On Optimal Test Signal Design and Parameter Identification Schemes for Dynamic Takagi-Sugeno Fuzzy Models Using the Fisher Information Matrix

Matthias Himmelsbach[1] · **Andreas Kroll[1]**

Abstract This paper is concerned with the analysis of optimization procedures for optimal experiment design for locally affine Takagi-Sugeno (TS) fuzzy models based on the Fisher Information Matrix (FIM). The FIM is used to estimate the covariance matrix of a parameter estimate. It depends on the model parameters as well as the regression variables. Due to the dependency on the model parameters good initial models are required. Since the FIM is a matrix, a scalar measure of the FIM is optimized. Different measures and optimization goals are investigated in three case studies.

1 Introduction

The most effective methods for simulation, prognosis, and control of technical systems are based on parametric models. In control design, these models are either used for simulation during the control algorithm design or its parameters are used to design a controller. Besides the conventional physically motivated modelling, system identification offers an alternative. Since the biggest effort in control design is the modelling process [1, 2], the ever-growing complexity of technological applications leads to the necessity of expert knowledge that is not freely available. This difficulty is supplemented by the fact that not all phenomena are well enough understood to be physically modelled.

This contribution is concerned with shaping the process of generating suitable data for the identification of parameters with low parameter uncertainty. For this purpose, the optimal experiment design for the identification of TS-models is investigated. Multisine, multistep and a combination of both signals are used as parametric signal models to relax the optimization problem. The impact of different scalar measures of the FIM are investigated as well as different experiment design procedures that group the TS model parameters into local model and partition parameters. Two independent goals, the reduction of the uncertainty and the improvement of the simulation quality on arbitrary validation data, are investigated. After a short introduction of the model class and the identification process, the method is presented followed by the presentation of three case studies, a simulated test system that has been used as a benchmark system before [3] as well as a test stand of a servo-pneumatic longitudinal drive (SPLD) that was used in Zaidi and Kroll [4], and a three-tank system to validate the results from the parameter study.

2 Literature Overview

The experiment design based on the works of Fisher at the beginning of the twentieth century has been steadily extended and improved [5–7]. This was driven by the need to analyze the significance of impact factors in regression applications. The industry has a high interest in generating

✉ Matthias Himmelsbach
matthias.himmelsbach@mrt.uni-kassel.de

Andreas Kroll
andreas.kroll@mrt.uni-kassel.de

[1] Department of Measurement and Control, Faculty of Mechanical Engineering, University of Kassel, Mönchebergstr. 7, 34125 Kassel, DE, Germany

system inputs to improve the data generation for parametrized models, e.g., to reduce the required experiment time to obtain a suitable model quality.

Experiment design can be categorized in five ways: model-based vs. model-free, online vs. offline, linear vs. nonlinear, static vs. dynamic, and bias vs. variance reduction. There are other goals that can influence the experiment design, e.g., structure selection or instrumentation. These factors will not be addressed in this contribution. In process model-free designs, space-filling approaches are used. In the field of dynamic test signal design, these approaches make use of standard test signals, comprehensively presented in [8, 9] and their modification possibilities w.r.t. the identification of nonlinear models in [10–12]. Model based approaches are often optimal experiment designs (OED) that use the model equations to minimize some criterion.

The minimization of a measure on the FIM is an OED approach and is the focus of this contribution. Kroll and Dürrbaum [13] and Dürrbaum and Kroll [14] are motivations for this work. In Kroll and Dürrbaum [13], the FIM is used to determine the optimum experiment design for the estimation of the parameters of a static TS model. Only the determinant is used as a scalar measure on the FIM. In Dürrbaum and Kroll [14], the aspect of using space-filling designs to remedy the effect of wrongly assumed partition parameters is introduced to static models. Although some of the aspects are already present in the static case this contribution addresses the dynamic case in a systematic way, analyzing the impact of different optimization procedures and FIM measures with respect to achievable uncertainty improvement as well as model prediction quality. The actual implementation of the design scheme for nonlinear dynamic systems is also addressed which is left open in many other contributions. The broad term "space-filling" design also means broadband excitation signals in the dynamic case, making the actual approach for dynamic systems very different from its motivation based in static systems. The problems that arise from the application of FIM-based experiment design to nonlinear system identification are addressed in other contributions but the actual implementation of such designs for dynamic systems requires a holistic view of the problem. In the seminal work [15] the problem is stated in a way making it similar to an optimal control problem. The emergence of a parameter-dependent FIM is briefly discussed but the complete ramifications of using dynamic models in context of the FIM are not addressed. In Hametner [16], FIM-based test signal design is done online to handle constraints locally. In Heinz [17], input signals are not the result of an optimization but selected w.r.t. model-based measures. Larsson and Hjalmarsson [18] addresses the structure of the optimization problem. In Fang and Shenton [19] the problem of backpropagation of the model output for the construction of the FIM is addressed for a model that is linear in its parameters (LiP) which reduces the complexity of the FIM significantly.

In recent works [20], confidence regions obtained from a posteriori analysis are parametrized and used for subsequent designs. [21] circumvents the calculation of the posterior by representing the desired samples as a reference trajectory, such that the input design task is transformed into an optimal control problem. Neilsen [22] uses the FIM to assess the uncertainty of other quantities than the model parameters that are more important to the applications dealt with. Pleşu [23] uses the FIM to assess the global sensitivity before any experimental design because the importance of high-quality initial models is emphasized. In Walz et al. [24], the importance of task-oriented experiment design is stressed to significantly reduce the experiment costs both in time and computation, because irrelevant system behavior is excluded from the modelling as well as the experiment design. Current works on fuzzy systems focus on the control design, such as Wang et al. [25] and Shen et al. [26] that make use of the persistent dwell-time switching law to design multi-objective fault-tolerant and H_∞ controllers.

The focus of this contribution is the investigation of design choices of the complete process from initial design to optimal experiments. In previously mentioned works that use an optimization procedure to generate an optimal input the problem of constraints is either bypassed using online designs or the problem of finding constraints is moved to the specific application. Without suitable constraints, no technically feasible signals will arise. In this contribution, parametrized signal models are used to reduce the number of optimization parameters as well as to exploit the inherent signal structure. This contribution specifically targets TS models and therefore their unique model structure can be exploited in the design process. This contribution addresses the details of using optimal model-based designs for dynamic nonlinear modelling like the necessity of a (high quality) initial simulation model (obtained using process model-free designs) to run the optimization. In many other works, the determinant is used as a measure for the FIM [27, 28]. In this contribution, the impact of different measures is investigated.

3 Problem Statement

The goal is to estimate parameters of locally affine TS models given a system

$$y(k) = f(y(k-1), \ldots, y(k-n),$$
$$\qquad u(k-1), \ldots, u(k-m)) + \epsilon(k) \qquad (1)$$

with additive noise $\epsilon(k)$, input $u(k)$, and output $y(k)$. In order to identify the model parameters with minimal uncertainty, a test signal design scheme is deployed to generate suitable data. The scheme is based on the FIM that is used to estimate the covariance matrix of the model parameters. In the case of independently and identically normally distributed noise (i.i.d.) with variance σ^2 and mean $\mu = 0$ and for an unbiased estimator the FIM is given as

$$I = \frac{1}{\sigma^2} \sum_{k=1}^{N} \left[\left[\frac{\partial \hat{y}(k)}{\partial \boldsymbol{\Theta}} \right] \cdot \left[\frac{\partial \hat{y}(k)}{\partial \boldsymbol{\Theta}} \right]^{\top} \right]_{\boldsymbol{\Theta} = \boldsymbol{\Theta}_0} \qquad (2)$$

with the number of observations N and the parameter vector $\boldsymbol{\Theta}$. In the case of a nonlinear dynamic system, the following holds: The FIM depends on the model parameters, therefore an initial parameter vector $\boldsymbol{\Theta}_0$ is necessary for the optimization scheme. The derivatives of the model output $\hat{y}(k)$ contain lagged values of the output and the input $u(k)$, therefore a change in $u(K)$ affects all values in $\hat{y}(k > K)$.

To assess the uncertainty, scalar measures of the FIM are used, which leads to the following optimization problem:

$$\boldsymbol{u}_{\text{opt}} = \arg \min_{\boldsymbol{u}} J(I) \qquad (3)$$

$\boldsymbol{u} = \{u(k)\}$ is the vector of the input timeseries and J is a scalar measure of the Fisher Information Matrix (FIM) I. To solve this problem, an (initial) simulation model is used to calculate the output values during the optimization which is similar to optimal control problems. However, due to the nonlinearity of the cost functions J, which will be introduced later, general purpose optimization tools must be used. To achieve feasibility, different signal models are used to reduce the number of optimization parameters. The focus of this contribution is the analysis of the impact of design choices like signal models, FIM measures, and optimization procedures on the ability of the test signals to generate suitable identification data, as shown in Fig. 1.
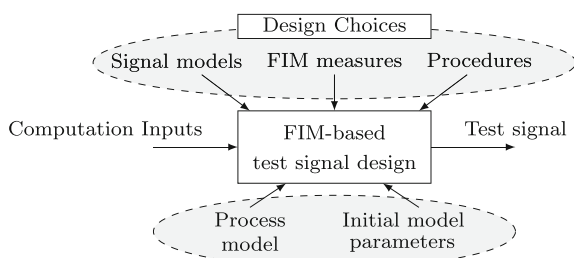

Fig. 1 Impact factors of the optimization scheme

The computation inputs describe all explicit inputs to the optimization procedure like the hyperparameterization of the optimization algorithm or the boundaries of the signal model parameter values.

In preliminary investigations, variation of these factors resulted in vastly different results, therefore the focus of this contribution is the impact analysis of said factors.

4 Model Class and Identification Process

4.1 Locally Affine Fuzzy Takagi-Sugeno Models

In this contribution, locally affine dynamical Takagi-Sugeno models are considered. They are universal approximators [29]. TS models are a superposition of c local models $\hat{y}_i(k)$ that are weighted by their respective fuzzy basis functions $\phi_i(k)$ (FBF). The prediction equation of such a model is:

$$\hat{y}(k) = \sum_{i=1}^{c} \phi_i(z(k), \boldsymbol{\Theta}_{\text{MF}}) \cdot \hat{y}_i(\boldsymbol{\varphi}(k), \boldsymbol{\Theta}_{\text{LM},i}) \qquad (4)$$

In (4), the FBF depend on the scheduling variable $z(k)$ and the fuzzy membership (MF) parameters $\boldsymbol{\Theta}_{\text{MF}}$ (see 10) which determine the partitioning into local models. The local models depend on the regression variable $\boldsymbol{\varphi}(k)$ and the i-th local parameter vector $\boldsymbol{\Theta}_{\text{LM},i}$. Local models in ARX configuration (autoregressive with exogenous input) can be written as:

$$\hat{y}_i(k) = \boldsymbol{\varphi}^{\top}(k) \cdot \boldsymbol{\Theta}_{\text{LM},i} \qquad (5)$$

with the regression variable:

$$\boldsymbol{\varphi}^{\top}(k) = [1 \quad y(k-1) \quad \cdots \quad y(k-n) \\ \qquad u(k-1-\tau) \quad u(k-m-\tau)] \qquad (6)$$

The dead time τ is not considered in the following and the i-th local parameter vector is comprised of the coefficients $a_{i,j}$ of the lagged outputs, $b_{i,j}$ of the lagged inputs and the affine term c_i.

$$\boldsymbol{\Theta}_{\text{LM},i}^{\top} = [c_i \quad a_{i,1} \quad \cdots \quad a_{i,n} \quad b_{i,1} \quad \cdots \quad b_{i,m}] \qquad (7)$$

This leads to an alternative formulation of (5):

$$\hat{y}_i(k) = \sum_{j=1}^{n} a_{i,j} y(k-j) + \sum_{j=1}^{m} b_{i,j} u(k-j) + c_i \qquad (8)$$

In locally affine TS models, the nonlinear behavior of the global system is encoded in the partitioning that is defined by the FBF. Therefore the choice of the scheduling variable $z(k)$ and the membership function parameters $\boldsymbol{\Theta}_{\text{MF}}$ define the global nonlinear structure. The FBF depend on the scheduling variable as well as the membership function

parameters. In practice, the scheduling variable is chosen using prior knowledge. Without the use of brute force methods [30], it is typically set to be equal to the regressor. The case of external scheduling is excluded in this contribution. In this contribution it is assumed that the scheduling space is part of the regression space. This is a plausible approach for the modelling of dynamic systems. Additionally, in this contribution it is assumed that $z(k)$ can be derived with a linear mapping from the regression vector.

The FBF are chosen as prototype-based membership functions of the fuzzy-c-means (FCM) cluster algorithm:

$$\mu_i(z(k), \boldsymbol{\Theta}_{\text{MF}}) = \left[\sum_{j=1}^{c} \left(\frac{\|z(k) - v_i\|}{\|z(k) - v_j\|} \right)^{\frac{2}{v-1}} \right]^{-1} \tag{9}$$

In (9), $v > 1$ is the fuzziness parameter that regulates how sharp the transitions between partitions are. It can be shown that membership functions (9) are normalized, meaning $\sum_{j=1}^{c} \mu_j(z(k), \boldsymbol{\Theta}_{\text{MF}}) \equiv 1 \, \forall z(k)$. It is possible to choose different distance norms $\|\cdot\|$. In this contribution, the Euclidean distance is used. The prototypes v_i make up the vector of membership function parameters:

$$\boldsymbol{\Theta}_{\text{MF}}^{\top} = \begin{bmatrix} v_1^{\top} & \cdots & v_c^{\top} \end{bmatrix}^{\top} \tag{10}$$

4.2 Identification Process

The identification of the parameters in (4) is conducted in four steps:

1. Determination of the hyperparameters n, m, c, v
2. Calculation of the initial partitioning through clustering
3. Calculation of the initial local model parameters using ordinary least-squares (LS) estimation
4. Optimization of the model parameters using parameter values from previous steps for initialization

Step 1. It can be conducted in various ways. Cluster validity measures like in Juhász et al. [31] should be used with caution in context of dynamic models, since these measures treat the problem as quasi-static which might not reflect the dynamical dependencies. In the model equations, v, c as well as the input and output lags n and m are hyperparameters. These parameters are determined by structure selection methods and prior knowledge.

Step 2. With the chosen scheduling variable and hyperparameters c and v, the initial partitioning can be calculated. By FCM-clustering [32] the prototypes v_i are obtained. Since the FCM-clustering algorithm converges locally, it is conducted multiple times with random initializations.

Step 3. For each data point $z(k)$ its membership $\mu_{i,k}(z(k))$ to each cluster is known. Using (5), (4) can be rewritten as follows:

$$\begin{aligned} \hat{y}(k) &= \sum_{i=1}^{c} \mu_i \boldsymbol{\varphi}^{\top} \boldsymbol{\Theta}_{\text{LM},i} \\ &= \underbrace{\begin{bmatrix} \mu_1 \boldsymbol{\varphi}^{\top} & \cdots & \mu_c \boldsymbol{\varphi}^{\top} \end{bmatrix}}_{\boldsymbol{\varphi}_{\text{E}}^{\top}} \cdot \underbrace{\begin{bmatrix} \boldsymbol{\Theta}_{\text{LM},1}^{\top} & \cdots & \boldsymbol{\Theta}_{\text{LM},c}^{\top} \end{bmatrix}^{\top}}_{\boldsymbol{\Theta}_{\text{LM}}} \\ &:= \boldsymbol{\varphi}_{\text{E}}^{\top} \cdot \boldsymbol{\Theta}_{\text{LM}} \end{aligned} \tag{11}$$

For readability, the arguments of the terms in (11) are dropped. In (11), $\boldsymbol{\varphi}_{\text{E}}^{\top}$ is called the extended regression vector and $\boldsymbol{\Theta}_{\text{LM}}$ the local model (LM) parameter vector. The evaluation of the model in $N - n$ datapoints in the case of $n \geq m$ leads to the following equations:

$$\underbrace{\begin{bmatrix} \hat{y}(n) \\ \vdots \\ \hat{y}(N) \end{bmatrix}}_{\hat{Y}} = \underbrace{\begin{bmatrix} \boldsymbol{\varphi}_{\text{E}}^{\top}(n) \\ \vdots \\ \boldsymbol{\varphi}_{\text{E}}^{\top}(N) \end{bmatrix}}_{\boldsymbol{\Phi}_{\text{E}}} \boldsymbol{\Theta}_{\text{LM}} \tag{12}$$

Setting the model output equal to the measured output $\boldsymbol{Y}^{\top} = \begin{bmatrix} y(n) \cdots y(N) \end{bmatrix}$ leads to an overdetermined system of linear equations. If a quadratic cost function is minimized, the resulting parameters $\hat{\boldsymbol{\Theta}}_{\text{LM}}$ are the local model parameters of the nonlinear ARX (NARX) model. The quadratic cost function in the prediction error method (PEM) framework is:

$$J_{\text{PEM}}(\boldsymbol{\Theta}_{\text{LM}}) = \left\| \hat{\boldsymbol{Y}}(\boldsymbol{\Theta}_{\text{LM}}) - \boldsymbol{Y} \right\|^2 \tag{13}$$

For the Euclidean norm, the minimizing argument is:

$$\hat{\boldsymbol{\Theta}}_{\text{LM}} = \arg \min_{\boldsymbol{\Theta}_{\text{LM}}} J_{\text{PEM}} = \left(\boldsymbol{\Phi}_{\text{E}}^{\top} \boldsymbol{\Phi}_{\text{E}} \right)^{-1} \boldsymbol{\Phi}_{\text{E}}^{\top} \boldsymbol{Y} \tag{14}$$

Step 4. Parameter values determined in steps 2 & 3 can be used to initialize the nonlinear optimization. The partition parameters from clustering and the local model parameters from (14) are not optimal w.r.t. the recursive model evaluation. The recursive model evaluation will be called simulation, since only initial conditions are needed. The local model parameters and the partition parameters are subsumed into the model parameter vector:

$$\boldsymbol{\Theta}^{\top} = \begin{bmatrix} \boldsymbol{\Theta}_{\text{LM}}^{\top} & \boldsymbol{\Theta}_{\text{MF}}^{\top} \end{bmatrix}^{\top} \tag{15}$$

The objective function of the nonlinear optimization problem is:

$$\hat{\boldsymbol{\Theta}} = \arg \min_{\boldsymbol{\Theta}} \sum_{k=1}^{N} (y(k) - \hat{y}(k, \boldsymbol{\Theta}))^2 \tag{16}$$

The nonlinear optimization is conducted using the MATLAB function `lsqnonlin` with the Levenberg-Marquardt algorithm.

4.3 Fisher Information Matrix for Dynamical TS Models

The Fisher Information Matrix I for i.i.d. normally distributed noise is given in (2). The Cramér-Rao-Lower-Bound (CRLB) relates the covariance matrix of the estimated parameters with the FIM:

$$\text{cov}\left(\hat{\boldsymbol{\Theta}}\right) \geq \left(I\left(\hat{\boldsymbol{\Theta}}\right)\right)^{-1} \tag{17}$$

Since the parameters of a TS model can be divided into two groups, the FIM is treated in the same way. The derivatives with respect to the local model parameters result in the extended regression vector:

$$\frac{\partial \hat{y}(k)}{\partial \boldsymbol{\Theta}_{\text{LM}}} = \boldsymbol{\varphi}_{\text{E}} \tag{18}$$

The derivatives in the NARX case with respect to the partition parameters are more complicated:

$$\frac{\partial \hat{y}(k)}{\partial \boldsymbol{\Theta}_{\text{MF}}} = \sum_{i=1}^{c}\left(\frac{\partial \mu_i(k)}{\partial \boldsymbol{\Theta}_{\text{MF}}}\hat{y}_i(k) + \mu_i(k)\underbrace{\frac{\partial \hat{y}_i(k)}{\partial \boldsymbol{\Theta}_{\text{MF}}}}_{=\boldsymbol{0}}\right) \tag{19}$$

$$\frac{\partial \mu_i(k)}{\partial \boldsymbol{\Theta}_{\text{MF}}} = \left[\frac{\partial \mu_i(k)}{\partial \boldsymbol{v}_1} \cdots \frac{\partial \mu_i(k)}{\partial \boldsymbol{v}_p} \cdots \frac{\partial \mu_i(k)}{\partial \boldsymbol{v}_c}\right] \tag{20}$$

$$\frac{\partial \mu_i(k)}{\partial \boldsymbol{v}_p} = \begin{cases} \dfrac{2\mu_i(k)^2}{v-1}\dfrac{z(k)-\boldsymbol{v}}{|z(k)-\boldsymbol{v}_i|}\left(\dfrac{1}{\mu_i(k)}-1\right) & i = p \\[4mm] -\dfrac{2\mu_i(k)^2}{v-1}\left(z(k)-\boldsymbol{v}_p\right)\dfrac{|z(k)-\boldsymbol{v}_i|^{\frac{2}{v-1}}}{|z(k)-\boldsymbol{v}_p|^{\frac{v+1}{v-1}}} & i \neq p \end{cases}$$

$$\text{with } v > 1 \tag{21}$$

5 FIM-Based Test Signal Design

In this section, the application of the FIM-based optimal test signal design to nonlinear dynamic TS models is presented.

5.1 Design Process

There are significant differences in procedures w.r.t. static and/or linear FIM-based OED problems of which not all have been addressed in the literature:

- A change in the input at time k impacts all future output values.

- The input sequence $\{u(k)\}$ is a trajectory.
- $\{u(k)\}$ is too long to be optimized for each $u(k)$.
- The FIM cannot be evaluated without model parameter values.

The consequences of these aspects are significant. In the case of a linear static problem, the FIM only depends on the choice of independent inputs, therefore an optimization procedure places these input values within specified bounds. In case of a nonlinear static problem, the optimization procedure is localized since model parameters are needed to calculate the FIM. In the case of a dynamic problem, the FIM contains system outputs that change based on the iterations of the input. Therefore, during the optimization a simulation model is needed to generate the system outputs in each iteration. The proposed test signal design method follows in five steps:

1. Initial signal model parameters $\boldsymbol{\beta}_0$ are used to generate a test signal $\boldsymbol{u}^{(0)}(\boldsymbol{\beta}_0)$ using the signal model.
2. The initial model $\mathcal{M}_0$ is used to generate the initial output $\boldsymbol{y}^{(0)}$.
3. The test signal, the output, and the initial model parameters are used to calculate the FIM.
4. A scalar measure on the inverse FIM is calculated.
5. IF: Termination criterion is not met, THEN: Signal parameters are adjusted, ELSE: Optimal test signal is found.

The procedure is presented in Fig. 2. The process model as well as the FIM calculation depend on the initial model parameters $\boldsymbol{\Theta}_0$ that have been obtained from process model-free identification. Since gradient-based optimization schemes converge to local minima, a particle swarm algorithm has been used. Since the absolute values differ w.r.t. various factors, a convergence-based criterion was chosen. A threshold of $\epsilon_{\text{termination}} = 10^{-4}$ for the relative change of the best solution has been selected as a termination criterion. This means that in each iteration step, the model has to be simulated for each particle, causing significant computational cost. Before different aspects of the method are clarified in the following subsections, problem (3) has to be reformulated to correspond to Fig. 2:

$$\boldsymbol{\beta}_{\text{opt}} = \arg\min_{\boldsymbol{\beta}} J(I(\hat{\boldsymbol{y}}(\boldsymbol{u}(\boldsymbol{\beta},\boldsymbol{\Theta}),\boldsymbol{u}(\boldsymbol{\beta},\boldsymbol{\Theta})))) \tag{22}$$

The goal is an optimal test signal, but the test signal depends on signal model parameters $\boldsymbol{\beta}$. Therefore the actual optimization is w.r.t. those parameters. The adjustment of the signal model parameters is subject to a measure on the FIM. The FIM depends on a test signal, a simulated model output as well as the initial model parameters. The simulated output itself depends on the test signal and the model parameters. Therefore, the optimization of the FIM is a
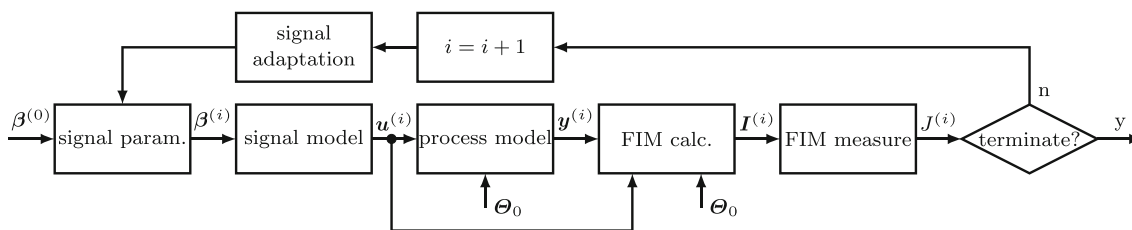
Fig. 2 Test signal Optimization scheme

highly complex, non-convex, and cascaded problem which is addressed in empirical studies in this contribution.

5.2 Measures on the Fisher Information Matrix

Due to the CRLB there is an inverse relation between the FIM and the covariance matrix of the parameter estimates. With i.i.d. normally distributed zero mean noise, the eigenvalues of the covariance matrix represent the spread in the direction of the half-axes of an ellipsoidal uncertainty region that is characterized by the eigenvectors. Therefore, the most commonly used measures on the FIM use eigenvalues. (2) can be viewed as a matrix that contains the sensitivity of the model output w.r.t. the model parameters. Therefore, the optimization goal is reformulated without rigorous statistical assumptions: The input should be chosen in a way that the model output is most sensitive to a change in the parameters. If the sensitivities are high, the values that any parameter can take are reduced and therefore indirectly the spread of the model parameters is reduced, even though "the spread" cannot be interpreted as an estimation of statistical quantities in a rigorous manner. If a matrix is assessed by a scalar measure, information is lost. Therefore, different measures will be compared:

– M1: Determinant of the inverse FIM J_{det}
– M2: Maximum Eigenvalue of the inverse FIM J_{eig}
– M4: Scaled trace of the inverse FIM J_{str}
– M3: Trace of the inverse FIM J_{tr}
– M5: Squared sum of the inverse FIM elements J_{sens}

The measures are implemented as follows:

$$J_{\text{det}} = \log\big(\det(\boldsymbol{I}^{-1})\big) = -\log(\det(\boldsymbol{I})) \tag{23}$$

$$J_{\text{eig}} = \max\Big(\text{eig}(\boldsymbol{I})^{-1}\Big) \tag{24}$$

$$J_{\text{str}} = \sum_{k=1}^{n_{\text{p}}} \frac{\lambda_{\boldsymbol{I},\text{init},k}}{\lambda_{\boldsymbol{I},\text{opt},k}} \tag{25}$$

$$J_{\text{tr}} = -\log(\text{tr}(\boldsymbol{I})) \tag{26}$$

$$J_{\text{sens}} = \sum_{k=1}^{n_{\text{p}}} \sum_{l=1}^{n_{\text{p}}} g_{k,l}^2 \tag{27}$$

In (23) and (26) the negative logarithm of the determinant and the trace is used to avoid calculating the inverse as well as to address the large range of values. In (24) the maximum of the inverse eigenvalues of the FIM are used. In (25), the eigenvalues are normalized to the corresponding eigenvalue of the original test signal to consider the relative change of the eigenvalues. In (27), $g_{k,l}$ are the elements of the inverse FIM and n_{p} is the number of model parameters. The sensitivity sum is the sum of squared elements of the FIM. The choice of measure is a multi-objective optimization problem that is addressed by scalarization through weighted aggregation.

5.3 Optimization and Identification Procedures

The parameters can be separated into the local model $\boldsymbol{\Theta}_{\text{LM}}$ and membership function parameters $\boldsymbol{\Theta}_{\text{MF}}$. The sensitivities w.r.t. different parameter groups of a TS model are on different scales which was confirmed by preliminary experiments. Two approaches to rescale the sensitivities have been investigated. *First*, the sensitivities were normalized to a percentage change in the parameter. This approach did not yield any improvements on the optimal experiment design. *Second*, the derivatives have been rescaled with the determinants of the block matrices of local model and partition parameters. This also did not improve the optimal experiment design.

Therefore, another path has been chosen in this contribution. The optimization of the local model parameters and the partition parameters is separated which leads to smaller matrices. The FIM for the local model parameters $\boldsymbol{I}_{\text{LM}}$ is just composed of the "squared" extended regression matrix. The FIM for the partition parameters $\boldsymbol{I}_{\text{MF}}$ is more complicated:

$$\boldsymbol{I} = \frac{1}{\sigma^2} \begin{bmatrix} \boldsymbol{I}_{\text{LM}} & \boldsymbol{I}_{\text{coupl}} \\ \boldsymbol{I}_{\text{coupl}}^{\top} & \boldsymbol{I}_{\text{MF}} \end{bmatrix} \tag{28}$$

with

$$\boldsymbol{I}_{\text{LM}} = \sum_{k=1}^{N} \boldsymbol{\varphi}_{\text{E}} \boldsymbol{\varphi}_{\text{E}}^{\top} = \boldsymbol{\Phi}_{\text{E}}^{\top} \boldsymbol{\Phi}_{\text{E}} \tag{29}$$

$$I_{\mathrm{MF}} = \sum_{k=1}^{N} \sum_{i=1}^{c} \left(\frac{\partial \mu_i}{\partial \boldsymbol{\Theta}_{\mathrm{MF}}}\right) \left(\frac{\partial \mu_i}{\partial \boldsymbol{\Theta}_{\mathrm{MF}}}\right)^{\top} \hat{y}_i(k)^2 \tag{30}$$

Those matrices lack the couplings in the complete FIM which is subject to the investigation. It is not clear what constitutes a fair comparison in this type of investigation. The following procedures have been determined:

– P1: local model and membership function parameters
– P2: only local model parameters
– P3: only membership function parameters
– P4: merged test signal
– P5: sequential

The mentioned procedures are combinations of identification schemes, optimization schemes of the FIM as well as the construction of the complete test signal. The procedures P1 - P5 have been selected specifically to reduce the number of possible combinations in a way to enable a fair comparison. P1 is the baseline and denotes the optimization of the FIM (28) and the simultaneous estimation of all model parameters in the optimization step. In the other procedures the block structure of the FIM has been exploited. In P2 and P3, the optimization of the test signal has been conducted w.r.t. the local model parameters and the partition parameters, using the partial FIMs (29) and (30), respectively. Since the signals resulting from P2 and P3 are optimized for the identification of the respective parameter group only those parameters have been estimated during the nonlinear optimization step.

Procedures P4 and P5 use the FIM optimization scheme from P2 and P3 but each only with half-length signals. In P4 the two half-length signals are concatenated to obtain a full-length signal. Even for multisine signals, the introduced step at the merging point has no significant impact. This signal is used for a simultaneous estimation of all model parameters. In P5 the half-length signals from P4 are used to sequentially estimate the membership function parameters based on a half-length P3-signal first, and then local model parameters based on a half-length P2 signal. FIM-based test signal design (for nonlinear models) is localized in the region of the true model parameters $\boldsymbol{\Theta}_0$ that are substituted by the initial model parameters $\boldsymbol{\Theta}_{\mathrm{pmf}}$. Therefore, the nonlinear optimization based on optimal datasets is initialized with the initial model parameters. FCM partitioning and NARX local model parameters might result in an initialization that has no connection to the optimal design and therefore no conclusions about the parameter uncertainty reduction could be drawn.

6 Case Studies

In this section the results of two extensive and one compact case studies are presented. The presented procedures are applied to a simulation case study (SIM) as well as a servo-pneumatic longitudinal drive test stand (SPLD) and in parts to a 3-tank system (3TK). Multisine, multistep and a combination of both signal types are considered as test signals. Different measures on the FIM as well as different optimization procedures to exploit the TS model structure are investigated. The method is assessed mainly by the simulation error on validation datasets and secondarily by the model parameter uncertainty.

6.1 Experimental and Identification Setups

For the two extensive case studies, an artificial simulation system in the form of a difference equation and a laboratory test stand have been used. For a third compressed case study another laboratory test stand was used.

6.1.1 Narendra System (SIM)

The system

$$
\begin{aligned}
y(k) =& \frac{y(k-1) \cdot y(k-2) \cdot (y(k-1) + 2.5)}{1 + y(k-1)^2 + y(k-2)^2} \\
&+ u(k-1) + \epsilon(k)
\end{aligned}
\tag{31}
$$

was first presented in Narendra and Parthasarathy [3] and serves as an identification benchmark. It is defined by a nonlinear 2nd order difference equation in the output $y(k)$. $u(k)$ is the input and $\epsilon(k) \in \mathcal{N}(0, 0.5^2)$ is chosen to be an i.i.d. normally distributed random variable. The noise variance has been chosen as $\sigma^2 = 0.5^2$ to obtain a relatively low signal to noise ratio (SNR).

The hyperparameters c and v of the model are chosen to be $c = 5$ and $v = 1.1$ based on the work in Gringard et al. [33]. Since the true system is known and the assumption of fuzzy TS models is that the nonlinear behavior is captured by the partitioning, the dynamic orders of the input and output of the TS model are chosen to be the same as in (31): $n = 2$ and $m = 1$. From (31) we also understand that the nonlinear behavior depends on the first and second lag of the output. Therefore, the regression vector is chosen as:

$$\boldsymbol{\varphi}^{\top}(k) = \begin{bmatrix} 1 & y(k-1) & y(k-2) & u(k-1) \end{bmatrix} \tag{32}$$

With the knowledge about the system's nonlinearity, the scheduling variable z is chosen as:

$$z(k)^{\top} = \begin{bmatrix} y(k-1) & y(k-2) \end{bmatrix} \tag{33}$$

This model has 30 parameters in total.

6.1.2 Servo-Pneumatic Longitudinal Drive (SPLD)

The SPLD test stand is described in detail in [4]. Further details can be found in [34]. The system input is the voltage applied to the servo valve and the output is the longitudinal position of the piston rod. The nonlinear behavior is due to the friction. The friction is reduced by dithering. W.r.t. the test signal design, a maximum frequency of $f_{max} = 3\,\mathrm{Hz}$ for the multisine signals as well as the slowest time constant $T_{slow} = 0.28\,\mathrm{s}$ for the multistep signals have been determined in preliminary experiments. To achieve a trade-off between reactivation effects of the FBF and smoothing model behavior, a fuzziness parameter of $v = 1.3$ has been selected. Using cluster-validation and assessing prediction errors, $c = 6$ local models are chosen. The orders of the input and output dynamics have been selected as $n = 4$ and $m = 3$, respectively, based on initial experiments. The scheduling vector is chosen equal to the regression vector. This results in 90 model parameters in total.

6.1.3 Three-Tank System (3TK)

The three-tank system [35] consists of three tanks that are serially connected. The system input is the volumetric flow rate q fed into the first tank by a pump. Since the fluid dynamics are significantly slower than the pump dynamics, the latter is neglected. The measured system output is the level h_2 of a tank on the other end. The slowest time constant $T_{slow} = 900\,\mathrm{s}$ has been used to determine the minimum and maximum lengths of the steps for S1 test signal design. In preliminary experiments $c = 4$, $n = 3$, $m = 1$, and the fuzziness parameter $v = 1.3$. The scheduling variable has been determined to be the first 3 lags of the output variable, resulting in 32 model parameters in total.

6.2 Test Signal Design

Table 1 shows the experiment design variables and how they are applied to the two extensive case studies. The signal model parameters for the multistep signals are the step lengths and step amplitudes, for the multisine signals the parameters are the amplitudes and phases of frequencies on a preselected grid. In the SIM and SPLD case study some of the design variations differ. The results from the SPLD as well as preliminary studies on the SIM case study have shown that the combined test signal (S3) is not suitable for this task and has been removed from experiment design for the SIM. The sensitivity sum (M5) was not used for the SPLD because the PSO was not able to consistently find a minimum, due to the high number of parameters. Therefore, M5 was removed from experiment design in the SPLD case study. The scaled trace (M2) and the scaled

Table 1 Experiment design variations

Symbol	Description	SIM	SPLD
Test signals			
S1	Multisine	Yes	Yes
S2	Multistep	Yes	Yes
S3	Combination	No	Yes
FIM measures			
M1	Determinant	Yes	Yes
M2	Eigenvalues	No	Yes
M3	Scaled Trace	No	Yes
M4	Trace	Yes	Yes
M5	Sensitivity Sum	Yes	No
Optimization Procedures			
P1	All parameters	Yes	Yes
P2	LM parameters only	Yes	Yes
P3	MF parameters only	Yes	Yes
P4	Concatenated signal	Yes	Yes
P5	Seq. ID (MF→LM)	Yes	Yes

trace (M3) did not yield any significant difference on the SIM case study and therefore have been removed from the SIM case study. The eigenvalues (M2) have not been used for the SIM case study, because preliminary experiments as well as the results from the SPLD have shown that the measure has no impact. In the 3TK case study, one combination (S2, M1, P3) is used to design an optimal test signal to indicate that the proposed method can be transferred to different systems. The variations for the artificial test system (SIM) result in 30 (2 (Signals) × 3 (Measures) × 5 (Procedures)) optimal designs, the variations for the SPLD test stand result in 108 (3 (Signal) × 4 (Measures) × 9 (Procedures)) optimal designs. The number of data points of the signals for the SIM case study is around 5000, for the SPLD around 2500, and for the 3TK around 6000. However, the signal lengths of all signals within the case studies are comparable and will therefore not be discussed further. For each signal type and system an initial test signal has been designed using process model-free methods. Repeated experiments have been conducted to choose the best model w.r.t. simulation quality on the validation dataset (that is used for assessment of the optimized models). This model is used to initialize the model-based designs.

6.3 Conducting the Experiments

To statistically investigate the impact of experiment design choices on the reduction of the uncertainty of the estimated parameters and the simulation quality, large amounts of data were collected. The initial models were obtained using

standard test signal design methods. For the optimization, a particle swarm optimization (PSO) algorithm is used, because preliminary investigations have shown that gradient-based methods cannot handle the search space. If a solution candidate is found by the particleswarm, a gradient-based algorithm (fmincon) efficiently finalizes the optimization. The MATLAB function particleswarm was used with a swarm size of 100. The inertia range has been set to $\begin{bmatrix} 0.1 & 1.1 \end{bmatrix}$. Both adjustment weights were set to 1.49. The meaning of those parameters is described in [36]. From preliminary works it is known that the PSO itself is robust w.r.t. random swarm initializations as well as for model parameter uncertainties.

For each of the initial test signals, repeated experiments have been conducted. For the artificial system, $N_{SIM} = 25$ repeats for the multisine and multistep signals each have been conducted. For the SPLD, each initial test signal has been applied $N_{SPLD} = 20$ times. On the 3TK system each test signal has been applied $N_{3TK} = 10$ times. The best performing models were selected for further use in the model-based design process and as a baseline for the simulation quality. The optimization of a test signal has only been initialized with a model identified with a signal of the same type. From the repeated experiments, information about the parameter uncertainty can be extracted for the process model-free test signal design.

The calculation of the optimal test signals as well as the identification and model evaluations have been conducted on an `Intel(R) Xeon(R) CPU E3-1270 v5 @ 3.60GHz 3.60GHz` PC with `64.0 GB of RAM`.

6.4 Results

The data collected in the parameter case study is used to assess the simulation quality as well as the parameter estimation uncertainty. Since the experiments have been conducted with many repeating experiments, "a model" refers to all models that have been identified from a dataset generated with the same test signal. This way, "a model" is defined by the combination of the signal model (S), the FIM measure (M) as well as the optimization procedure (P). Fundamentally, the RMSE is used to assess the simulation quality:

$$J_{RMSE} = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (\hat{y}(k) - y(k))^2} \tag{34}$$

All models have been evaluated on several validation datasets. Since there is no normalization method that is invariant w.r.t. different signals, the change relative to the initial model's simulation quality is used. This enables a fair comparison over different validation signals. The parameter estimation uncertainty is assessed by comparing

the empirical covariance matrices based on optimal experiments to the empirical covariance matrices based on the respective process model-free experiment. As an additional assessment, the method is applied to a system with different dynamics and evaluated based on the simulation error as well as direct uncertainty analysis.

6.4.1 Simulation Quality

The simulation quality is assessed in two ways. The first criterion is the fraction of improvement by at least 25 % over the respective initial model's performance, which will be called "success rate". Figures 3, 4, 5, and 6 show the percentages of models reaching this threshold.

Figures 3 and 4 show the results for the Narendra system, Figs. 5 and 6 for the SPLD. The simulation error improvements are either grouped by the procedures or by the FIM measures.

Tables 2, 3, 4, 5 show the average improvements of the successful models regarding the first criterion. The numbers show to which percentage of the initial model's performance the optimal model's performance can be lowered by the method. Figure 7 shows an evaluation of several models on a validation datasets for the compact 3TK case study. The FIM-based model outperforms both the model based on a PRBS as well as the PMF-based model.

6.4.2 Parameter Uncertainty

The investigation of the parameter uncertainty yielded the result that there is no significant change of uncertainty w.r.t. different FIM measures. Therefore, the results have been aggregated. A threshold is used to determine the fraction of models that have reduced uncertainty of at least 25 %. The uncertainty is assessed the measures on the empirical covariance matrices of the optimal models as well as the initial models. It is presented for M4 in Table 6. The covariance matrices are calculated separately for each parameter group. Since P2 and P3 do not make changes to the partition parameters and the local model parameters, respectively, the corresponding fields are left out. P5 is disregarded in this analysis because it is just the sequential
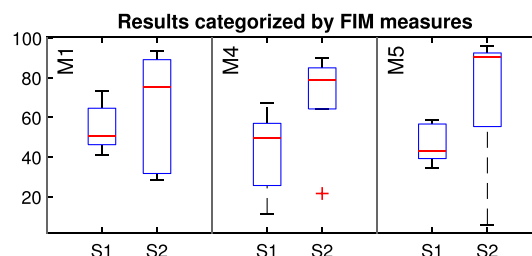

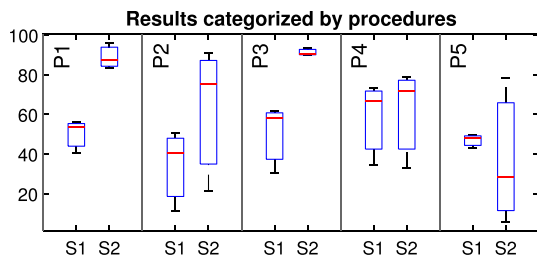**Fig. 3** Reduction of the RMSE in % by at least 25 % w.r.t. the initial model for different FIM measures (SIM)

Fig. 4 Reduction of the RMSE in % by at least 25 % w.r.t. the initial model for different procedures (SIM)
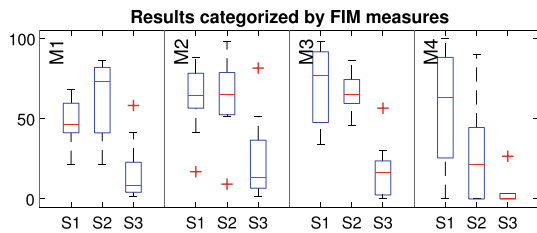

Fig. 5 Reduction of the RMSE in % by at least 25 % w.r.t. the initial model for different FIM measures (SPLD)
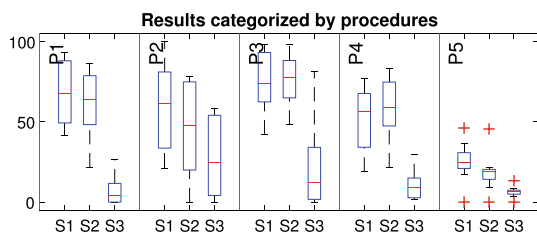

Fig. 6 Reduction of the RMSE in % by at least 25 % w.r.t. the initial model for different procedures (SPLD)

Table 2 Avg. Improvement by FIM measure in % (SIM)

	M1	M4	M5
S1	43.87	47.06	53.20
S2	42.28	44.94	40.49

Table 3 Avg. improvement by procedure in % (SIM)

	P1	P2	P3	P4	P5
S1	48.89	54.85	54.19	38.32	46.45
S2	37.71	45.57	41.04	46.01	44.10

Table 4 Avg. improvement by FIM measure in % (SPLD)

	M1	M2	M3	M4
S1	43.66	43.22	45.73	46.00
S2	37.64	38.04	36.50	45.20
S3	61.92	63.01	65.91	66.72

Table 5 Avg. improvement by procedure in % (SPLD)

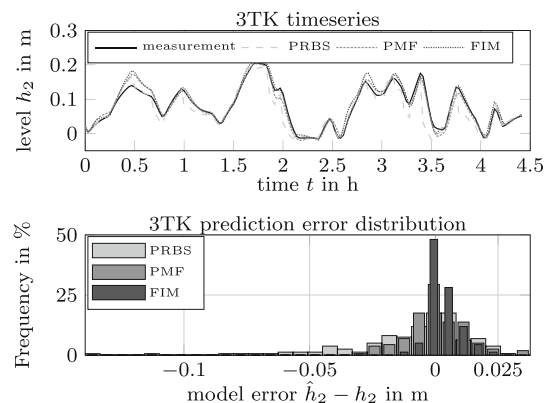|	P1	P2	P3	P4	P5
S1	44.31	48.19	40.28	47.37	48.38
S2	37.74	35.99	32.74	40.15	43.79
S3	64.97	63.11	61.23	67.48	66.60

Fig. 7 Simulation results on the 3TK system

6.5 Discussion

The choice of the signal model is significant for the resulting model quality, therefore it is important to characterize the type of excitation. The choice can have an impact in three ways. First, a signal model may not be suitable for identifying the plant in general. Second, using a plant model based on a signal type for optimal experiment design of another signal type yields worse results. Third, choosing an input signal type that fits the operating scenario is more likely to cover the relevant behavior. For the Narendra system, multistep signals performed superior, as can be seen in Figs. 3 and 4. S1 and S2 performed comparably for the SPLD, but multisine-based models do not perform well for the Narendra system. From Figs. 4 and 6 it follows that P1 and P3 measures perform very well both on the SPLD as well as the Narendra system, disregarding the overall mediocre performance of S1 models on that test system. P2 focusses on the local model parameters, resulting in a high variability of the partition parameters

identifications with P2 and P3 signals. To verify the results from the extensive case study, models based on different test signal designs have been analyzed. Table 7 shows the measure M1 of the models. FIM-based design yields more accurate parameters than identification based upon a PRBS as well as PMF design.

Table 6 Fraction of models in % with reduced uncertainty using M4 of at least 25 %

SIM		P1	P2	P3	P4
S1	LM	78.33	71.67	-	80.00
	MF	73.33	-	86.67	90.00
S2	LM	58.67	59.67	-	70.00
	MF	64.33	-	78.33	75.00
SPLD		P1	P2	P3	P4
S1	LM	83.86	62.50	-	75.53
	MF	86.91	-	89.29	85.12
S2	LM	78.65	51.57	-	81.25
	MF	70.24	-	85.12	73.81
S3	LM	69.28	49.48	-	51.57
	MF	89.89	-	69.65	83.93

Table 7 Uncertainty assessment of the 3TK with M1

Signal	Binary	PMF	FIM
J_{det}	92.0046	51.8493	28.3429

and therefore of the simulation results. The signal results in P4 is not optimal w.r.t. either parameter group or therefore resulting in mediocre simulation performance. The breakdown w.r.t. the measures in Figs. 3 and 5 implies that the choice of the measure has no significant impact on the results, which was verified by an analysis of the variance (ANOVA).

Table 6 exemplifies the results regarding the uncertainty. It can be observed that the goal of reducing the parameter uncertainty was achieved in the majority of the experiments. However, accurately estimated parameters do not necessarily result in a high simulation quality. Considering the results for the uncertainty reduction of P3 as well as the simulation results, focussing on the membership function parameters leads to better models.

There are several reasons why P3 yields the best results. The membership function parameters determine the nonlinear behavior of the TS model. The quality of the local model parameter estimation is limited by the estimation of the membership function parameters. If this is not considered during the design, local model parameters cannot be estimated accurately. The optimization problem becomes more complex using all model parameters, explaining the inferior performance of P1 compared to P3.

S3 underperforms in every case. The signals have been observed during optimization. During optimization all signal model parameters are treated equally, leading to "chaotic" behavior that is not present if the two signal are optimized individually. For the compact case study of the three-tank system (3TK) an optimal test signal was designed (S2, M1, P3). The uncertainty was reduced considering M1 as shown in Table 7. The assessment of the simulation quality based on the 3TK case study is exemplary shown in Fig. 7.

6.6 Conclusions

The targeted test signal design goal is the reduction of parameter estimation uncertainty as well as the improvement of the simulation quality. The design performance regarding both goals has been investigated. Table 6 shows that the first design goal is achieved, in principle. However, this does not necessarily lead to an improvement of the simulation quality. This is not a case of a bias-variance trade-off but rather that the partitioning conditions the estimation of the local model parameters. Even though local model parameters might be estimated accurately, a low confidence in the partitioning results in worse global simulation properties. Therefore, the design should emphasize the partition parameters. The estimation of the local model parameters is easier and their consideration during the optimal test signal design increases the complexity of the optimization problem in a way that is detrimental to the solution. The choice of the FIM measure does not significantly impact the results, therefore the measure should be chosen considering the computational complexity. Even though P3 is only optimal w.r.t. the partition parameters, it performs better than P1. This is an interesting result, since the computational costs reduce significantly, if the FIM measure is chosen correctly. FIM-based test signal design for TS models assumes that at least the model structure is roughly correct and that the partition parameters are already in a region around their "optimal value" or a computationally demanding robust design (such as sequential) must be applied. This again stresses the importance of model-free test signal design before attempting an optimal design. FIM-based optimal experiment design for dynamic TS models should include a well-designed preceding model-free design stage, consider the operating conditions, and focus on the estimation of partition parameters to achieve low uncertainty as well as high simulation quality models. In a previous study [33], the method has been applied to an electro-mechanical throttle. The results are comparable, even though less design choices were studied.

7 Summary and Outlook

In this contribution a novel method for FIM-based optimal experiment design for the identification of nonlinear dynamic Takagi-Sugeno multi-models was presented. The impact of design choices was analyzed on a simulated system as well as a laboratory test stand and tested on another laboratory test stand. It was demonstrated that the approach leads to improvements in parameter uncertainty, which was the stated goal, but also to significant improvements in simulation quality.

The presented method is a holistic approach, considering all aspects of FIM-based test signal design for the identification of nonlinear dynamic models combined and analyzing the impact of design choices, whereas many other contributions focus on a specific aspect and assume the choice of measure and signal model. It was shown that the design choices have a significant impact on the achievable results, especially exploitation of the unique multi model structure. It was also shown that different measures on the FIM do not result in significant changes either in uncertainty or simulation quality but that the computational cost can be reduced by choosing less complex measures.

References

1. Gevers, M.: Identification for control: from the early achievements to the revival of experiment design. Eur. J. Control **11**(4–5), 335–352 (1998)
2. Hussain, M.A.: Review of the applications of neural networks in chemical process control—simulation and online implementation. Artif. Intell. Eng. **13**(1), 55–68 (1999)
3. Narendra, K.S., Parthasarathy, K.: Identification and control of dynamical systems using neural networks. IEEE Trans. Neural Netw. **1**(1), 4–27 (1990)
4. Zaidi, S., Kroll, A.: On interval-valued-data type-1 Takagi-Sugeno fuzzy systems for uncertain nonlinear dynamic system identification. Automatisierungstechnik **64**(6), 418–427 (2016)
5. Atkinson, A., Donev, A.: Optimum Experimental Designs. Clarendon Press, Oxford (1992)
6. Bandemeer, H., Bellmann, A.: Statistical Experimental Designs (in German), 4th edn. Teubner, Stuttgart (1994)
7. Fedorov, V.: Theory of Optimal Experiments. Academic Press, New York (1972)
8. Godfrey, K.: Perturbation Signals for System Identification. Prentice Hall, Hoboken (1993)
9. Pintelon, R., Schoukens, J.: System Identification: A Frequency Domain Approach, 2nd edn. Wiley, New York (2012)
10. Ebert, T., Fischer, T., Belz, J., Heinz, T.O., Kampmann, G., Nelles, O.: Extended deterministic local search algorithm for maximin latin hypercube designs. In: Proc. 2015 IEEE Symposium Series on Computational Intelligence, pp. 375–382. Cape Town, South Africa (2015)
11. Gringard, M., Kroll, A.: On the systematic analysis of the impact of the parametrization of standard test signals. In: Proc. IEEE Symposium Series of Computational Intelligence 2016. Athens, Greece (2016)
12. Heinz, T.O., Ebert, T., Nelles, O.: Pole determination for Takagi-Sugeno fuzzy Laguerre models. In: International Conference on Fuzzy Systems: FUZZ-IEEE. New Orleans, USA (2019)
13. Kroll, A., Dürrbaum, A.: On joint optimal experiment design for identifying partition and local model parameters of Takagi-Sugeno models. In: Proc. IFAC Symposium on System Identification, pp. 1427–1432. Beijing, China (2015)
14. Dürrbaum, A., Kroll, A.: On robust experiment design for identifying locally affine takagi-sugeno models. In: IEEE International Conference on Systems, Man, and Cybernetics (SMC 2016). Budapest, Hungary (2016)
15. Prozato, L.: Optimal experiment design and some related control problems. Automatica **44**(2), 303–325 (2008)
16. Hametner, C., Stadlbauer, M., Deregnaucourt, M., Jakubek, S., Winsel, T.: Optimal experiment design based on local model networks and multiplayer perceptron networks. Eng. Appl. Artif. Intell. **26**(1), 251–261 (2013)
17. Heinz, T., Nelles, O.: Comparison of excitation signals for nonlinear identification tasks (in German). In: Proc. 26th Workshop Computational Intelligence, pp. 139–158. Dortmund, Germany (2016)
18. Larsson, C., Hjalmarsson, H.: On optimal input design for nonlinear FIR-type systems. In: Proc. 49th IEEE Conf. on Decision and Control. Atlanta, USA (2010)
19. Fang, K., Shenton, A.T.: Constrained optimal test signal design for improved prediction error. IEEE Trans. Autom. Sci. Eng. **11**(4), 1191–1202 (2014)
20. Mukkula, A.R.G., Paulen, R.: Optimal experiment design in nonlinear parameter estimation with exact confidence regions. J. Process Control **83**, 187–195 (2019)
21. Umenberger, J., Schön, T.: Nonlinear input design as optimal control of a Hamiltonian system. IEEE Control Syst. Lett. **4**(1), 85–90 (2020)
22. Neilsen, T.B., Transtrum, M.K., Van Komen, D.F., Knobles, D.P.: Optimal experimental design for machine learning using the Fisher information matrix. J. Acoust. Soc. Am. **144**(3), 1730–1730 (2018)
23. Pleşu, V., Şerban Agachi, P.: Optimal experimental design based on global sensitivity analysis **24**, 63–68 (2007)

24. Walz, O., Djelassi, H., Mitsos, A.: Optimal experimental design for optimal process design: a trilevel optimization formulation. AIChE J. **66**(1), e16788 (2020)

25. Wang, J., Xia, J., Shen, H., Xing, M., Park, J.: H_∞ synchronization for fuzzy markov jump chaotic systems with piecewise-constant transition probabilities subject to pdt switching rule. IEEE Trans. Fuzzy Syst. pp. 1–1 (2020)

26. Shen, H., Xing, M., Wu, Z.G., Xu, S., Cao, J.: Multiobjective fault-tolerant control for fuzzy switched systems with persistent dwell time and its application in electric circuits. IEEE Trans. Fuzzy Syst. **28**(10), 2335–2347 (2020)

27. Cock, A.D., Gevers, M., Schoukens, J.: D-optimal input design for nonlinear FIR-type systems: a dispersion-based approach. Computing Research Repository Arxiv: abs/1703.08401 (2017)

28. Deflorian, M., Zaglauer, S.: Design of experiments for nonlinear dynamic system identification. IFAC Proc. Vol. **44**(1), 13179–13184 (2011)

29. Ying, H.: Sufficient conditions on uniform approximation of multivariate functions by general Takagi-Sugeno fuzzy systems with linear rule consequent. IEEE Trans. Syst. Man Cybern. Part A **28**(4), 515–520 (1998)

30. Kahl, M., Kroll, A., Kästner, R., Sofsky, M.: Application of model selection methods for the identification of a dynamic boost pressure model. In: Proc. 17th IFAC Symposium on System Identification (SysID), pp. 829–834. Beijing, China (2015)

31. Juhász, S., Legány, C., Babos, A.: Cluster validity measurement techniques. In: Proc. International Conference of Artificial Intelligence, Knowledge Engineering and Data Bases, pp. 388–393. Madrid, Spain (2006)

32. Bezdek, J.C., Ehrlich, R., Full, W.: Fcm: the fuzzy c-means clustering algorithm. Comput. Geosci. **10**(2), 191–203 (1984)

33. Gringard, M., Kroll, A.: Optimal experiment design for identifying dynamical Takagi-Sugeno models with minimal parameter uncertainty. In: Proc. 18th IFAC Symposium on System Identification (SysID). Stockholm, Sweden (2018)

34. Gringard, M., Kroll, A.: On the homogenization of different test signal classes for the identification of dynamical Takagi-Sugeno models: Methods, signal properties and comparative analysis (in German). at - Automatisierungstechnik **67**(10) (2019)

35. Schrodt, A., Kroll, A.: On iterative closed-loop identification using affine takagi-sugeno models and controllers. Int. J. Fuzzy Syst. **19**(6), 1978–1988 (2017)

36. MathWorks: Optimization Toolbox User's Guide (2019)

Matthias Himmelsbach was a research associate in the measurement and control department at the University of Kassel. His research focus is test signal design for nonlinear dynamic systems.

Andreas Kroll is head of the measurement and control department at the University of Kassel. His main research areas are non-linear identification and control methods, computational intelligence, telemetry technology and sensor data fusion.