuzzy regression algorithms (e.g. [3] or [4]) could be employed to identify fuzzy-valued finite-impulse-response models. Their major drawback is that they typically fail to provide a motivation for the choice of the membership functions of the regression parameters which are usually assumed to be of triangular shape. Other approaches for identifying fuzzy-valued model parameters from data can be found in [5], [6] and [7]. In addition to the aforementioned problems, they deal with highly non-convex optimization problems which can lead to computational infeasibilities. A methodology for identifying interval predictor models has been presented in [8] which yields comparable results – yet it does not provide a fuzzy membership function encoding further information about the involved uncertainties.

The presented approach is based on possibilistic regression analysis [9] and provides a tool to analyze the necessary parameter variations under the assumption that possible residuals stem from unmodeled dynamics – rather than from noise – leading to the conclusion that all the dynamics contained in the data should be accounted for by the fuzzy-valued parameters. Simultaneously, meaningful information is encoded in their membership function.

The resulting fuzzy-valued finite-impulse-response model is then employed to perform a possibilistic stability analysis and for the synthesis of control laws that are robust against the identified uncertainties. An illustrative example showing the advantages compared to the classical least-squares approach is provided.

2 Possibility Theory and Fuzzy Sets

Possibility theory provides a framework for measuring uncertainties similar to probability theory. The possibility measure $\text{Pos} : 2^\Omega \rightarrow [0, 1]$ on the universe of discourse Ω is based on the three axioms

1. $\text{Pos}(\Omega) = 1$,
2. $\text{Pos}(\emptyset) = 0$,
3. $\text{Pos}(A \cup B) = \max[\text{Pos}(A), \text{Pos}(B)]$ for two disjoint subsets $A, B \subseteq \Omega$.

From these axioms, it follows that the possibility of an event A is defined by the maximum possibility of all elements contained therein, i.e.

$$\text{Pos}(A) = \max_{a \in A} \text{Pos}(\{a\}) = \max_{a \in A} \pi(a) , \quad (1)$$

where π is called a possibility distribution. If A is uncountably infinite, the max-operator is replaced by the sup-operator. Since this measure is not self-dual, the necessity measure defined by

$$\text{Nec}(A) = 1 - \text{Pos}(\Omega/A) \quad \forall A \subseteq \Omega \quad (2)$$

is introduced additionally. For a detailed discussion of the extensive theoretic background of possibility theory, refer to e.g. [10].

Fuzzy sets were initially introduced by Zadeh [11]. They differ from classical sets in the way that the classical characteristic function, which assumes values of either 1 or 0, is generalized to a membership function $\mu : \Theta \rightarrow [0, 1]$ where Θ is a universal set (here $\Theta = \mathbb{R}^n$), formalizing the notion of gradual memberships. Each fuzzy set is uniquely defined by its membership function. The space of fuzzy sets on the reals is denoted $\mathcal{F}\mathbb{R}$, and the space of fuzzy vectors on \mathbb{R}^n is denoted $\mathcal{F}\mathbb{R}^n$. A detailed discussion of fuzzy sets and the important subclass of fuzzy numbers can be found in [12]. Additionally, the support of a fuzzy set $\tilde{x} \in \mathcal{F}\mathbb{R}^n$ is the closure of all elements with non-zero membership

$$\text{supp}(\tilde{x}) = \overline{\{\mathbf{x} \in \mathbb{R}^n : \mu_{\tilde{x}}(\mathbf{x}) > 0\}} . \quad (3)$$

Fuzzy theory is strongly connected to possibility theory [13] as hinted above. More precisely, the membership function $\mu_{\tilde{x}}$ of a possibilistic variable, i.e. a fuzzy variable $\tilde{x} : \Omega \rightarrow \mathbb{R}$, induces a possibility distribution $\pi_{\tilde{x}}$, according to

$$\mu_{\tilde{x}}(x) = \text{Pos}(\{a \in \Omega : \tilde{x}(a) = x\}) = \pi_{\tilde{x}}(x) , \quad (4)$$

in the same way a probability density function of a random variable induces a probability distribution. The value of $\pi_{\tilde{x}}(x)$ is the possibility that \tilde{x} assumes the value x , and the resulting possibility measure is

$$\text{Pos}_{\tilde{x}}(X) = \sup_{x \in X} \mu_{\tilde{x}}(x) \quad \forall X \subseteq \mathbb{R} . \quad (5)$$

The subscript denotes a possibility measure induced by a fuzzy variable.

3 Possibility Propagation

Fuzzy arithmetic, i.e. the forward propagation of fuzzy variables or possibility distributions through deterministic mappings, is fundamentally based on the extension principle formulated by Zadeh in [14]. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a mapping and let $\tilde{x} \in \mathcal{F}\mathbb{R}^n$ be a fuzzy vector with membership function $\mu_{\tilde{x}}$. The membership function $\mu_{\tilde{y}}$ of the fuzzy output $\tilde{y} \in \mathcal{F}\mathbb{R}^m$ defined by

$$\tilde{y} = f(\tilde{x}) \quad (6)$$

is given by

$$\mu_{\tilde{y}}(\mathbf{y}) = \sup_{\mathbf{x} : \mathbf{y} = f(\mathbf{x})} \mu_{\tilde{x}}(\mathbf{x}) . \quad (7)$$

If f is not surjective, it is not possible to find an \mathbf{x} satisfying $\mathbf{y} = f(\mathbf{x})$ for all $\mathbf{y} \in \mathbb{R}^m$. In this case, the membership of the respective element \mathbf{y} is simply zero. A proof of the extension principle employing possibilistic arguments is given in [15]. Several implementations for efficient computations with fuzzy sets and fuzzy numbers exist and are discussed e.g. in [12] or in [5].

4 Finite-Impulse-Response Models

Any time-discrete LTI (linear time-invariant) system \mathcal{G} is uniquely defined by its impulse-response coefficients $g[\cdot]$, allowing to compute the output signal $y[\cdot]$ at the k -th time instant for a given input signal $u[\cdot]$ by the infinite discrete-time convolution equation

$$y[k] = \sum_{n=-\infty}^{\infty} g[n]u[k-n]. \quad (8)$$

A thorough discussion of this type of system representation can be found e.g. in [1]. Therein, it is stated that

"merely knowing the impulse response $g[\cdot]$ is sufficient to predict the response of the system \mathcal{G} to an arbitrary input."

Here, causality of the system in question, i.e. $g[k] = 0$ for all $k < 0$, is assumed. For asymptotically stable systems, it can furthermore be shown that the impulse response coefficients $g[k]$ tend to zero for large k and can be considered negligible after some M time instants. The infinite sum then reduces to the finite convolution

$$y[k] = \sum_{n=0}^M g[n]u[k-n] \quad (9)$$

which is known as the finite-impulse-response (FIR) model of the system.

5 System Identification

Given a measured input $u[\cdot]$ and output $y[\cdot]$ at $N + 1$ sampling instants, the standard approach to the identification of the FIR model given in Eq. (9) is to include some additive measurement noise ν . Accordingly, the measured output signal is the superposition of the true output signal being the average of the input signal weighted with the finite-impulse response and the perturbation, i.e.

$$y[k] = \sum_{n=0}^M g[n]u[k-n] + \nu[k]. \quad (10)$$

Since the input for negative time instants is generally unknown, this equation can usually only be evaluated for $k \geq M$. Defining the vector of FIR coefficients $\boldsymbol{\theta}$, the output vector \mathbf{y} and the input matrix $\boldsymbol{\Psi}$ respectively by

$$\boldsymbol{\theta} = (g[0] \quad \dots \quad g[M])^T, \quad \mathbf{y} = (y[M] \quad \dots \quad y[N])^T, \quad \boldsymbol{\Psi} = \begin{pmatrix} u[0] & \dots & u[M] \\ \vdots & \ddots & \vdots \\ u[N-M] & \dots & u[N] \end{pmatrix}, \quad (11)$$

the noise vector $\boldsymbol{\nu}$ can be expressed by $\boldsymbol{\nu} = \mathbf{y} - \boldsymbol{\Psi}\boldsymbol{\theta}$. If the noise is white, i.e. it is uncorrelated, has a zero mean and finite variance, the maximum likelihood estimator for $\boldsymbol{\theta}$ is given by the solution to the ordinary least-squares problem

$$\min_{\boldsymbol{\theta}} \|\mathbf{y} - \boldsymbol{\Psi}\boldsymbol{\theta}\|_2, \quad (12)$$

whose analytic solution is given by

$$\boldsymbol{\theta}^{\text{LS}} = (\boldsymbol{\Psi}^T \boldsymbol{\Psi})^{-1} \boldsymbol{\Psi}^T \mathbf{y}. \quad (13)$$

A detailed description of the problem formulation, several solution methods and further implications are presented in [2]. This estimator is able to reproduce a reasonable approximation of the actual system dynamics. Evidently, the quality of the estimated FIR model depends largely on the quality of the data, i.e. the frequencies contained in the test signal, the signal-to-noise ratio, the 'whiteness' of the noise, etc. The choice of good test signals is a science in itself and covered e.g. in [1].

It is a justifiable point of criticism of this method that it is not able to provide robust predictions. If, for instance, the data did not contain noise but were obtained from the identification of a high-order system with complex (perhaps non-linear) dynamics and long settling time, one would still not expect the predictions to exactly match the data. This can partially be attributed to the fact that the assumption of uncorrelated noise is unwarranted in this case since the prediction error is actually resulting from system dynamics that are not accounted for in the model.

6 Possibilistic Regression Analysis

Possibilistic regression analysis [9] provides a tool to overcome the problems mentioned above by including the possibility that every sample is actually an undisturbed result of the system dynamics.

The first step is to find the $K = N - M + 1$ elementary parameters θ^i that exactly reproduce the output samples

$$y_i = y[i + M - 1], \quad \forall i = 1, \dots, K \quad (14)$$

from the correlating inputs and yield the minimum deviation from the remaining data points. The i -th row of Ψ is denoted by

$$\psi_i = (u[i - 1] \quad \dots \quad u[i + M - 1])^T, \quad \forall i = 1, \dots, K \quad (15)$$

and provides these inputs. Hence, all admissible parameters that produce a predicted output signal passing through the respective output sample y_i are geometrically located on the line

$$\mathcal{S}^i = \{\theta \in \mathbb{R}^M : \psi_i^T \theta - y_i = 0\}, \quad \forall i = 1, \dots, K. \quad (16)$$

Minimum deviation from the remaining data is ensured by employing e.g. the least-squares objective function. Of course, it is possible to choose other p -norms, that may be more appropriate for a given problem, such as the 1- or the ∞ -norm. Even the minimization of the distance to the least-squares estimate

$$J(\theta) = \|\theta - \theta^{\text{LS}}\|_2 \quad (17)$$

would be acceptable in order to ensure a fuzzy set with a tight support.

Choosing the 2-norm, the problem of finding the elementary parameters can be formulated as solving the K convex optimization problems

$$\min_{\theta^i \in \mathcal{S}^i} \|\mathbf{y} - \Psi \theta^i\|_2, \quad \forall i = 1, \dots, K. \quad (18)$$

The Lagrange function of this constrained quadratic optimization problem is given by

$$\mathcal{L}(\theta^i, \lambda) = (\Psi \theta^i - \mathbf{y})^T (\Psi \theta^i - \mathbf{y}) + \lambda_i (\psi_i^T \theta^i - y_i), \quad \forall i = 1, \dots, K \quad (19)$$

and the analytic solution yields

$$\begin{pmatrix} \theta^i \\ \lambda^i \end{pmatrix} = \begin{pmatrix} 2\Psi^T \Psi & \psi_i \\ \psi_i^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} 2\Psi^T \mathbf{y} \\ y_i \end{pmatrix}, \quad \forall i = 1, \dots, K. \quad (20)$$

It is also convenient to set $\theta^0 = \theta^{\text{LS}}$. Evidently, the support of a fuzzy-valued parameter $\tilde{\theta}$ should be an outer approximation of the convex hull of these elementary parameters. The most intuitive – yet weak – approximation is the hypercube

$$\mathcal{C} = \{\theta \in \mathbb{R}^M : \theta_j^{\min} \leq \theta_j \leq \theta_j^{\max}, \forall j = 1, \dots, M\} \quad (21)$$

where

$$\theta_j^{\min} = \min_{i=0, \dots, K} \theta_j^i \quad \text{and} \quad \theta_j^{\max} = \max_{i=0, \dots, K} \theta_j^i \quad (22)$$

are the element-wise minima and maxima of the elementary parameters, respectively. Generally, other techniques, leading to alternative approximations of the convex hull, such as bounding hyperspheres or -ellipsoids, are also possible.

Intuitively, the objective values

$$J^i = \|\mathbf{y} - \Psi \theta^i\|_2 \quad \forall i = 0, \dots, K \quad (23)$$

represent a measure for the plausibility of the corresponding elementary parameters. The larger J^i is, the less plausible, i.e. consistent with the data, the respective elementary parameter is due to a larger deviation measured by the 2-norm. E.g. the least plausible elementary parameter has the objective value

$$J^{\max} = \max_{i=0, \dots, K} J^i. \quad (24)$$

This sensitivity information can be encoded in the membership function of the fuzzy FIR coefficients $\tilde{\theta}$ by constructing a minimizing set [16]

$$\mathbf{m} : J \mapsto \frac{J - J^{\max}}{J^{\text{LS}} - J^{\max}} \quad (25)$$

and defining the membership function

$$\mu_{\tilde{\theta}} : \theta \mapsto \max [0, \mathbf{m} (\|\mathbf{y} - \Psi\theta\|_2)] . \quad (26)$$

This particular choice for the membership function yields the least-squares solution as the nominal parameter and every elementary parameter is contained in the support by construction. Most importantly, *the membership of the parameters is directly derived from their consistency with the data*. Moreover, it ensures that the reference signal $y[\cdot]$ is contained in the fuzzy output domain of the fuzzy-valued finite-impulse-response (fFIR) model. A detailed discussion of the possibilistic regression method is presented in [9].

7 Possibilistic System Analysis

The presented identification procedure allows for robust predictions of the output of the identified system. This follows from the fact that all dynamics that are possibly contained in the data are accounted for by at least one parameter contained in the support of the fFIR coefficients $\tilde{\theta}$. If the $M + 1$ past inputs are known, it is possible to predict $\tilde{y}[k]$ by evaluating

$$\tilde{y}[k] = \sum_{n=0}^M \tilde{\theta}_n u[k-n] . \quad (27)$$

This fFIR model can be used to perform a possibilistic system analysis. As an example, consider the task to design a stable feedback control law $u[k] = -Cy[k]$. Stability can be achieved if the largest pole of the closed loop λ^{\max} is located inside the unit circle. The fuzzy-valued largest closed-loop pole $\tilde{\lambda}_{\max}$ is computed by applying the extension principle to the solution operator $E : \theta \mapsto \lambda_{\max}$ finding the largest root of the closed-loop characteristic polynomial

$$p(z) = z^M + C \sum_{n=0}^M \tilde{\theta}_n z^{M-n} , \quad (28)$$

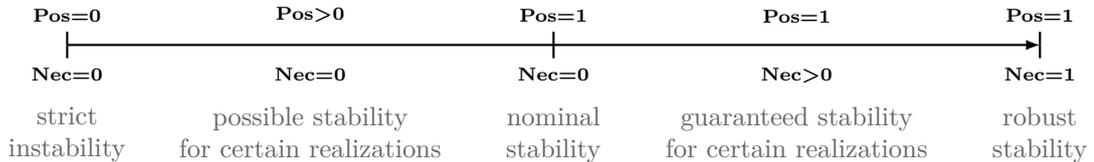
i.e. by evaluating $\tilde{\lambda}_{\max} = E[\tilde{\theta}]$. The (stability) event

$$S : \|\tilde{\lambda}_{\max}\|_2 < 1 \quad (29)$$

depends on the uncertain fFIR coefficients $\tilde{\theta}$, and thus, it is uncertain itself. Consequently, it has to be evaluated possibilistically, i.e. the degrees of possibility and necessity to which the event S occurs have to be evaluated according to

$$\text{Pos}_{\tilde{\lambda}_{\max}}(S) = \sup_{\lambda : \|\lambda\|_2 < 1} \mu_{\tilde{\lambda}_{\max}}(\lambda) . \quad (30)$$

In general, the following cases can be distinguished:



In particular, if stability is to be guaranteed robustly, then the closed loop has to fulfill

$$\text{Nec}(S) = 1 . \quad (31)$$

This translates to the whole support of $\tilde{\lambda}_{\max}$ being inside the unit circle. Evidently, the identified fFIR model admits a smaller feasible set of robust controllers compared to the least-squares FIR model since the latter is contained in the former. More precisely, the fFIR model will not admit any controller that would destabilize any of the elementary dynamics described by the elementary parameters θ^i . *All the possible dynamics contained in the data are stabilized*. For further insight and a detailed explanation of the concept of possibilistic stability analysis, refer to [17].

8 Application

For the sake of clearness, a simple example will be given in this section. Suppose a time-discrete LTI system is given by the difference equation

$$y[k] - \frac{3}{2}y[k-1] + \frac{11}{16}y[k-2] - \frac{3}{32}y[k-3] = u[k-1] - u[k-2] + \frac{2}{9}u[k-3]. \quad (32)$$

It will be shown that it suffices to identify an fFIR model of order $M = 4$ for a reasonable approximation. In order to do so, the system is excited with a random input signal of length $N = 50$ with zero mean and unit variance.

The signals predicted by the identified elementary FIR coefficients θ^i are shown in Fig. 1. The close fits suggest that the model is chosen with a sufficient number of degrees of freedom permitting a tight approximation of the input/output behaviour.

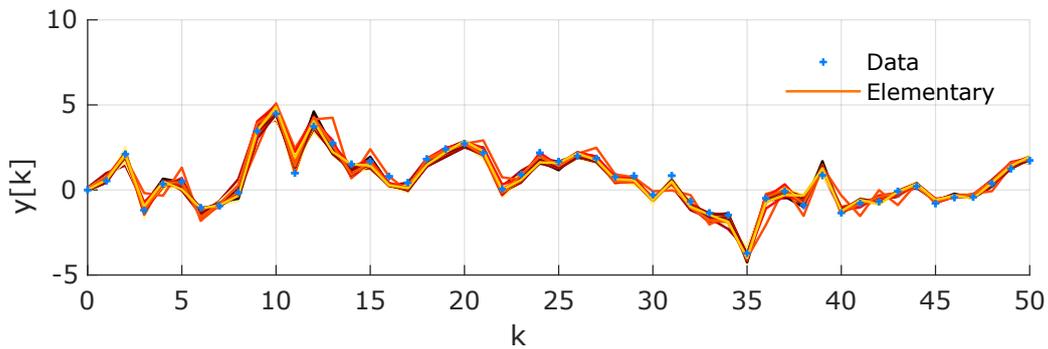


Figure 1: Signals predicted by the elementary parameters.

The resulting fFIR coefficients $\tilde{\theta}$ are shown in Fig. 2. The true value of the impulse response is contained in the fuzzy coefficients and the identified uncertainty in the fFIR coefficients compensates for the neglected coefficients of the true impulse response.

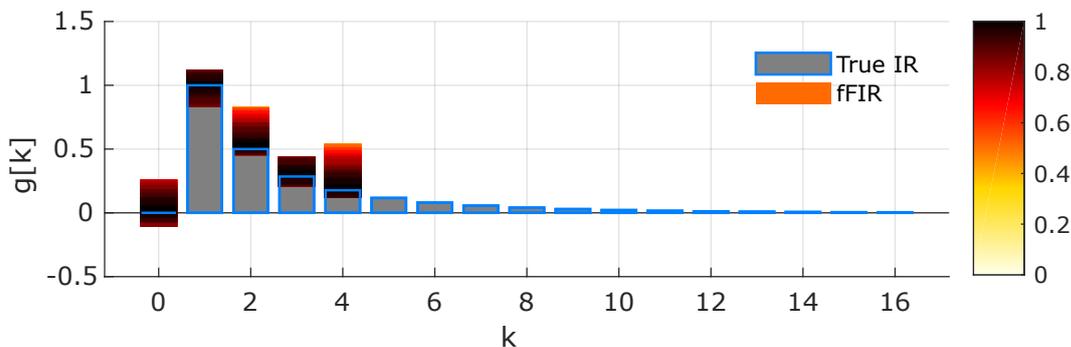


Figure 2: (Fuzzy-valued) Finite-impulse-response coefficients.

For validation purposes, the system is excited with a chirp input $u[k] = \sin\left(\frac{1}{20}k^2\right)$. The identified fFIR model reliably predicts the system output as shown in Fig. 3 since the support of the predicted signal covers the entire reference signal with an adequate amount of uncertainty contained in the prediction supporting the confidence in the identified fFIR model.

Finally, a controller $u[k] = -Cy[k]$ can be synthesized for the identified model. The fuzzy-valued largest closed-loop pole $\tilde{\lambda}_{\max}$ depending on the feedback gain C is shown in Fig. 4. A possibilistic evaluation of the stability event S is depicted in Fig. 5. Notice, that the procedure provides a reliable estimate for the region of stabilizing controllers. The possibilistic regression algorithm guarantees stability in the region where $\text{Nec}(S) = 1$ which is contained in the true region where the system is stable. In contrast, the FIR model identified by the classical least-squares procedure predicts stability where the true system would actually be unstable.

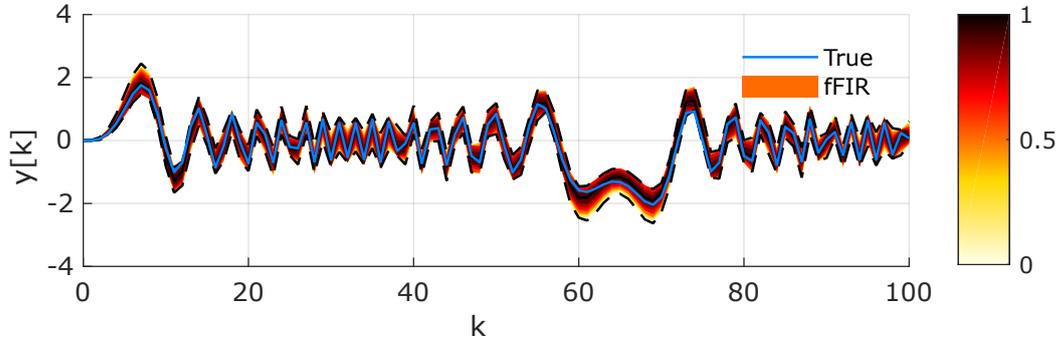


Figure 3: Validation: Predicted response to chirp input.

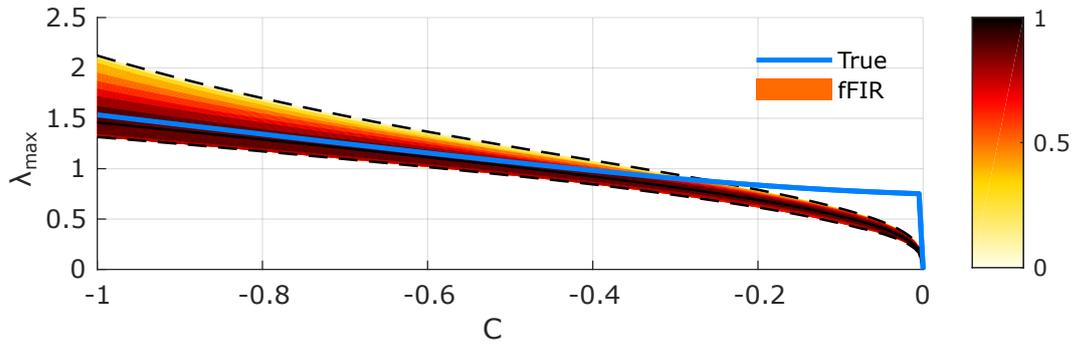


Figure 4: Largest closed-loop pole depending on the feedback gain C .

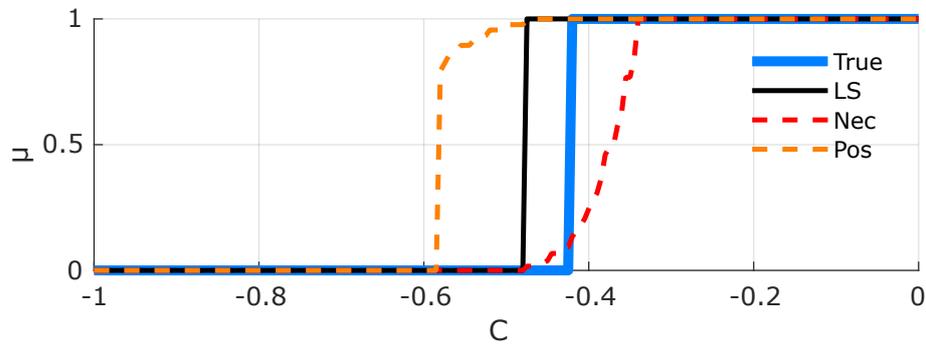


Figure 5: Evaluation of closed-loop stability for different feedback gains C .

9 Conclusion

Despite the exemplary nature of the proposed procedure, it is generally applicable to a broad range of problems. Any other identification technique based on linear regression, such as auto-regressive modeling, can be performed in the same manner and other properties of the fuzzy-valued dynamical systems can be analyzed by means of possibilistic measures. The use of the proposed framework is recommended when reliable predictions about the completely unknown system behavior are required and the uncertainties are not assumed to stem from stochastic noise.

Ultimately, the presented approach extends the paradigm of building reliable input/output models by finding (fuzzy) set-valued model parameters in order to account for system-inherent uncertainties. Increasing computational power facilitates the evaluation of these models, and thus, their analysis can become a valuable tool for any engineer and scientist aiming at overcoming the limitations inherent to crisp-valued models.

Acknowledgements

The authors gratefully acknowledge the support provided by the German Research Foundation (DFG) in the framework of the research project no. 319924547.

References

- [1] A. K. Tangirala, *Principles of System Identification: Theory and Practice* (Crc Press, 2014).
- [2] R. Isermann and M. Münchhof, *Identification of Dynamic Systems* (Springer-Verlag Berlin Heidelberg, 2012).
- [3] H. Tanaka, I. Hayashi, and J. Watada, Possibilistic linear regression analysis for fuzzy data, *European Journal of Operational Research* **40**, 389 (1989), ISSN 0377-2217.
- [4] H. Ishibuchi and M. Nii, Fuzzy regression using asymmetric fuzzy coefficients and fuzzified neural networks, *Fuzzy Sets and Systems* **119**, 273 (2001), ISSN 0165-0114.
- [5] N.-P. Walz, *Fuzzy Arithmetical Methods for Possibilistic Uncertainty Analysis*, Ph.D. thesis, Institute of Engineering and Computational Mechanics, University of Stuttgart (Shaker-Verlag, Aachen, 2016).
- [6] A. Hanselowski, S. Ihrle, and M. Hanss, A Fuzzy Model Updating Technique Motivated by Bayesian Inference, in *Proceedings of the UNCECOMP 2015, 1st ECOMMAS Thematic Conference on Uncertainty Quantification in Computational Sciences and Engineering* (2015), 548–559.
- [7] T. Haag, J. Herrmann, and M. Hanss, Identification procedure for epistemic uncertainties using inverse fuzzy arithmetic, *Mechanical Systems and Signal Processing* **24**, 2021 (2010).
- [8] L. G. Crespo, S. P. Kenny, and D. P. Giesy, Interval predictor models with a linear parameter dependency, *Journal of Verification, Validation and Uncertainty Quantification* **1**, 021007 (2016).
- [9] D. Hose and M. Hanss, Fuzzy linear least squares for the identification of possibilistic regression models, *to appear in Fuzzy Sets and Systems* (2018).
- [10] D. Dubois and H. Prade, *Possibility theory: an approach to computerized processing of uncertainty* (Plenum Press, 1988).
- [11] L. A. Zadeh, Fuzzy sets, *Information and Control* **8**, 338 (1965).
- [12] M. Hanss, *Applied Fuzzy Arithmetic - An Introduction with Engineering Applications* (Springer, Berlin, 2010).
- [13] L. A. Zadeh, Fuzzy sets as a basis for a theory of possibility, *Fuzzy sets and systems* **100**, 9 (1999).
- [14] L. A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning-i, *Information Sciences* **8**, 199 (1975), ISSN 0020-0255.
- [15] B. Liu, *Theory and practice of uncertain programming* (Springer, 2002).
- [16] L. A. Zadeh, Maximizing sets and fuzzy markoff algorithms, *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **28**, 9 (1998), ISSN 1094-6977.
- [17] D. Hose, M. Mäck, and M. Hanss, A Possibilistic Approach to the Optimization of Uncertain Systems, in *Proceedings of the ICVRAM-ISUMA 2018, 3rd International Conference on Vulnerability and Risk Analysis and Management, and the 7th International Symposium on Uncertainty Modeling and Analysis* (2018).